



University of Bradford eThesis

This thesis is hosted in [Bradford Scholars](#) – The University of Bradford Open Access repository. Visit the repository for full metadata or to contact the repository team



© University of Bradford. This work is licenced for reuse under a [Creative Commons Licence](#).

SPATIAL ASSOCIATION IN ARCHAEOLOGY

Development of statistical methodologies and computer techniques for spatial association of surface, lattice and point processes, applied to prehistoric evidence in North Yorkshire and to the Heslerton Romano-British site.

A Thesis submitted for the Degree of Doctor of Philosophy.

By

Michael Anthony

~~M.A.~~ KELLY B.Tech.(Hons)

21079044-

Postgraduate School of Studies in Physics,
University of Bradford.

March 1986

Spatial Association in Archaeology

M.A. Kelly

The thesis investigates the concepts of archaeological spatial association within the context of both site and regional data sets.

The techniques of geophysical surveying, surface distribution collection and aerial photography are described and discussed. Several new developments of technique are presented as well as a detailed discussion of the problems of data presentation and analysis.

The quantitative relationships between these data sets are explored by modelling them as operands and describing association in terms of operators. Both local and global measures of association are considered with a discussion as to their relative merits.

Methods for the spatial association of regional lattice and point processes are developed. A detailed discussion of distance based spatial analysis techniques is presented.

101728

BD 034561690 2



CONTENTS

	Page
Acknowledgements	7
<u>PART 1</u>	8
1.0 Introduction	9
1.1 The importance of spatial association in archaeology	9
1.2 Objectives of the thesis	12
1.3 Structure of the thesis	15
2.0 Physical background	16
2.1 Geomorphology of North Yorkshire	16
2.2 Geomorphology of Heslerton	19
3.0 Archaeological background	22
3.1 Archaeology of North Yorkshire	22
3.2 Archaeology of Heslerton	24
<u>PART 2</u>	26
4.0 Data presentation	27
4.1 Data structure	27

4.2	Methods of illustrating structured data	29
4.3	Contouring methods	33
4.4	BAGS	43
4.5	Discussion	43
5.0	Geophysical surveying	48
5.1	Objectives of geophysical surveying	48
5.2	Methods of geophysical surveying	49
5.3	Conventional resistance surveying	51
5.4	Automatic logging for geophysical surveying	54
5.5	A microcomputer system for resistance surveying	55
5.5.1	Hardware	55
5.5.2	Software	58
5.5.3	Field operation	59
5.6	The Heselton geophysical survey	61
5.7	Discussion	65
6.0	Surface distributions	70
6.1	Cause and effect	70
6.2	Methods of surface collection	75
6.3	The Heselton surface collection scheme	79
6.4	Heselton surface collection results	80
6.5	Discussion	86
7.0	Aerial photography	90
7.1	Aerial photography in archaeology	90

7.2 Problems in aerial photography	91
7.3 Solutions to problems in aerial photography at Heslerton	92
7.4 Discussion	97
<u>PART 3</u>	99
8.0 Site association operands	100
8.1 Operands and operators	100
8.2 Transforms of operands	101
8.3 Inverse transforms of operands	112
8.4 Measuring the archaeological significance of operands	117
8.5 Discussion	131
9.0 Site association operators	133
9.1 Types of operators	133
9.2 Global operators	134
9.2.1 Bivariate correlation	136
9.2.2 Multivariate correlation	140
9.2.3 Bivariate regression	142
9.2.4 Multivariate regression	145
9.3 Local operators	147
9.3.1 Similarity maps	149
9.3.2 Local association maps	155
9.4 Discussion	159

10.0 Regional association for lattice and point processes	168
10.1 Lattice processes	168
10.1.1 Methods of contouring scattered data	169
10.1.2 Association of lattice processes	172
10.2 Point processes	174
10.3 Classical K th nearest neighbour analysis	175
10.4 Problems with the classical analysis	178
10.5 A new correction for K th order neighbours	182
10.5.1 Simulation experiments	183
10.5.2 Mean K th nearest neighbour distance correction	185
10.5.3 Sampling variance correction	190
10.5.4 Comparison	197
10.6 Association of point processes	200
10.7 Discussion	207
<u>PART 4</u>	209
11.0 Conclusion	210
11.1 Summary	210
11.2 Discussion	214
Appendix A Software outlines	218
A.1 Program outline for contour_plot	219
A.2 System outline for BAGS	221
A.3 Program outline for RSCS	224

A.4 Program outline for apgrid	228
A.5 Program outline for f_extract	229
A.6 Program outline for local_a_map	230
A.7 Program outline for k_nna	231
Appendix B BAGS user manual	233
Appendix C RSCS user manual	256
Appendix D Standard point data file	291
Bibliography	296

ACKNOWLEDGEMENTS

Without the help of many people this thesis could not have been under taken. In particular thanks must go to J. Haigh, my supervisor, who provided objective support throughout my research. D. Powlesland, the brightest dressed archaeologist in Britain, for his help in the organisation of fieldwork at Heslerton. P. Dale for his technical expertise. A. Spencer for her invaluable help and support. And finally the Cyber 170-720 mainframe, at the University of Bradford Computer Centre, for its uncomplaining acceptance of the following text.

PART 1

OF SPATIAL ASSOCIATION IN ARCHAEOLOGY

The study of archaeology could be described as the study of the present through the past. Philosophers might argue that the history of archaeology is in fact the history of contemporary society. The early antiquarians of the nineteenth century were mainly clergymen and landed gentry, whose attitudes towards archaeology reflected their Victorian interests in philanthropy and self improvement. These antiquarians were mainly concerned with the collection of artifacts with little interest in their context. It was not until the latter part of the nineteenth century that the importance of relative chronology and hence stratigraphy became generally accepted. This concept dominated archaeological thinking well into the twentieth century, with emphasis being placed on typologies and chronological schemes. The importance of horizontal relationships only became widely acknowledged by the middle of this century. Up to this point archaeology was in general characterised by prominent individuals (successors to the Victorian antiquarians) rather than prominent ideas. The social changes of the 1960's however, produced drastic changes in archaeology. Protests about the accelerating destruction of finite archaeological resources were voiced and an effective organisation for management was started. Academically the 'New Archaeology' appeared with an emphasis on theory and methodology.

Today the 'New Archaeology' has not fulfilled all its early

promise. Pioneer works, such as Clarke's Analytical Archaeology (1968), have only proved useful in stimulating interest in areas outside traditional archaeology. This interest has developed into a move towards a 'Theoretical Archaeology'. From the mid 1970's onwards an important influence in this development has been information technology. This has developed from a novel academic application to a basic archaeological tool. By the end of this century computer technology has the potential to be the fundamental basis of all professional archaeology. This development is a simple reflection of the accelerating changes in our present society.

The concept of spatial relationships is now recognised as a basic parameter of archaeological methodology. Archaeological excavation recovers both physical evidence and more importantly its spatial context. The spatial relationships that are defined fossilise the processes operational before deposition and also incorporate any post-depositional transformations. Formal spatial analysis uses statistics and mathematics to describe and analyse in simple terms the patterning of spatial distributions. This formal analysis is usually uniquely identified with intra-site distribution studies. Spatial information is however more general, ranging from cultural distributions to elemental distributions on artifacts. The techniques available are also more widespread than often recognised. Traditional spatial analysis (Hodder and Orton, 1976) is only one aspect of spatial problems in archaeology.

An important area within spatial archaeology is that of association. This has unfortunately been confined to traditional analyses with the result that there has been little development of the concept in archaeology. Association is a powerful conceptual

tool in that it is normally very invariant to transformations and disturbances. For example the spatial distribution of material on an archaeological site may be substantially disturbed by chemical, physical and biological agents. The original spatial structure will be obscured, possibly destroyed. In contrast the spatial relationships of association between various categories of material will often be changed very little. Thus although the global structure is lost, the local infra-structure is often preserved.

The problems inherent in association analysis are manifold. The first problem is to decide what to measure association between. The choice of data must obviously be archaeologically meaningful, yet very little work has been attempted to define association between different types of data. A site may be surveyed by aerial photography, geophysics and surface collection, but no attempt is made objectively to analyse the spatial relationships between these data types. The resulting extracted information is therefore substantially reduced as can be demonstrated by Boole's inequality (Mood, Graybill and Boes, 1974). In association terms this indicates that it is more efficient and profitable to extract information from a set of data types treated as a group rather than treating each data type individually.

The second problem is to define a measure of association. Two basic approaches are possible here, global and local analysis. A global measure will simply characterise the spatial relationships of the data sets. This only becomes useful if the global approach is used at a synthesis level, allowing large scale comparison of large numbers of relationships. A local measure will provide much more information about the structure within and between the data sets.

This is much more useful archaeologically at an interpretive level.

The third problem is to represent the results of association analysis. For global analysis this is relatively straightforward, a series of statistics or simple graphs will often be sufficient. Local analysis presents more difficult problems in that a lot of the detailed spatial information is preserved and must be displayed. This necessitates the use of complex two dimensional representations of analyses.

The fourth problem is the archaeological interpretation of the results. By necessity the methodologies used must be based on statistical and mathematical models, but these models must relate to archaeological premises. This logically leads back to the first problem of data definition. Thus the individual problems cannot be isolated but must be studied as a whole.

1.2 OBJECTIVES OF THE THESIS

The main objective of this thesis is to study the problems of spatial association outlined above, in the context of two major application areas in archaeology. The first of these is site survey. This includes all archaeological techniques that aim to map the extent and characteristics of archaeological information without excavation. Traditionally such techniques have been used to locate sites for subsequent excavation. However the increasing sophistication of these techniques and the accelerating expense of excavation has led many archaeologists to use site survey as a primary data collection tool rather than just for simple site prospecting. The first half of this thesis will therefore

concentrate on survey techniques; principally geophysical surveying, surface collection and aerial photography.

To give practical examples of the use of these techniques a study area at Heslerton, North Yorkshire (figure 1.1) has been used. This lies within the Heslerton Parish Project (Powlesland, 1982) which is a long term project aimed at studying the interaction of man and the Yorkshire landscape from the Mesolithic to Medieval periods. The study area of one square hectare lies on the edge of a Romano-British ladder settlement and encloses areas of both occupation and cultivation as indicated by aerial photography. This study area was chosen for two reasons. Firstly, planned future excavations will allow detailed validation of survey results. Secondly, the responses of the various survey techniques will provide useful information on the future applicability of these techniques to larger areas of the project.

The second half of this thesis will develop methods of association for survey data as well as for the second major application area in archaeology, that of regional distribution analysis. The spatial interpretation of regional distribution maps is important as:

'The distribution map lies behind some of the most central themes in archaeology such as trade, diffusion and culture.',
(Hodder and Orton, 1976).

Using both point and lattice type data from North Yorkshire's archaeological record as illustrative material, the problems and methods of spatial association at a regional level will be studied.

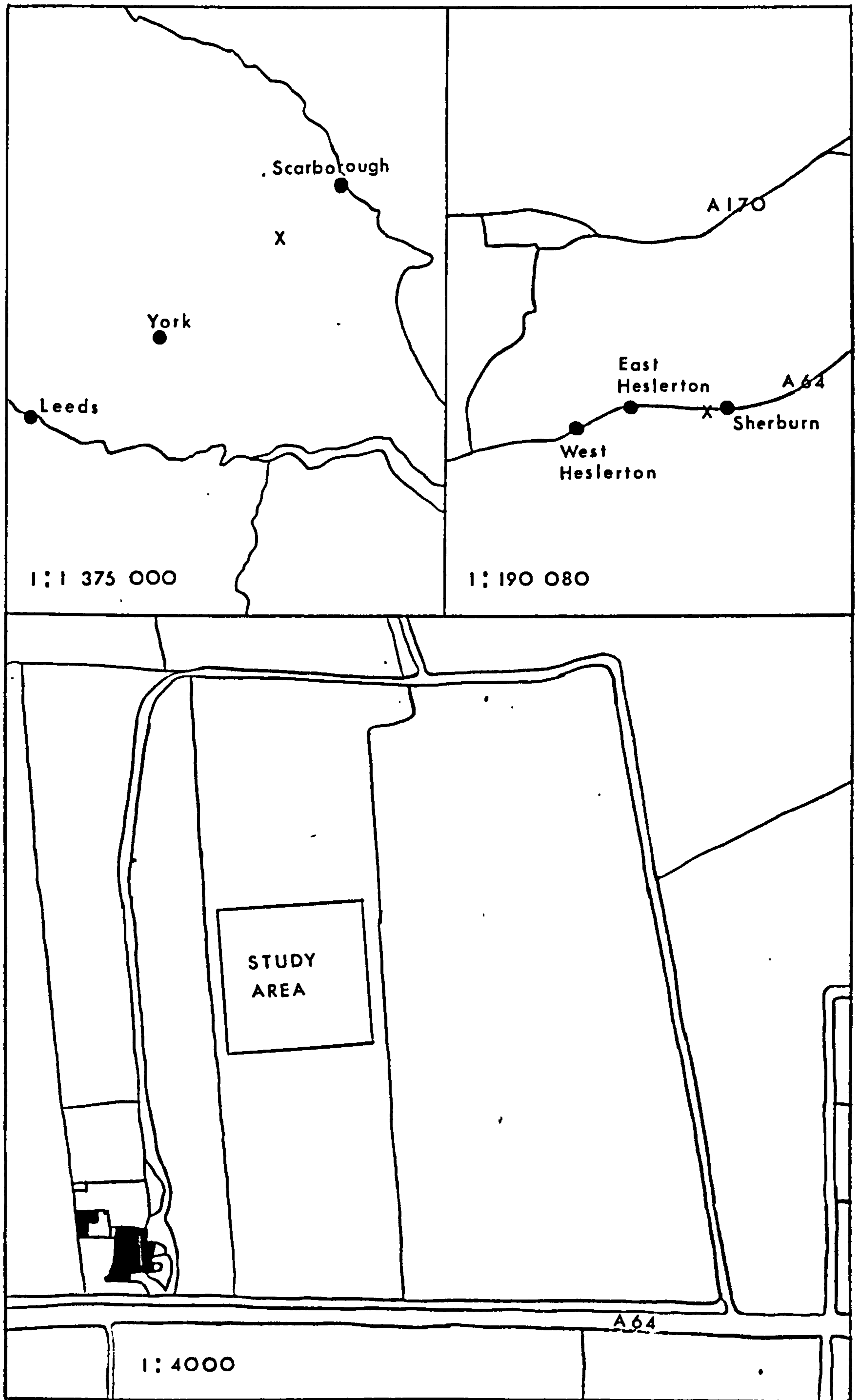


Figure 1.1 Heslerton study area.

1.3 STRUCTURE OF THE THESIS

The thesis is divided into four parts. Part one acts as an introduction. As well as outlining the aims of the thesis a summary of both the physical and archaeological background to the Heslerton study area and North Yorkshire in general is given.

Part two presents the survey techniques. It begins by looking at the methods of presenting survey data. It then considers in turn the problems and methods of geophysical surveying, surface distribution collection and aerial photography.

Part three contains the development of association methods. These are first applied to the survey techniques of part two through the concepts of site operands and operators. Regional association for lattice and point processes is then considered.

Finally, part four presents a summary and discussion of the thesis. Appendices containing the outlines of software used in the thesis and copies of associated user manuals are also included.

2.0 PHYSICAL BACKGROUND

2.1 GEOMORPHOLOGY OF NORTH YORKSHIRE

North Yorkshire can be divided into several distinct geomorphological areas as indicated in figure 2.1 (see Bastwood (1953); Trueman (1967)). The first of these is the North Yorkshire Moors which is an upland area of 30 miles by 15 miles extent. Much of this heather moorland is over 330 metres high and is cut by deep steep sided dales. Most of these dales run north to south with Newtondale, formed by glacial meltwater, being the only continuous channel across the moors. Geologically the North Yorkshire Moors are formed from a variety of rocks. Deltaic sandstones predominate with shales, ironstones, grits and limestones also present. The hard shales to the west of the moors produce a sharply rising scarp in the Hambleton Hills. While the ironstone bands of the Middle Lias to the north of the moors in the Cleveland Hills, provide ores for the Tees valley.

To the south of the North Yorkshire Moors lies the Vale of Pickering. The morphology of this flat bottomed vale has been primarily determined by glaciation. Mounds of glacial deposit near the coast indicate the location of an ice dam at the eastern end of the vale. This led to the development of a wide lake which eventually drained through a spillway to the west of the vale, now seen as a deep gorge in the Howardian Hills through which the river Derwent flows. This eventual drainage left extensive alluvial deposits covering the wide outcrop of Kimmeridge clay which

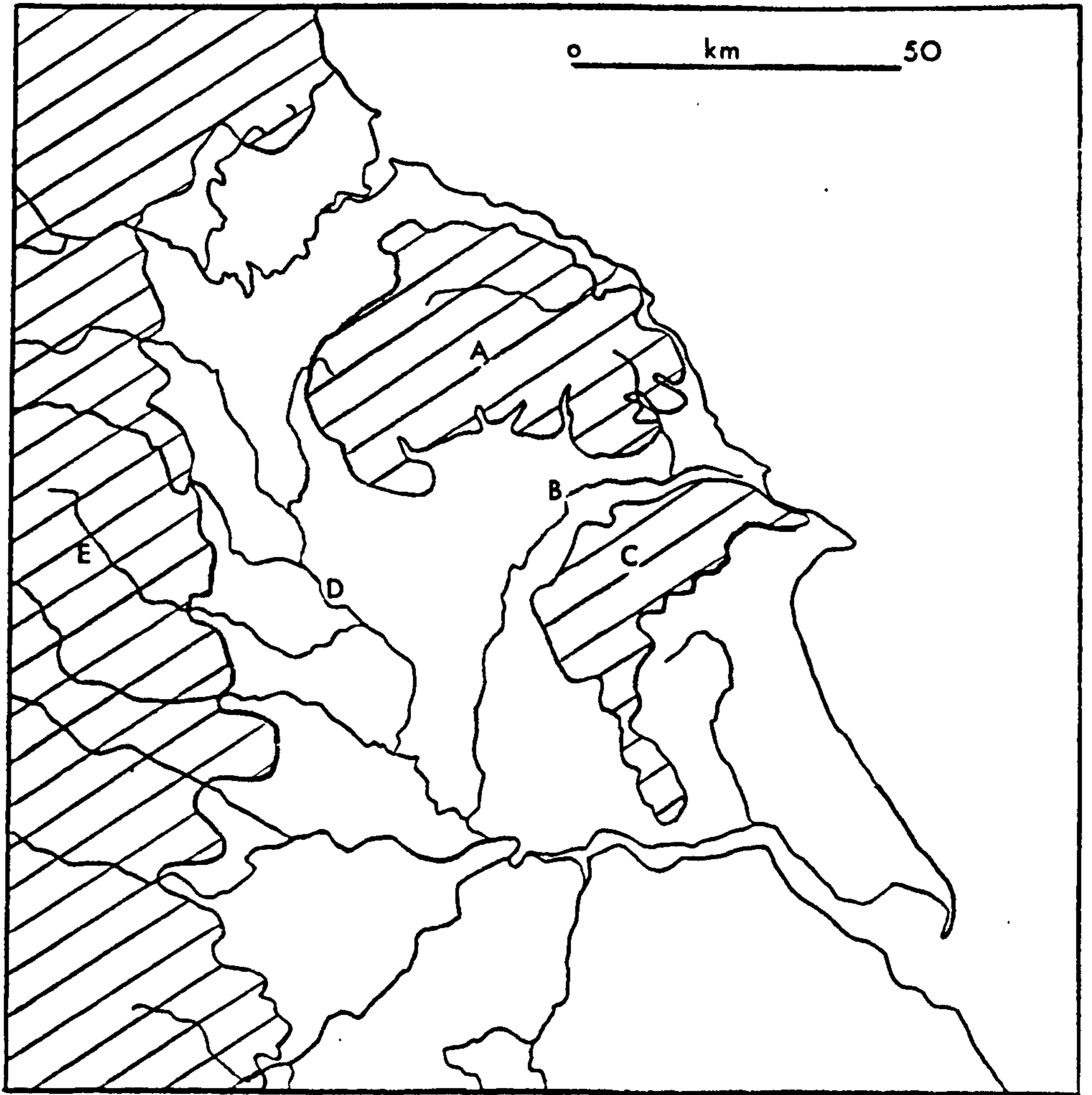


Figure 2.1 Geomorphological areas of North Yorkshire.
A - North Yorkshire moors, B - Vale of Pickering,
C - Yorkshire Wolds, D - Vale of York, E - Pennines.

underlies the vale.

To the south of the Vale of Pickering lies the crescent shaped Yorkshire Wolds. This chalk downland reaches heights of 240 metres with steep escarpments to the west and north. The chalk is much harder than that found in South England with many beds resembling limestone. This has produced many deep waterless valleys although the Great Wold Valley, which cuts across the wolds, does contain a stream. The whole area is extensively cultivated which presents severe plough damage problems to the many archaeological sites found in this area (see section 3.1).

In the west of North Yorkshire lies the mountainous country of the Pennines. Structurally this area can be split into the Northern Pennines and the Southern Pennines with the boundary near Skipton. The Southern Pennines are further subdivided into the Central Pennines and the Peak District while the Northern Pennines consist of the Askrigg block and the Alston block divided by the Stainmore Gap. Geologically the Pennines are dominated by Carboniferous limestones, Millstone Grit and Coal Measures. This has led to a varied morphology ranging from the dales cut by east-flowing streams to mountain limestone country. Many of the drier areas support heather moors while the wetter land is often covered by peaty bogs or mosses.

Separating the Pennines from the North Yorkshire Moors and the Yorkshire Wolds is the Vale of York. This broad lowland of low rolling hills and scarps is geologically younger than the Coal Measures of the Pennines. Extensive flooding has left rich soils for cultivation. In former times this flooding caused the Vale of

York to be a marshy barrier between the Pennines and East Yorkshire. Fortunately the York and Escrick terminal moraines provided a natural elevated highway which allowed easy passage across the Vale of York.

2.2 GEOMORPHOLOGY OF HESLERTON

The parish of Heselton covers an area of approximately 3000 hectares. This is roughly divided into equal areas of the Yorkshire Wolds and the Vale of Pickering. However six distinct geomorphological areas are evident within the parish on careful examination (Powlesland, 1982).

The wold top (area 1 of figure 2.2) is the largest geomorphological zone and consists of an area of open downland. It is situated above the 150 metre contour and contains dry valleys with deposits of alluvial sands and gravels. Extensive cultivation has eroded many archaeological features as indicated by levelled crop mark sites.

The wold scarp (area 2 in figure 2.2) lying between the 150 metre and 90 metre contours is characterised by the steep north facing slope of the wolds. This has restricted the formation of good soils on the chalk and thus severely limits cultivation.

The wold foot geomorphological area (area 3 in figure 2.2) lies above the 50 metre contour. It includes both basal red chalk and Speeton clay deposits, the junction of which provides a spring line. The present day villages of East and West Heselton are situated along this spring line.

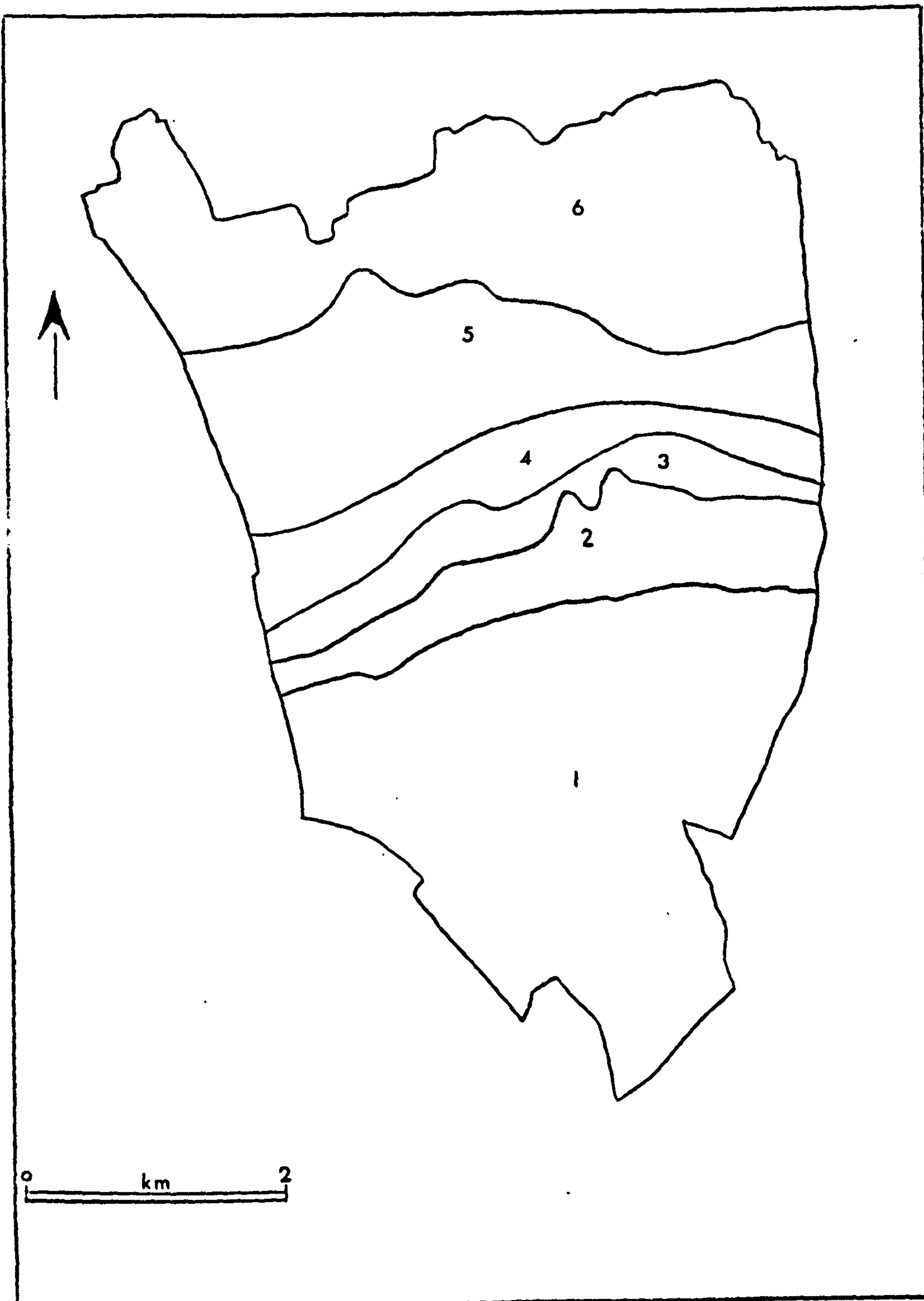


Figure 2.2 Geomorphological areas of Heslerton.

Below the wold foot lies an area of aeolian deposits (area 4 in figure 2.2). Blown sands, derived from an extensive deposit of post glacial sands and gravels beneath and to the north, have sealed old ground surfaces with up to three metres of overburden.

The deposit of post glacial sand and gravel constitutes the dry vale geomorphological area (area 5 in figure 2.2). The light soils are extensively cultivated and provide suitable conditions for aerial photography. However modern agricultural methods threaten any identifiable crop mark sites. The Heslerton study area used in the thesis is situated in this geomorphological area.

The wet vale is the final geomorphological area (area 6 in figure 2.2). This consists of a deposit of lacustrine clays with extensive overlaying peat deposits. Within this marshy area a series of low islands may be identified. In other parts of the Vale of Pickering these islands have been used for settlement.

3.0 ARCHAEOLOGICAL BACKGROUND

3.1 ARCHAEOLOGY OF NORTH YORKSHIRE

The Mesolithic period in North Yorkshire is mainly represented by microliths found on the North Yorkshire Moors and the Pennines. There has also been several excavations of occupation areas, the most important being Star Carr in the Vale of Pickering. This eighth millennium B.C. site has provided invaluable environmental and organic artifactual evidence for the Mesolithic period in Britain.

The Neolithic period is characterised by funerary and ritualistic monuments. These long barrows, henges and cursuses are concentrated on the Yorkshire Wolds but extend north into the North Yorkshire Moors and south into Lincolnshire. The association of Neolithic sites with light easily cultivated soils emphasises the importance of agriculture in this period. The visible presence of upstanding Neolithic earthworks on the wold top has also meant that historically, archaeological interest has tended to concentrate on the Yorkshire wolds, much to the detriment of other potentially important archaeological areas within North Yorkshire.

From the Beaker period onwards there is increasing evidence of occupation in lower-lying areas. Many ring ditches in valley positions are visible on aerial photographs and field surveys. It has been suggested (Dunn, 1976) that many of these features could be associated with the Beaker period.

From the Early Bronze Age to the Middle Bronze Age settlement evidence is rarely found. The most widespread remains of this period are round barrows. In the Late Bronze Age however cremations appear and finds of metal work are common with a marked concentration in Holderness. Settlement sites are now evident. Open settlements in the form of lake villages are found in Holderness (Varley, 1968) and hillforts consisting of palisaded enclosures on steep sided knolls such as at Staple Howe, Devils Hill and Grimthorpe (Stead, 1968) are also found. The Late Bronze Age/Early Iron Age therefore presents a picture of more widely distributed settlement. This probably reflects a population expansion, possibly causing strains in society and thus leading to the building of the hillforts.

From the seventh century B.C. Hallstatt C influences are evident on sites around the Vale of Pickering as indicated by finds of metal work and pottery. By the mid fifth century B.C. inhumations under a small barrow surrounded by a square ditch appear. These are the principle characteristic of the Arras culture. Chariot burials are also associated with this period and there is substantial evidence for arable farming.

Late Iron Age traditions seem to continue into the Roman period especially in rural areas such as the Vale of Pickering (Ramm, 1978). The main Roman influences are concentrated at urban and military centres, such as York, and on the prime agricultural land as indicated by villas. Historically the native population of the eastern side of North Yorkshire belong to the Parisi.

3.2 ARCHAEOLOGY OF HESLERTON

The use of excavation, field walking and aerial photography within the parish project, as well as previous archaeological work, has allowed a comprehensive survey of the archaeology of Heslerton. Although not complete, and lacking in detail in places, it provides a good summary of the range and extent of archaeological information within the parish.

The Mesolithic period is represented by microliths, indicating hunting activities, and quantities of flint working debris. Although the latter indicates the manufacture of microliths, no evidence of actual occupation has yet been found.

There is evidence for Late Neolithic occupation in the form of rubbish pits and associated field systems. Finds include flint implements and pottery, mostly Peterborough/Grooved ware.

The Early Bronze Age presents possible occupation evidence in the form of a ring gully that could be associated with a round house. There are also two independent barrow cemeteries dated to this period, that contain inhumations and cremations with Beakers and Food Vessels.

The Middle Bronze Age is represented only by finds of dung tempered pottery. However there is an open settlement with associated field systems in the Late Bronze Age/Early Iron Age.

The Roman period is dominated by a continuous series of ladder settlements and field systems which can be traced for about 15 km along the southern edge of the vale. These are dated to the third

and fourth centuries A.D. The Heslerton study area lies across one of these settlements.

There is an important Anglian inhumation cemetery which contains a rich variety of grave goods. There are also areas of Anglian plough marks and field boundaries but no evidence as yet of actual occupation.

The later medieval period saw the development of the present day settlement pattern. Strip parishes rising from the vale up into the wolds were formed, with the villages such as East Heslerton, West Heslerton and Sherburn sited on the spring line just below the wold scarp (see section 2.2).

PART 2

4.0 DATA PRESENTATION

4.1 DATA STRUCTURE

All data have two characteristics of type and structure. Type is defined by the measurement system or collection scheme with which data are recorded. Structure defines how different data types are related; for example, simple data types such as height, width and weight can be structured in a multidimensional scaling to form a single new data structure that characterises artifact class (Hodson, Sneath and Doran, 1966).

In archaeology most of the data types used are structured. This structure can be complex and needs to be explicitly defined if statistical and mathematical analytical techniques are to be applied. However any definition will be subjective so that, strictly speaking, only a conceptual data structure can be defined.

In the Heslerton study area the data sets involved; geophysical surveying, surface distributions and aerial photography, have complex conceptual structures. If these data sets are to be collected in a reasonable manner and analysed, then these conceptual structures must be simplified.

All the data sets are collected from a three dimensional space, as the surface of a field is not exactly planar, but in most cases it is an acceptable simplification to assume that collection is from a two dimensional plane. Therefore a collection scheme can assume that each data set is collected from a plane. Furthermore the set

of collection planes can be assumed to be coincident to form a single survey plane (this is the fundamental assumption needed for association analysis in part 3). This plane is limited or bounded by the survey area, at Heslerton a one hectare square. It has an alignment, an origin and is assumed to lie parallel to the overall ground slope. The data sets can be viewed as functions defined on the survey plane:

$f(x,y)$ where x and y are coordinates in the survey plane.

It is impossible to collect function values for every point in the survey plane. A sensible collection scheme must involve sampling. As the spatial context of the data is important the most appropriate sampling strategy is that of systematic sampling (Plog, 1978). It is noteworthy that geophysical surveying traditionally uses this sampling technique. The collection scheme is therefore defined as a systematic sampling of function values within the survey plane.

Associated with the collection scheme are several sources of possible error:

- 1) plane approximation for each data set
- 2) coincidence of planes
- 3) defining the sampling point on the ground
- 4) locating the collection point onto the sampling point
- 5) collecting the data
- 6) transcribing the data to storage

It is impossible to ensure that errors do not accrue; if they are not to be ignored the effect of errors must be allowed for, in one

of two ways:

1) Error estimates are made in the field, and are incorporated in all subsequent analyses.

2) Tolerance levels are predefined for fieldwork. Quality control is applied to keep fieldwork within these limits. The effects of these tolerances on all subsequent analyses are estimated.

Option one has the advantage of allowing flexible fieldwork. It has the disadvantage that allowance for errors in analyses may be complex. Option two has the advantage of fixed fieldwork standards and relatively easy incorporation of errors into analyses. It has the disadvantage that fieldwork of greater precision than the tolerance levels will not be distinguishable as such.

For the Heslerton study area option two is used for two reasons. Firstly fixed fieldwork standards offer a more rigorous and efficient approach to archaeological fieldwork. They also strengthen the practicality of compatibility or portability of data between different studies. Secondly the use of fixed tolerance levels allows a greater flexibility in analyses than the use of variable error estimates.

4.2 METHODS OF ILLUSTRATING STRUCTURED DATA

The methods described in this section relate to the conceptual structure defined in the previous section i.e. a function defined on a plane which is systematically sampled. Such data can be illustrated or displayed in two basic forms; three dimensional and

two dimensional displays.

In a three dimensional display (figure 4.1) the data is represented by points whose x and y coordinates specify the position in the plane and whose z coordinate is proportional to the value of the reading. The data points are joined by line segments (or possibly curve segments) to produce a surface image. This sort of display looks impressive but it is difficult to extract useful information from it. Parts of the surface may be hidden or obscured by other parts, the whole perception of the surface is distorted by the perspective viewpoint which must be used to compress the display into two dimensions.

In a two dimensional display (figure 4.2) the data retain their x and y coordinates but the value of the reading is shown by an intensity. In one form of display (figure 4.2.a) each data point is assigned a surrounding uniform area or pixel, the pixels for neighbouring data points being contiguous. The z coordinate of a data point is represented by an intensity for its associated pixel. The intensity can be represented in various ways; different colours, different shades of grey (grey scales), different density of symbols (dot density). The boundaries between pixels can be well defined to emphasis the sampling ancestry of the data, or the boundaries can be smeared to give a smooth display that visually indicates the function sampled.

In another form of display (figure 4.2.b) each reading is retained as a point. The sampled function is numerically estimated from the data by fitting either a global function to the data points (as in trend surface analysis) or a series of contiguous local

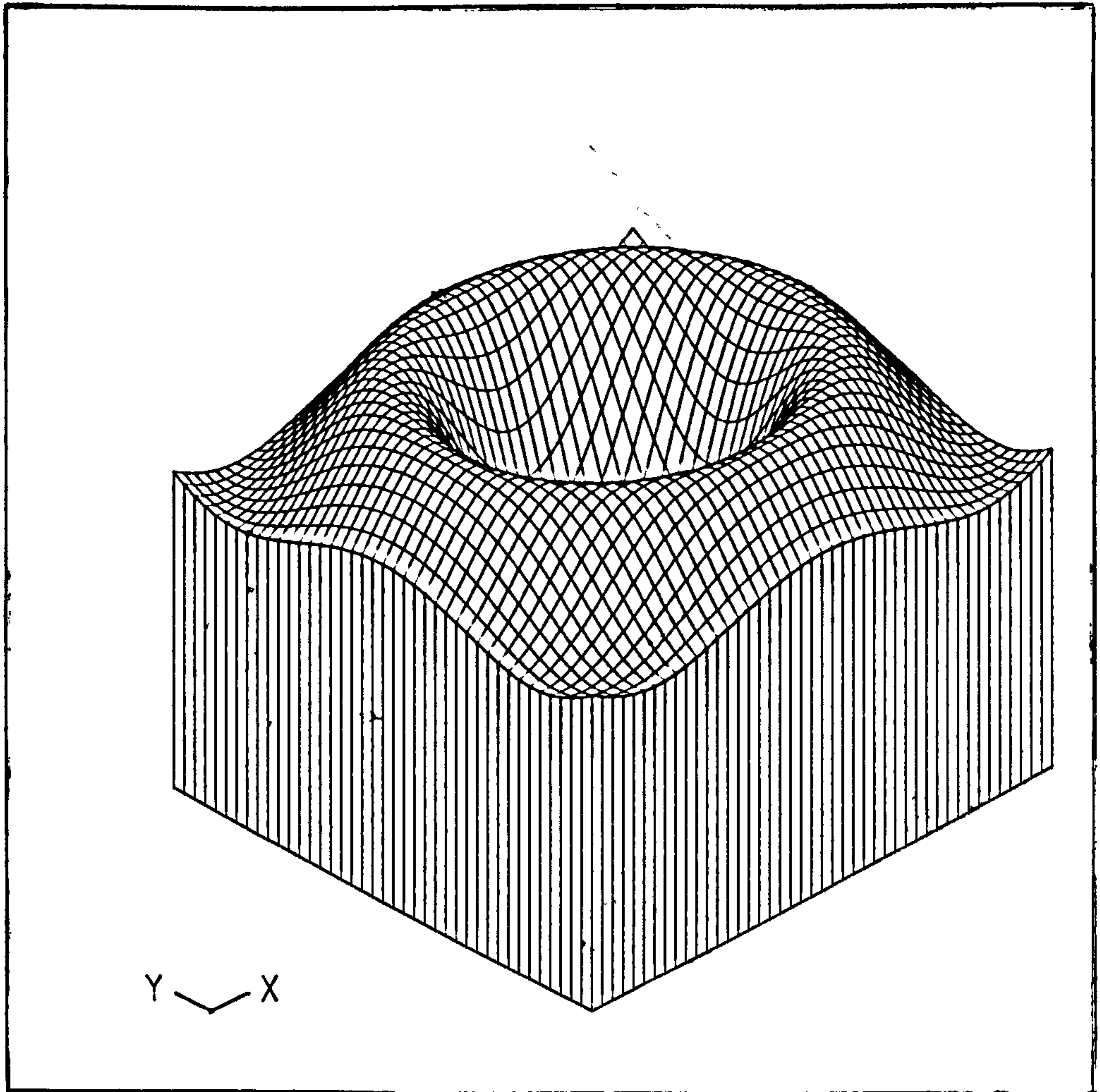


Figure 4.1 Three dimensional display of a function defined on a plane.

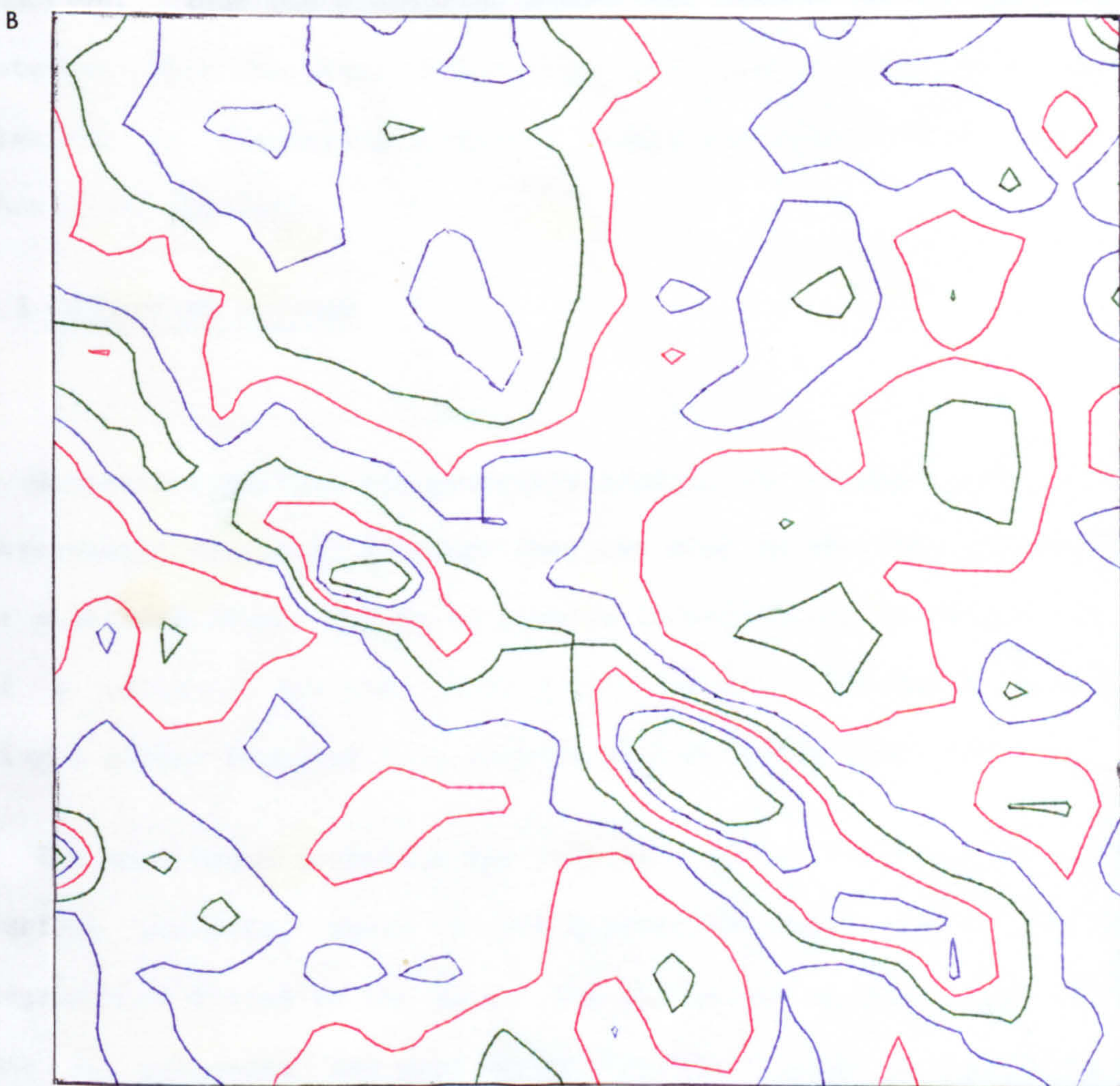
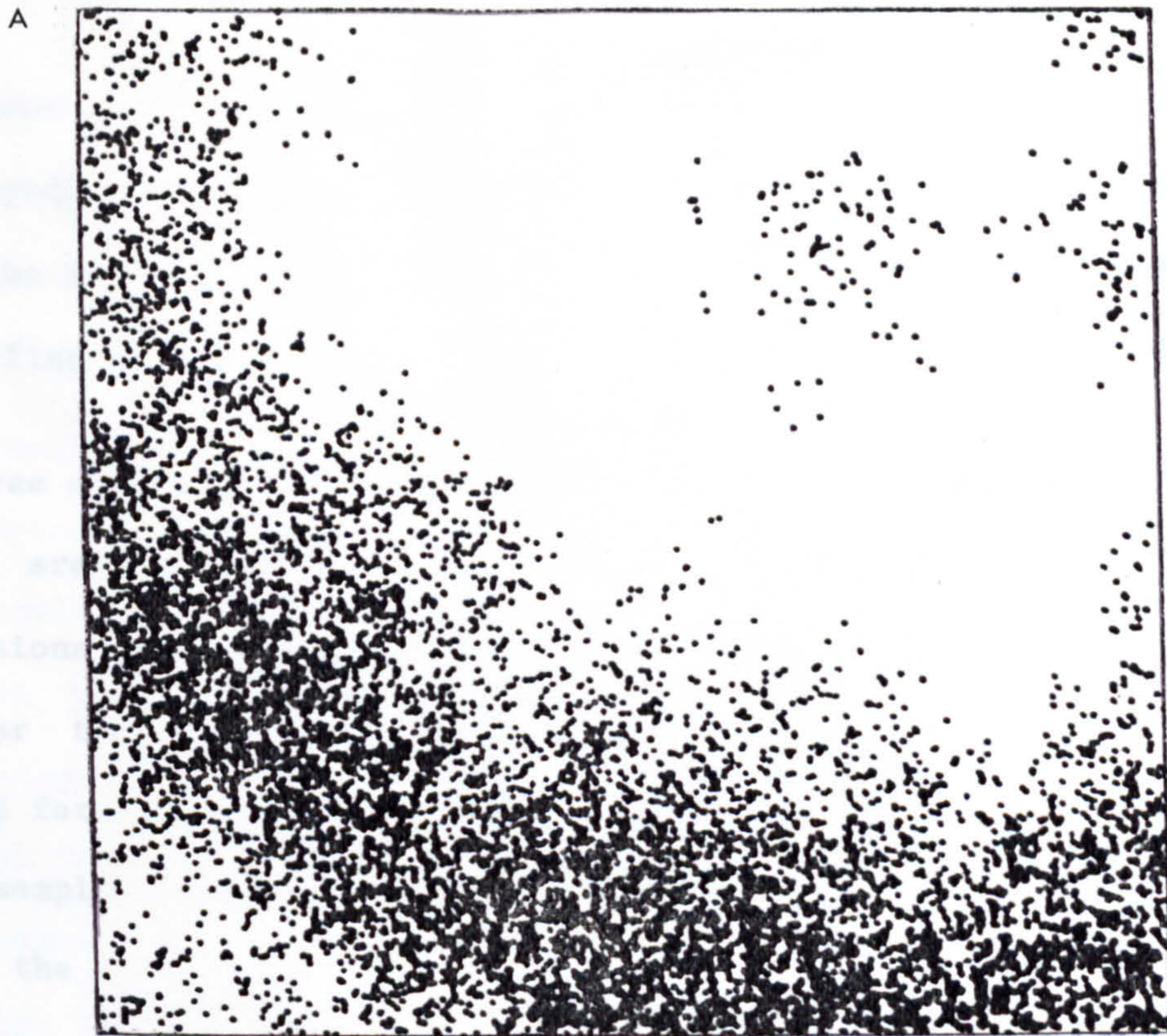


Figure 4.2 Two dimensional display of a function defined on a plane.

A) Dot density plot.

B) Contour plot.

functions (as in spline surfaces). The estimated function is then displayed by drawing contours which link points in the plane that have the same estimated function value. Different contours may be identified by labelling or by using different colours or line forms.

Three dimensional display methods are not used for the Heslerton study area because of the disadvantages outlined above. For two dimensional display both pixel and contouring techniques are used. However the contouring method is generally favoured over the pixel method for two reasons. Firstly both methods are trying to display the sampled function. The pixel method does this purely visually while the contouring method must first mathematically model the function. Thus the contouring method has greater latent analytical potential than the pixel method. Secondly contour displays are more flexible in displaying different ranges and subsets of information than pixel displays.

4.3 CONTOURING METHODS

Contouring methods are generally applied to regular data grids (systematic samples), although they can also be applied to irregular or scattered data. The first problem in contouring is the selection of a suitable function form to fit the data. It can be either a single global function or a network of contiguous local functions.

The best known technique for fitting a global function is trend surface analysis, where a two dimensional polynomial of suitable degree n is fitted to the data. The technique has had wide spread use in geography and many archaeological studies have used it. A typical example is that of Larson (1975) who fitted trend surfaces

to the distribution of tools in quadrats from Olduvia Gorge. The major disadvantage of trend surface analysis and global functions in general is that they do not fit the data consistently. Typically a trend surface will only fit a small section of the data satisfactorily, usually in the centre, while the remaining parts of the trend surface are totally disjoint from their associated data (the curly edge phenomenon). Hodder and Orton (1976) supply a classic example of this on page 167, figure 5.53 which shows a third-degree trend surface fitted to the distribution of Romano-British Oxford products. As well as indicating a high concentration of products around the Oxford kilns the trend surface also shows an equally high concentration of products in the middle of the English Channel. Theoretically a polynomial of degree n can be found to fit the data satisfactorily, but for data of realistic complexity, n will be large and the computations involved excessive. Therefore in general global functions are not really suitable for contour plotting.

Local functions can be defined as a patchwork of functions that individually fit a small set of local data satisfactorily and communally have some degree of continuity at their patch joins. The simplest local function is a linear one:

$$f(x,y) = a + bx + cy,$$

which describes a plane. The local function is defined on a plotting cell which consists of four grid point or node values (figure 4.3). Unfortunately the simple linear function can only be fitted to three nodes. If the selection of these three nodes is not done carefully then artifacts may be introduced into the contouring.

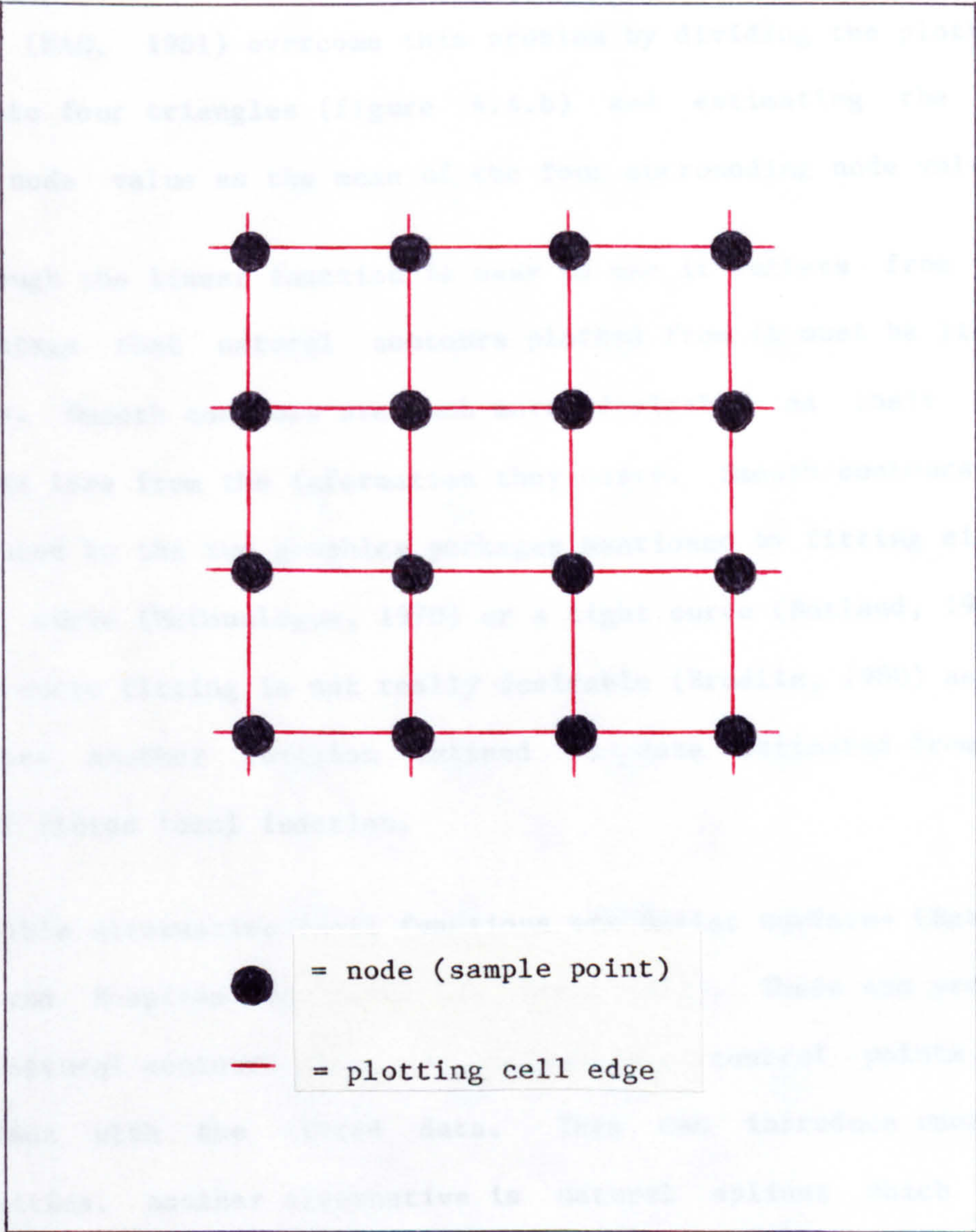


Figure 4.3 Definition of square plotting cells.

Such a situation occurs with the contouring routines in the GHOST graphics package (Noland, 1979), when the rectangular plotting cells are divided into two triangles along the diagonal (figure 4.4.a), there is a tendency for the resultant contours to be slanted in the diagonal direction. The contouring routines in another graphics package (NAG, 1981) overcome this problem by dividing the plotting cells into four triangles (figure 4.4.b) and estimating the new central node value as the mean of the four surrounding node values.

Although the linear function is easy to use it suffers from the disadvantage that natural contours plotted from it must be linear segments. Smooth contours are much more desirable as their form distracts less from the information they carry. Smooth contours can be produced by the two graphics packages mentioned by fitting either a loose curve (McConalogue, 1970) or a tight curve (Butland, 1980). However curve fitting is not really desirable (Brodlie, 1980) as it introduces another function defined on data estimated from the original fitted local function.

Possible alternative local functions are Bezier surfaces (Bezier, 1974) and B-spline surfaces (de Boor, 1978). These can produce smooth natural contours but are defined by control points not coincident with the fitted data. This can introduce unwanted complexities. Another alternative is natural splines which have coincident control and data points. Different orders of spline give different orders of continuity at the patch joins. With zero-order continuity two surfaces simply meet. With first-order continuity two surfaces have the same tangent at the join. With second-order continuity two surfaces have the same tangent and curvature at the join. McLain (1974) uses bicubic splines, giving second-order

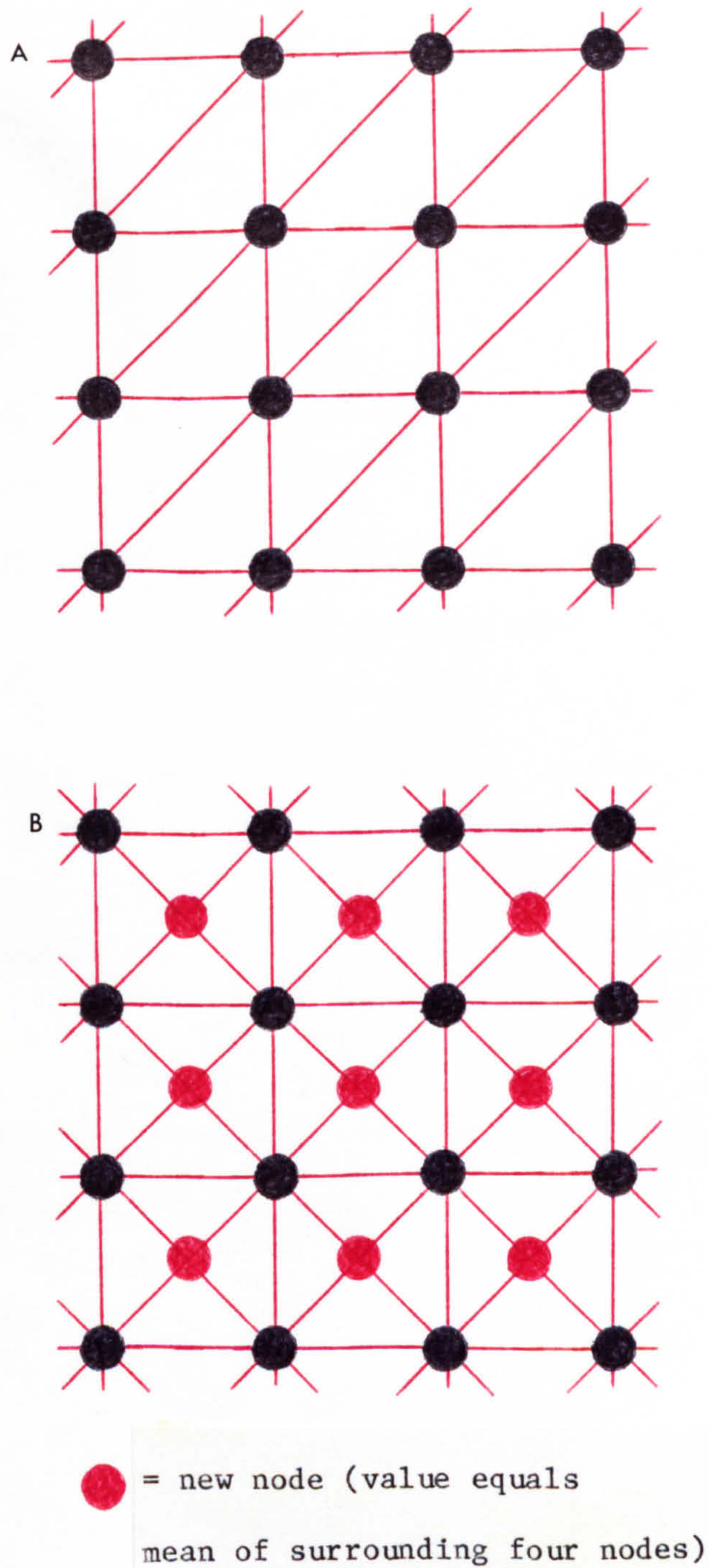


Figure 4.4 Simple linear function fit.

A) GHOST graphics package.

B) NAG graphics package.

continuity at patch joins, for contouring. The results are impressive but the calculations involved are long. Given the quality of the majority of archaeological data a coarser spline is more appropriate.

The bilinear spline is such a suitable choice:

$$f(x,y) = a + bx + cy + dxy.$$

The bilinear spline has a low order of continuity (zero-order) but it is easily fitted to the four nodes of a plotting cell, its linearity along the edges of the plotting cell simplifies the calculation of the points where a contour enters the cell and leaves it (entry and exit points). Contours take the form of rectangular hyperbolae (figure 4.5.a) within the plotting cell. With some node values asymptotes will be present in the cell, and where two asymptotes intersect a saddle point is defined (figure 4.5.b).

A curved contour $f(x,y)=c$, where c is a constant, may be plotted through a cell using the following algorithm (Kelly and Haigh, 1984).

Given a starting point $P_0(x_0,y_0)$ which is known to be on the contour, calculate the two derivatives:

$$u_0 = \frac{\partial f}{\partial x}(x_0, y_0),$$

$$v_0 = \frac{\partial f}{\partial y}(x_0, y_0);$$

the normal to the contour is in the direction of the gradient vector (u_0, v_0) and the tangent is in the direction of the vector $(v_0, -u_0)$.

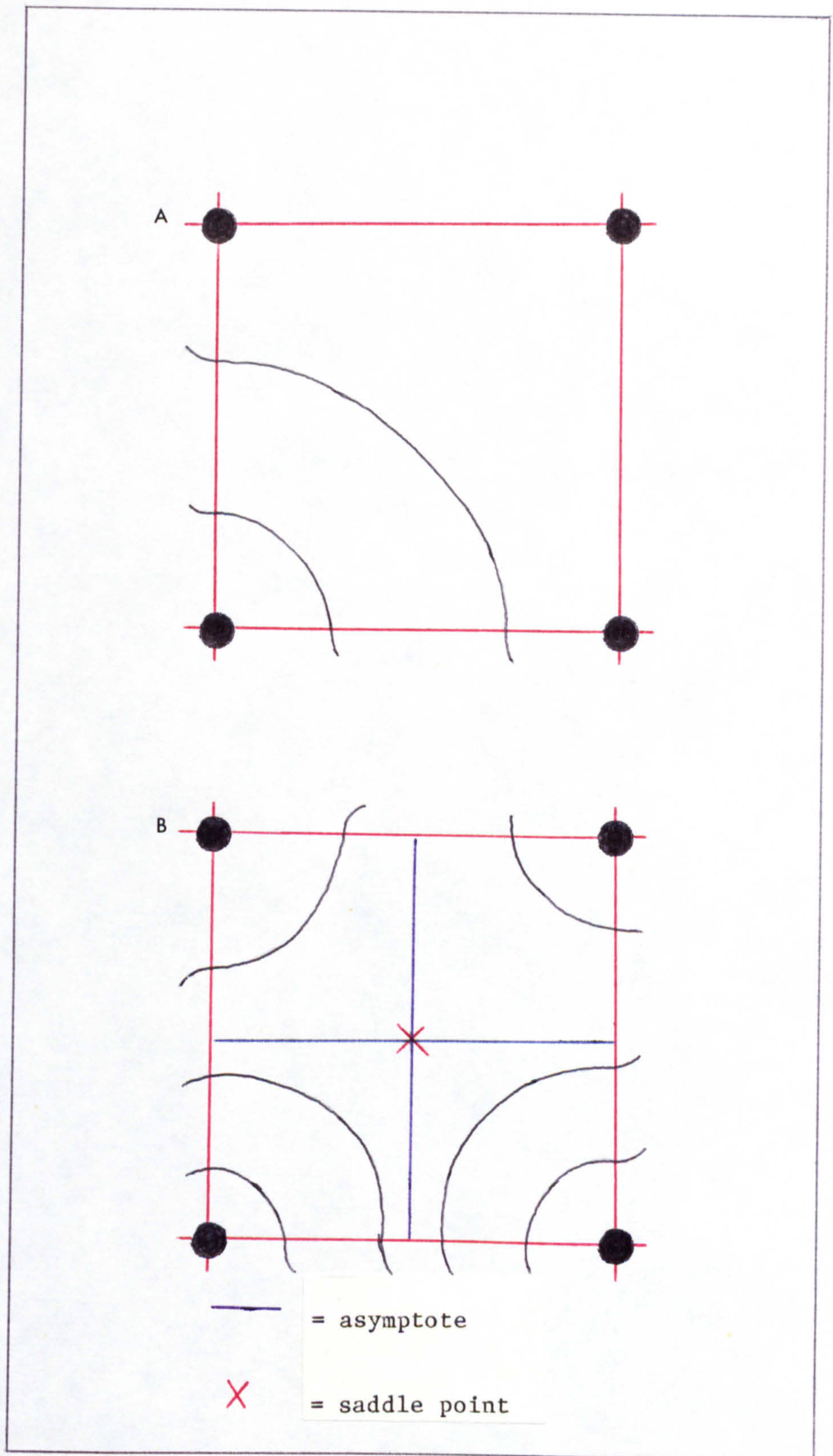


Figure 4.5 Bilinear spline contours.

A) Contour form.

B) Possible asymptotes and saddle point.

To move an estimated distance h along the contour, first calculate the point $P_1(x_1, y_1)$ on the tangent at P_0 , where:

$$x_1 = x_0 + \frac{v_0 h}{\sqrt{u_0^2 + v_0^2}},$$

$$y_1 = y_0 - \frac{u_0 h}{\sqrt{u_0^2 + v_0^2}};$$

(it may be necessary to reverse the signs of the two increments to ensure that the contour is traversed in the correct direction). To get back on to the original contour from the tangent at P_0 , apply incremental corrections along the normal at P_1 :

$$\delta x = u_1 k,$$

$$\delta y = v_1 k,$$

where

$$u_1 = \frac{\partial f}{\partial x}(x_1, y_1),$$

$$v_1 = \frac{\partial f}{\partial y}(x_1, y_1),$$

$$k = \frac{c - f(x_1, y_1)}{u_1^2 + v_1^2}$$

The point $P_2(x_1 + \delta x, y_1 + \delta y)$ should now lie close to the required contour, so that the line segment P_0P_2 may be plotted as part of the contour. The whole operation is then repeated using P_2 as the new starting point.

The bilinear spline offers an effective local function for contouring. It produces more natural contours than simple linear functions with or without curve fitting. It is easier to implement than the more sophisticated splines such as the bicubic. Although it has zero-order continuity and may produce cusps where plotting cells join, this may be appropriate to the quality of most sampled archaeological data. Severe cusps give a clear indication of poor data, which would be disguised by splines of higher-order continuity. For these reasons the bilinear spline has been chosen in the implementation of the following contouring program.

Contours are plotted within cells using the algorithm outlined above. The initial point for the plotting algorithm is taken to be the contour entry point. This is calculated by simple linear interpolation along the appropriate cell side. The contour is then plotted using a fixed step-length of one-fifth the cell size. This gives a good compromise between smoothness and number of calculations. When a plotting point is calculated which lies outside the cell, the point is recalculated as the exit point on the appropriate side. Simultaneously the exit point becomes the entry point to the next cell. Problems can occur when a contour passes close to a saddle point and the plotting algorithm becomes unstable. These problems are solved by detecting the point of instability and plotting directly to the exit point.

Contours are traced individually through the mesh of plotting cells using a technique outlined by Sutcliffe (1980). A table of possible intersections of a contour with each cell is set up. By simple linear interpolation those intersections that occur are marked. By scanning the edges of the mesh the start of open

contours can be located and traced through the mesh, removing intersection markers. Any remaining markers belong to closed contours which can then be plotted.

There is a problem when a contour passes through a node point, it is difficult to decide which cell it enters next. The problem is overcome by altering any such degenerate nodal values by 1%, given the quality of most archaeological data this will not noticeably effect the subsequent contours.

Initially the contouring program was intended for geophysical data. Historically this has been collected at Bradford in the form of 20 X 20 grids of readings (see 5.3). The contouring of large sites requires that data from a large number of grids should be accessible to the computer simultaneously, but there is clearly a limit to the number of grids which can be held in primary storage. The problem of limited storage is circumvented by augmenting each grid with an additional row from its neighbour above and an additional column from its neighbour to the right. The grids can then be brought into primary store one at a time; as each is plotted its contours will join exactly to those of its neighbours. This method obviously involves more complex data management than simple contouring.

The structure of the program (`contour_plot`) is outlined in appendix A.1. The program is coded in Fortran 77 for two reasons. Firstly the plotting primitives available at Bradford University Computer Centre are only accessible as Fortran subroutines. Secondly Fortran is well suited to the large and frequent number crunching tasks involved in the program. An example of the contours

produced by program `contour_plot` is shown in figure 4.6.

4.4 BAGS

BAGS (Bradford Archaeological Geophysics System) is a software package built around the `contour_plot` program described in the previous section. It is designed to offer easy and simple control of all data management tasks relating to the storage, selection, processing and presentation of data files. Again it was initially designed for geophysical data, but as subsequent chapters show it is general enough to be applicable to most data grids. The system structure is outlined in appendix A.2. It consists of a series of programs (written in Fortran 77 for compatibility with `contour_plot`) and CCL procedures. CCL or Cyber Control Language consists of operating system commands that may be stored in a procedure (or macro) and subsequently invoked (CDC, 1982). They are specific to the computer on which the system is implemented, the Cyber 170-720 mainframe at the University of Bradford Computer Centre. The accompanying user manual for BAGS is included in appendix B, this explains some of the terminology used in the system outline.

4.5 DISCUSSION

This chapter has outlined a basic conceptual data structure for the data sets involved in the Heslerton study area. Different methods of presentation for this structure have been considered. The contouring method has been put forward as the 'best' method and a detailed theoretical development and practical implementation for it has been given. But in what sense is this method best? The

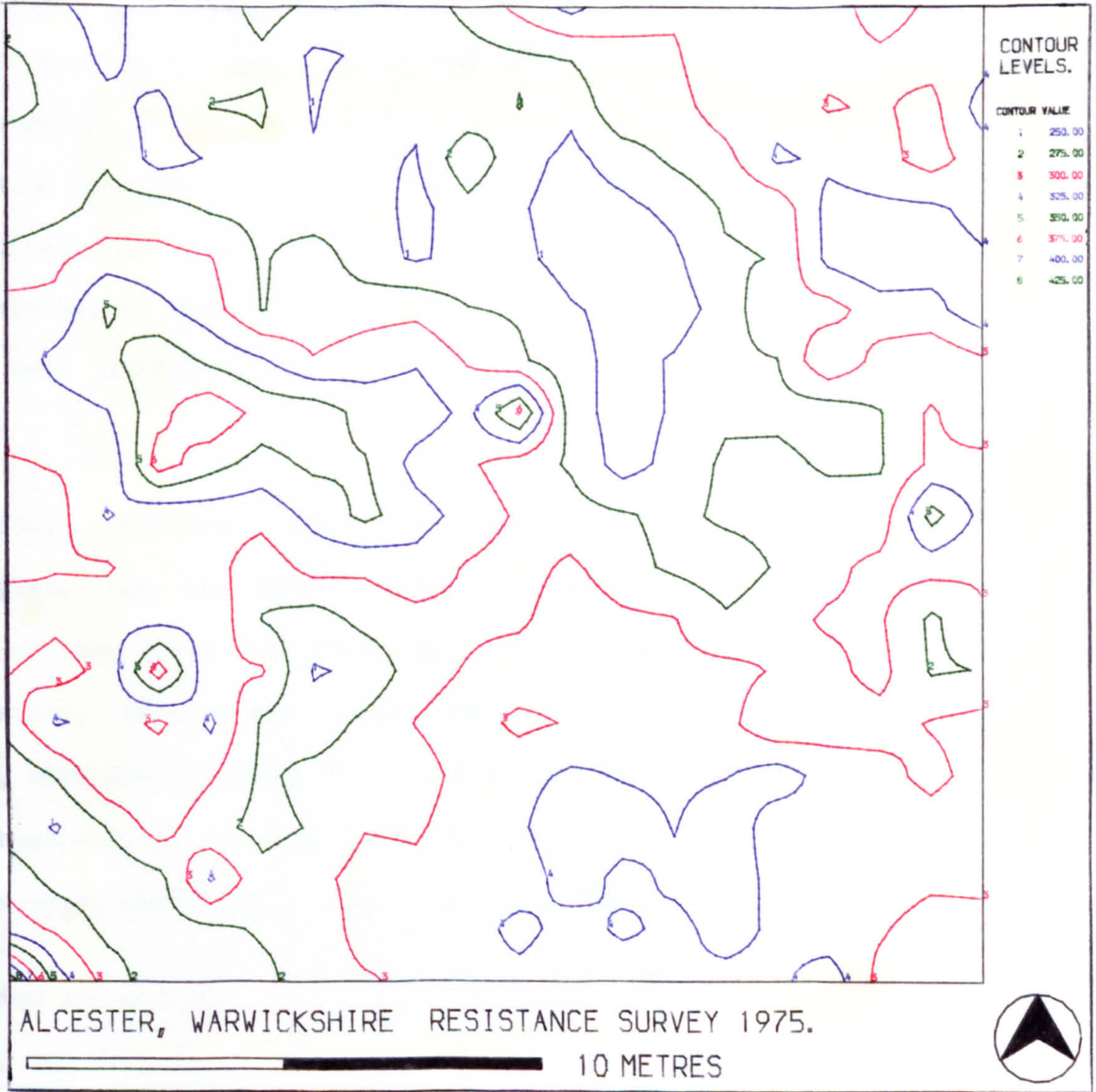


Figure 4.6 Contours produced by program contour_plot.

obvious answer is that it should be best in satisfying archaeological needs. However these needs are far from being well defined.

When presented with a data structure as described in this chapter, most archaeologists would attempt to identify features within it. The concept of features is important to archaeologists. They physically exist on excavations, are recorded on plans, identified on maps and aerial photographs. A feature is a solid entity, it has a definite closed boundary, so that the region inside the boundary may be distinguished from the region outside. This seems a very different data structure to a systematically sampled function.

To identify features within the sampled function data most archaeologists use the pixel method of display and alter the contrast of the data to give a very 'black and white' representation. This form of display is literally produced with dot densities, with the resulting black and white areas being equated with archaeological features. For this reason many archaeologists regard the pixel method as the best method of display.

This view is not really justifiable as the pixel method is being distorted to produce display forms that resemble features. The contour method does this naturally, as the contours produced act as boundaries between distinct valued regions. By distorting the pixel method to produce a simple display the majority of the information carried by the data is lost. In contrast the contour display method retains all its information. It must be accepted that at some stage a 'black and white' picture of features will be synthesised, but the

contour method provides a much more flexible and critical form for this synthesis than the pixel method.

Many archaeologists also forget that the data are a sample, which can lead to erroneous interpretation. The distance between sampling points defines a maximum information frequency (Ash, 1965). Thus in geophysical surveying at one metre intervals only detectible features with an instrument response greater than two metres can be located. The assignment of single readings to features such as post holes is not justifiable.

The detection of features can be regarded as a form of local synthesis, where the interest lies in defining significant partitions of the data. It is also of interest to look at the overall trends or characteristics of the data, a procedure which could be termed global synthesis. One method for doing this, trend surface analysis, has already been presented and criticised. But other methods can be used based on the concept of filtering. Techniques for filtering are normally used to help clarify local synthesis (Linnington, 1970), however they can work equally well for global synthesis. Data is assumed to carry a signal, which is of interest, and noise which obscures the signal. By filtering out the noise the strength of the signal is increased. But by taking the filtering a stage further the signal itself can be simplified.

For example the use of fourier transforms (Scollar, 1970) allows a complex function to be represented as a series of simple sine and cosine functions. By removing all the high frequency components the remaining low frequency function characterises the overall trend of the data. Yet again the contour method offers the best form of

presentation for these trends.

5.0 GEOPHYSICAL SURVEYING

5.1 OBJECTIVES OF GEOPHYSICAL SURVEYING

The term geophysical surveying covers a range of techniques which aim to discover and elucidate archaeological features below the ground without the necessity of excavation. Some of these techniques have been applied to archaeology from a relatively early date (Aitken, 1958) and many have been subsequently developed within archaeology itself (Clark, 1975).

Until recently these techniques have been referred to as 'prospecting' tools. This in part betrays their geological heritage, but it is also attributable to the traditional archaeological preoccupation with site discovery and delimitation. However over the last decade many archaeologists have moved away from this discrete view of the archaeological landscape to a more continuous one involving non-site archaeology (e.g. Aston and Rowley, 1974; Foley, 1981). Geophysical techniques have moved with this trend and are now more accurately referred to as 'survey' tools. Surveys of increasing size are being undertaken, a typical example is that of a magnetometer survey of a Romano-British settlement at Tiddington (Heathcote and Kelly, 1979) covering twenty thousand square metres. Archaeologists have become increasingly sophisticated in their expectations of geophysical techniques; from simple site location, to a prerequisite for excavation, to an independent mapping tool with clarification by sampling excavation.

5.2 METHODS OF GEOPHYSICAL SURVEYING

Geophysical surveying techniques were originally used by geologists for the investigation of geological formations in the base and drift geology. In archaeology only relatively shallow features below the surface are of interest. Therefore all the techniques used must be operational only for shallow depths.

Magnetic methods measure either the absolute magnetic flux, as in a proton magnetometer, or the relative magnetic flux change, as in a fluxgate magnetometer (Aitken, 1974). In both cases the magnetic properties of buried features such as ditches, rubble and kilns cause local variations in the earth's magnetic flux. This variation is obviously a continuous function defined on the ground surface, that can be recorded at discrete points. This is in fact the standard station survey method, in which a grid of points or stations is established on the ground (usually at intervals of 0.5, 1 or 2 metres). The instrument is then placed above each station and a reading taken. The fluxgate instrument is also capable of continuous readings, where it is run along a series of transects continuously recording magnetic flux changes on an X-Y plotter.

Resistance methods pass an electric current through the ground. The electrical resistance of volumes of earth are then measured, as different archaeological features such as pits and walls will have different resistance characteristics. The current generation and the resistance measurement are accomplished by inserting rods or probes into the ground. Different arrangements of probes (called arrays) can be used to discriminate different depths and fields of

measurement (Lynam, 1970). The resistance values, as in the magnetic methods, form a continuous function defined on the ground surface. However this function can only be recorded at discrete points as in a standard station survey.

For a standard station survey, geophysical readings are recorded at discrete uniform intervals. The resultant data have the structure of a systematically sampled function which is defined on a plane, the basic data structure introduced in 4.1. In a continuous transect survey, as is possible with a fluxgate magnetometer, the resultant data have a different structure. The density of data along the transects is much higher than between them; in fact the transects really represent profiles. This makes the two dimensional display techniques of 4.2 inappropriate. The usual method of representation is a stacked line display, which is a three dimensional display form and suffers from the inherent problems of interpretation outlined in 4.2. Therefore although in practical terms continuous transect survey offers a more rapid form of recording than station survey, the resultant data can be less amenable to interpretation than for the station survey method.

At Bradford both magnetic and resistance methods of geophysical surveying are available. For the Heselton study area magnetic methods were found to be unresponsive as there was no magnetic contrast between archaeological features and undisturbed ground. Theoretically resistance methods should also have been inappropriate, as the post glacial sands and gravels of the dry vale in which the study area is situated should provide poor earth resistance contrast. It was found however that under suitable climatic conditions, specifically late autumn, earth resistance

contrast was strong enough for resistance surveying to be feasible. For this reason a resistance method was chosen for geophysical survey in the Heslerton study area.

5.3 CONVENTIONAL RESISTANCE SURVEYING

The geophysical surveying technique used for the Heslerton survey was the two-probe resistance survey method. This uses a Bradphys earth resistance meter (Aspinall and Walker, 1975) which is designed to detect earth resistance variations due to the presence of shallow archaeological features. The meter has connections for four ground electrodes, two of which are used to force a constant electric current through the ground, while the potential difference is monitored between the remaining two. The potential difference is measured in a high-resistance circuit so that it is effectively independent of any contact resistance at the potential measuring electrodes. Since the current in the ground is kept constant, the potential drop between the two measuring electrodes is directly proportional to the earth resistance between them, and hence the instrument may be calibrated directly in terms of earth resistance.

The meter has two sensitivity ranges; on the first range full-scale deflection corresponds to an earth resistance of 10 ohms, while on the second range it corresponds to 100 ohms. An offset control extends the range further without reducing the sensitivity; the offset increases in steps, each corresponding to a full-scale deflection, so that resistances between 0 and 100 ohms can be measured on the first range and those between 0 and 1000 ohms can be measured on the second.

The ground electrodes are connected to the earth resistance meter in a two-probe configuration (Aspinall and Lynam, 1970). One current probe and one measuring probe remain stationary. They are set at a distance of several metres apart and are positioned well away from the survey area. The remaining current and measuring probes are termed the movable probes. They are fixed to a frame at a distance of 0.5 metres apart. The frame has handles to allow it to be moved, and a base plate onto which the earth resistance meter is usually secured. Wire leads connect the movable probes to the meter and a cable drum provides a connection to the distant stationary probes.

In comparison to other probe configurations, such as the Wenner, the two-probe offers considerable advantages. It is an easy and quick resistance survey method as the movable probes are moved as a pair (hence the name) to each station. Also the measured earth resistances relate to a localised volume of earth directly below the movable probes, making the interpretation of the results more straightforward than for other arrays.

Surveying is carried out in blocks or grids of 20 X 20 stations, arranged contiguously to cover the survey area. The usual practise is to locate on the ordnance survey map a base line from which the corners of the grids may be measured and marked with pegs. The 400 stations within each grid are then established at 1 metre intervals by means of tapes and marked lines. An absolute error of up to 25 centimetres for the location of an individual station point within the whole survey is acceptable.

Having established the survey grid the meter and probes are set

up. A single station is selected and marked as a calibration point. The movable probes are inserted at each station within a grid, the measured earth resistance being recorded on a record sheet. At the end of a grid a reading is taken at the calibration point. If there is any drift this is corrected and noted on the record sheet. At some stage on most surveys the limit of the cable drum is reached and the stationary probes must be moved to a closer location. This is done with the movable probes at the calibration point, the distance between the stationary probes at the new location being adjusted to give the calibration reading. A new calibration point can be established if necessary.

When the survey is completed the recorded resistance readings are usually typed into a computer system for processing, often to produce descriptive displays such as dot densities or contours. These displays may be used for interpretation, in an attempt to identify high resistance features such as walls and roads, and low resistance features such as pits and ditches.

When results have to be recorded manually, there are many problems associated with resistance surveying. The process is slow, since the results must be read from the meter and then noted down on the record sheet, and inaccurate, since errors are likely to arise at both stages. Furthermore there is no rapid means to produce a descriptive display for preliminary interpretation of the data. When the data are to be submitted for processing on a computer, normally the case for the large surveys which are becoming increasingly common, data preparation may take an excessive length of time. All these problems may be avoided by the use of an automatic logging system which transfers resistance values from the

meter directly into an on-site computer.

5.4 AUTOMATIC LOGGING FOR GEOPHYSICAL SURVEYING

The idea of automatic data capture or logging for geophysical surveying in archaeology is not new; Scollar and Kruckeberg (1966) used a logging system for magnetic surveying. Becker (1978) used a tape drive to log magnetic readings on-site. The readings could be later down loaded from the tape onto a computer for processing. Freeman (1981) designed a dedicated on-site display device to produce dot density displays; unfortunately the geophysical readings still needed to be recorded by hand and entered into the display device manually. Most recently Sowerbutts and Mason (forthcoming) have produced an automated system for small scale geophysical survey. The system consists of a microcomputer with disc drives and monitor, comprehensive graphics facilities, automatic logging and transducers to locate the position of the survey instrument.

All the systems described, except for the last one, do not supply comprehensive enough solutions to the problems of implementing on-site automatic logging and processing. Sowerbutts' and Mason's system although comprehensive is very cumbersome, being mounted in a wheel-barrow which also contains its supporting car battery, and too expensive to be within the budget of most archaeological organisations. A more suitable archaeological solution would be to use an easily portable and relatively inexpensive computer in the field, and support it with carefully developed facilities on more sophisticated computers. In association with co-workers at Bradford, the author has developed such a system, which is described

below.

5.5 A MICROCOMPUTER SYSTEM FOR RESISTANCE SURVEYING

The RSCS (Resistance Survey Computer System) is designed to log and process data from a two-probe resistance survey. In addition to the earth resistance meter and probes; a small portable computer, an analogue to digital converter (A.D.C.) package and a klaxon are used.

5.5.1 HARDWARE

The microcomputer is the Epson HX-20, which is a compact (A4 size) machine, weighing approximately 1.7 kg and featuring a full-size keyboard, a small LCD screen, a printer, a microcassette drive, and serial and RS-232C ports. It is powered by a rechargeable Ni-Cd battery allowing continuous periods of up to 50 hours operation. Such long operating periods are attained through the use of CMOS technology, which provides non-volatile memory and low power consumption. The HX-20 contains two 6303 processors, essentially CMOS versions of the Motorola 6800; there are 16K bytes of RAM, and a further 16K bytes may be added as a separate unit, which is attached to the main computer through a system bus connector. A Microsoft style BASIC is provided with the computer.

The Epson HX-20 was chosen for the automatic logging system for several reasons. Its compactness and its independent power supply make it very suitable for field use. Protected file areas can be defined within the computer's memory so that data can be accumulated

rapidly and efficiently, before being transferred to microcassette for permanent storage. The microcassette drive can be controlled from BASIC and can be made to simulate a very slow floppy disc drive. The built-in printer allows data to be processed and displayed in the field. Finally the bus connector, designed for the extension RAM, gives access to the system bus to allow the connection of the A-D converter package.

As the Bradphys earth resistance meter was not originally designed to be interfaced to a computer, some modifications were required to connect it to the A.D.C. Initially these modifications were incorporated in the meter itself (Kelly, Dale and Haigh, 1984), but were then packaged with the A-D converter. This package fits onto the bus connector on the side of the HX-20. A water-proof connector and lead takes the analogue voltages for the resistance reading, offset value and scale to the A-D converter package (figure 5.1). The scale voltage is passed directly to the A.D.C., while the resistance reading and offset value are both passed through filters (to eliminate noise) and amplifiers before reaching the A.D.C. The A.D.C. is an 8 bit 8 channel converter that can accept up to 5 volts input. This gives a sensitivity of 20 mv per bit, with conversion times of 1 ms or less. The A.D.C. is operated by accessing a predefined memory location. The address of this location is decoded by the address logic within the A-D converter package to activate the A.D.C. The converted voltages for the 8 channels are then memory mapped to 8 memory locations where they can be accessed from software. The A-D converter package also contains two other facilities. When the movable probes are removed from the ground, the resistance meter gives an effectively infinite earth

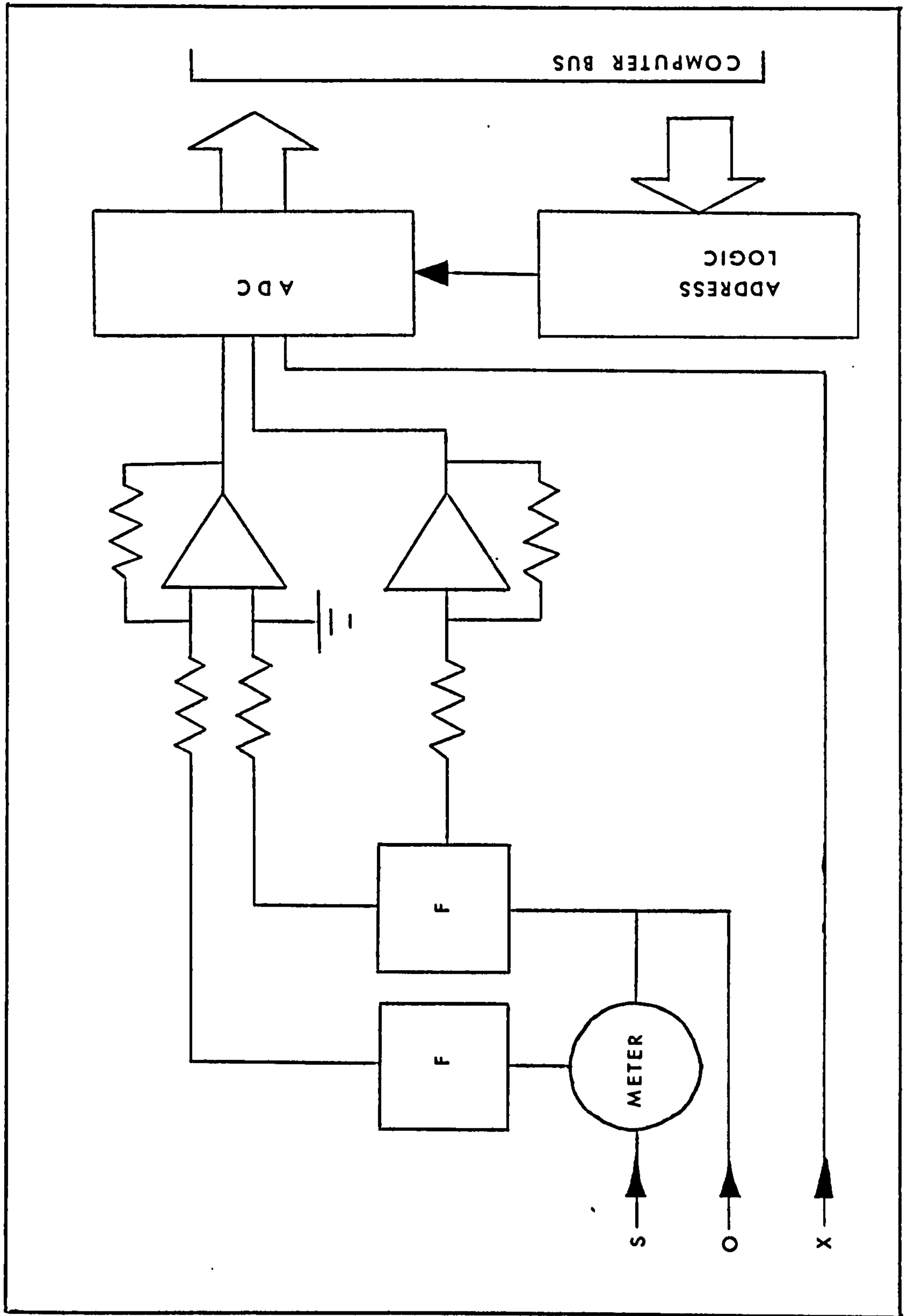


Figure 5.1 A-D converter package.

S - signal, O - offset, X - scale, F - filter.

resistance reading. When this condition occurs a LED is illuminated on the A-D converter package and an A.D.C. channel is set high. This allows the software to detect when the probes are being moved and thus prevent resistance readings being taken. Another predefined memory location activates an output from the A-D converter package when accessed from software, which sounds an external klaxon.

5.5.2 SOFTWARE

All the software for the system is written in BASIC except for a machine code routine which handles communications with the A-D converter package. The system software is divided into a number of functions which are initiated by pressing an appropriate function key at the top of the HX-20 keyboard. These functions are described in the system user manual (appendix C). An outline of the software structure for the RSCS is given in appendix A.3.

Although the software works adequately, programming problems are foreseen in possible future system developments, where extra external devices such as location transducers are added. The extra control and communication functions would necessitate a large amount of programming in machine code. A more attractive alternative would be to implement the system in FORTH, as this is ideally suited to the development of real time control systems (Brodie, 1981). FORTH is the best known example of a threaded interpretive language (TIL). A TIL can be used in interpretive mode for development, then in incremental compile mode to produce final production quality code. The expandable keyword list makes developed software both fast and efficient. A ROM based FORTH is available for the EPSON HX-20.

5.5.3 FIELD OPERATION

The RSCS is operated as a standard two-probe resistance survey. The only differences are that the meter and attached computer are positioned by the stationary probes and the movable probes are now connected to the meter via the cable drum (figure 5.2). This arrangement saves the computer from the vibration that might be expected from the moving probes, and also enables it to be protected during inclement weather conditions. It is possible for one operator to run the whole system, with the meter and computer both mounted on the moving probes, but it is much more efficient to run a two-man operation, with one operator moving the probes and the other supervising the data logging.

The survey procedure is exactly the same as for normal surveying except that when a resistance reading is logged the klaxon will automatically sound once to signal to the probe operator to move to the next station. If the computer operator tries to log a reading when the movable probes are out of the ground or are not making proper contact, the klaxon will automatically sound twice and the reading will be ignored. Once the readings for a grid have been logged, the grid is stored as a file on the microcassette drive (after any necessary drift corrections). Because the microcassette drive is relatively slow, the storage may take two to three minutes, but this occupies a period when the operators are busy preparing to survey the next grid.

The processing facilities of the RSCS allows the grids logged during the day to be displayed for preliminary interpretation before

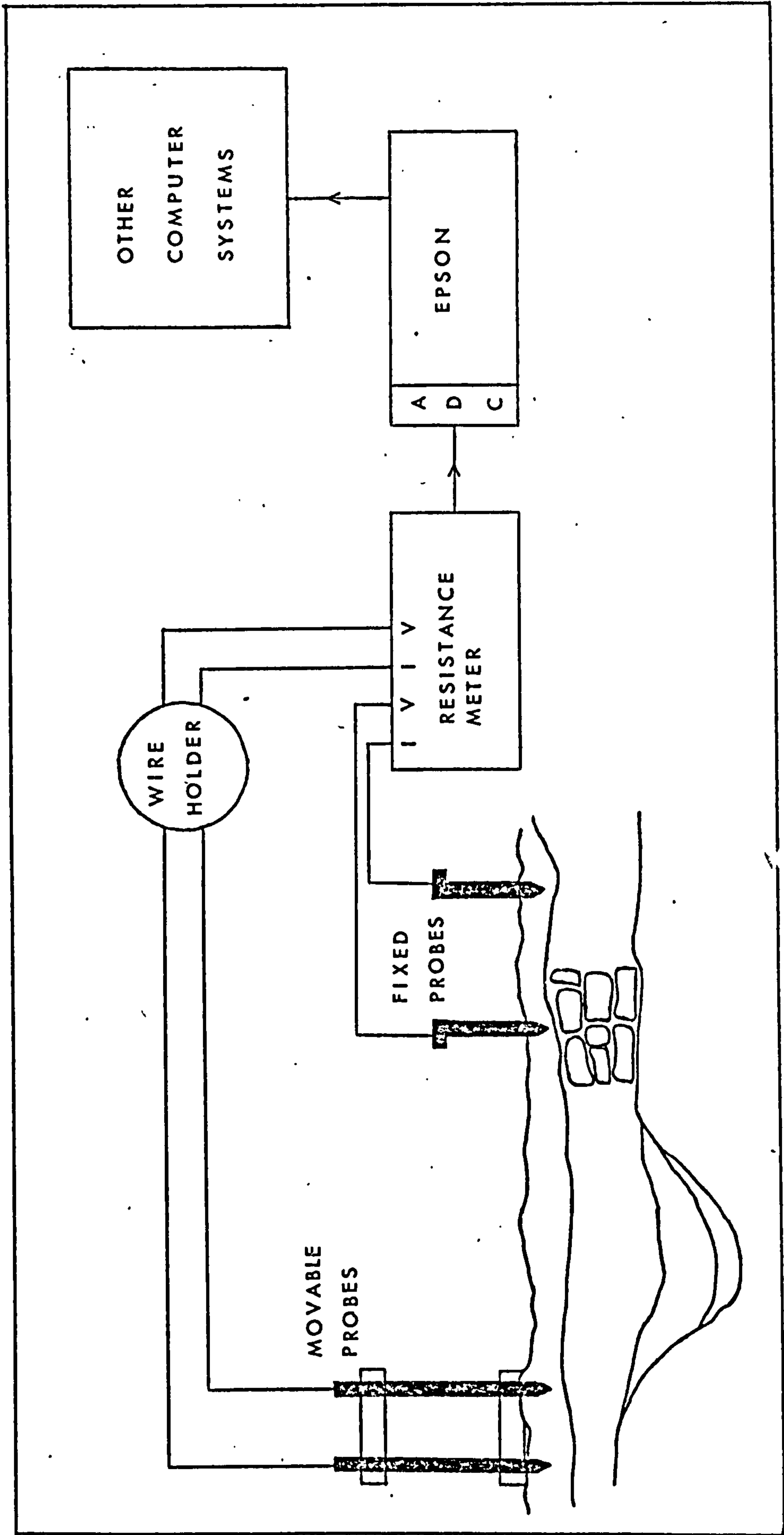


Figure 5.2 RSCS survey rig.

the next day's surveying. Any gross errors or mistakes can be easily detected and corrected before the survey is finished.

5.6 THE HESLERTON GEOPHYSICAL SURVEY

A resistance survey of the Heselton study area was made, using the RSCS, on the 19th and 20th of November 1983. The weather conditions were generally clear and bright for the two days, with a few occasional short showers. The ground conditions were firm but very moist, as there had been considerable rainfall over the preceding two weeks. The ground had been ploughed then harrowed flat for the planting of winter wheat, which was just visible as seedlings. For the survey the study area was divided into twenty five 20 metre by 20 metre grids, each containing 400 stations at 1 metre intervals. The survey was accomplished without any technical problems with the computer system or the surveying equipment.

The resistance readings logged are presented in dot density form in figure 5.3 (as produced by RSCS) and in contour form in figure 5.4 (as produced by BAGS after transfer). The most prominent features visible on both displays are the vertical striations. These run the entire length of the survey area in a N-S direction and seem to be regularly spaced at an interval of several metres. The explanation for these striations was discovered in the excavation of summer 1984. The excavation included a trench which had adjacent areas of untouched and stripped plough soil (figure 5.5). Ditches were discovered in the trench running into both areas. A resistance survey of the untouched plough soil area showed the familiar striations and very little else (top of figure 5.5).

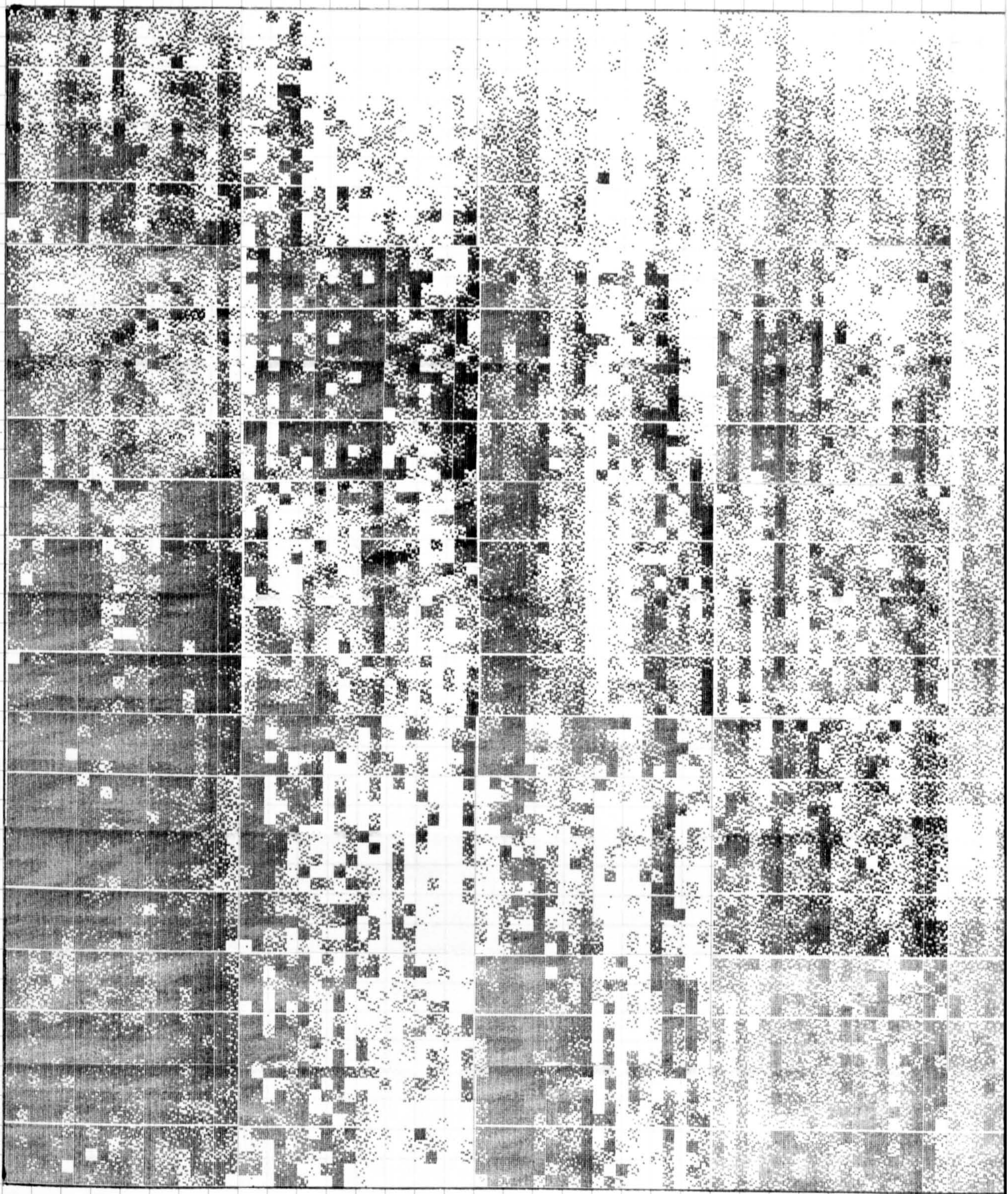


Figure 5.3 Dot density display of Heselton resistance survey.

Display range 136-169 ohms.

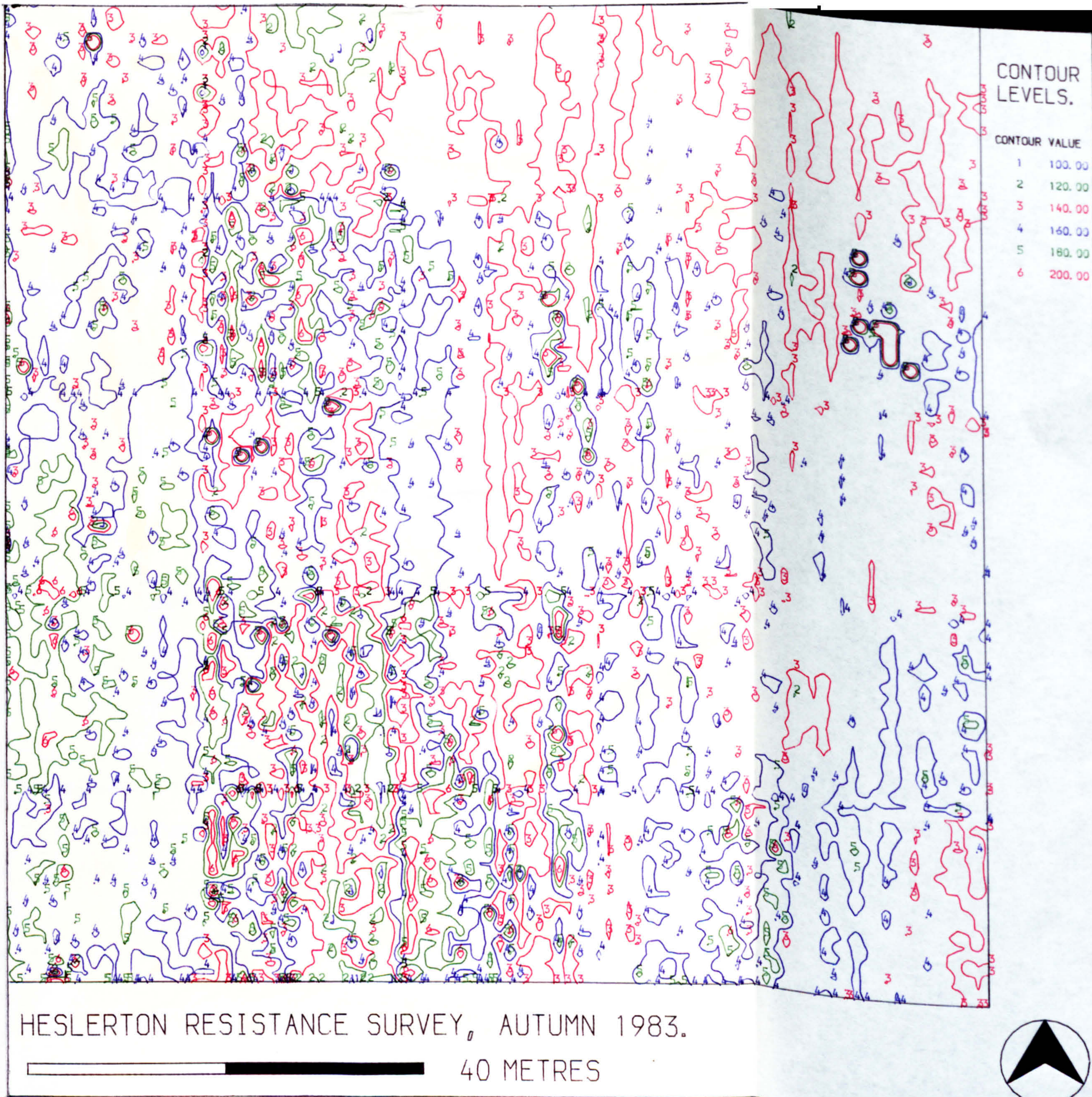


Figure 5.4 Contour display of Heselton resistance survey.

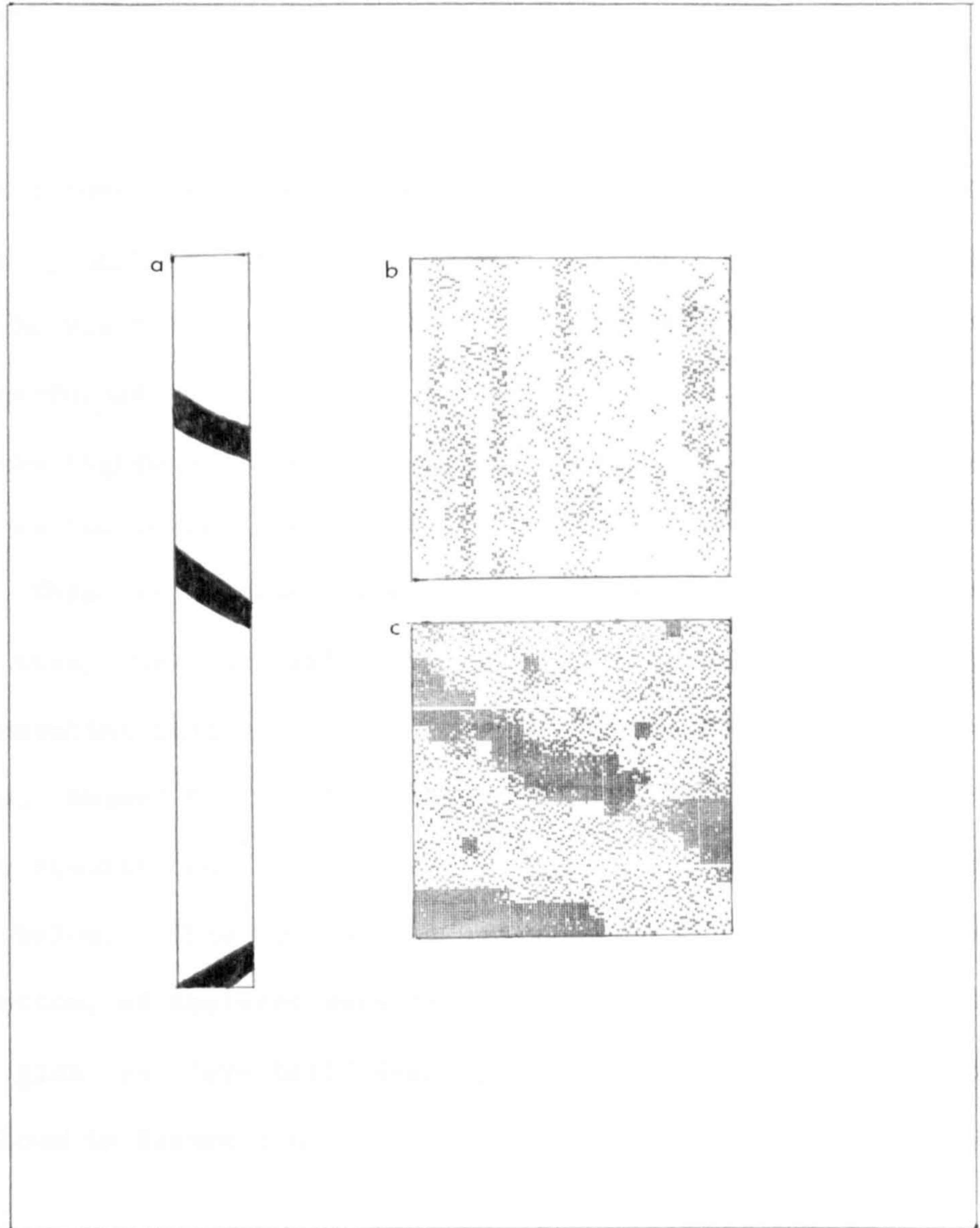


Figure 5.5 Investigation of striations.

A - excavated trench, B - plough soil, C - plough soil removed.

In contrast a resistance survey of the area stripped of its plough soil clearly showed the ditches without the striations (bottom of figure 5.5).

The striations are therefore most probably caused by the effects of modern ploughing in the top soil. As the sandy soil is well drained, the visibility of features to the resistance survey method will be determined by their water retention characteristics. It can be seen from figure 5.5 that the ditches which would normally be regarded as low resistance features, are visible as high resistance features. This is because the ditch fills offer much better drainage than the surrounding undisturbed ground (it was noted during excavation that the ditch fills were excessively dry). These striations, caused by the ploughing of the top soil, are saturating the survey results and substantially masking any archaeological features below. This makes the identification, let alone the interpretation, of features very difficult. But an attempt has been made to give an 'eye-ball' description of the features, with the results shown in figure 5.6.

5.7 DISCUSSION

This chapter has considered the aims of geophysical surveying in archaeology and considered the techniques available to achieve them. For the Heslerton study area the two-probe resistance survey method was found most appropriate. The inherent problems of this survey method, and resistance methods in general, were discussed and a solution put forward in the form of the RSCS. This system offers a cheap viable solution to the problems of data logging and on-site



Figure 5.6 Interpretation of Heslerton resistance survey.

processing. In economic terms the cost of the system will be recoverable over half a dozen medium sized surveys through the savings in unnecessary data preparation. This does not take into account the reduced survey expenses due to increased survey speed, nor increased efficiency as errors in collecting data and transcribing data to storage are eliminated.

Although technically the RSCS worked flawlessly on the Heslerton survey, the resultant logged data are disappointing in that no archaeological features are clearly visible. This is mainly due to the marginal soil characteristics and conditions for the geophysical survey method used. However most geophysical surveys are marginal in the sense that features are hard to define. It is the smaller minority of surveys with obvious outstanding features that tend to be presented as representative of the results of geophysical surveying. This has influenced the development of data processing techniques which have not progressed very far from the early work with its emphasis on descriptive displays (e.g. Wilcock, 1970). No methods exist for the objective interpretation of geophysical data. Interpretation itself can be broken down into two distinct processes. Feature extraction in which significant features in terms of the survey method are identified. Feature interpretation in which extracted features are equated with possible archaeological features such as pits and ditches. The presentation of data as descriptive displays, such as dot densities or contours, allows these processes to be carried out on an intuitive basis. However for the majority of surveys this intuitive basis is often inadequate.

More sophisticated methods of data processing possibly lie in the

techniques of image processing. In image processing the descriptive display methods are formed from radiometric operations, in which the relative values or intensities of stations or pixels are manipulated (Star, 1985). In contrast stretching the highest and lowest values are forced to the extremes of the display representation, and all intermediate values are linearly varied. In density slicing only a specified range of values are displayed. A combination of both techniques is implicitly used in dot density displays.

Image processing also uses spatial operations in which the spatial relationships between values are used in processing them. Filtering is an important spatial operation which attempts to remove noise from a signal (see 4.5). For the Heslerton survey data the archaeological features represent the signal and the striations represent the noise. It would obviously be of great value to be able to remove this noise from the survey, as the archaeological features are poorly defined. An attempt was made to do this using Fourier transform filtering as implemented in BAGS by S.Ipson (A.2). Figure 5.7.a shows an area of the survey unfiltered. A two dimensional Fourier transform was applied to the data, the high frequency components removed, and then the inverse transform applied to reconstitute the data. The resultant data (figure 5.7.b) shows no marked improvement over the unfiltered data. This was found to be the general case for a variety of filtering experiments with the Fourier transform. This tends to indicate the strength of the noise and the weakness of the underlying signal in the Heslerton survey data.

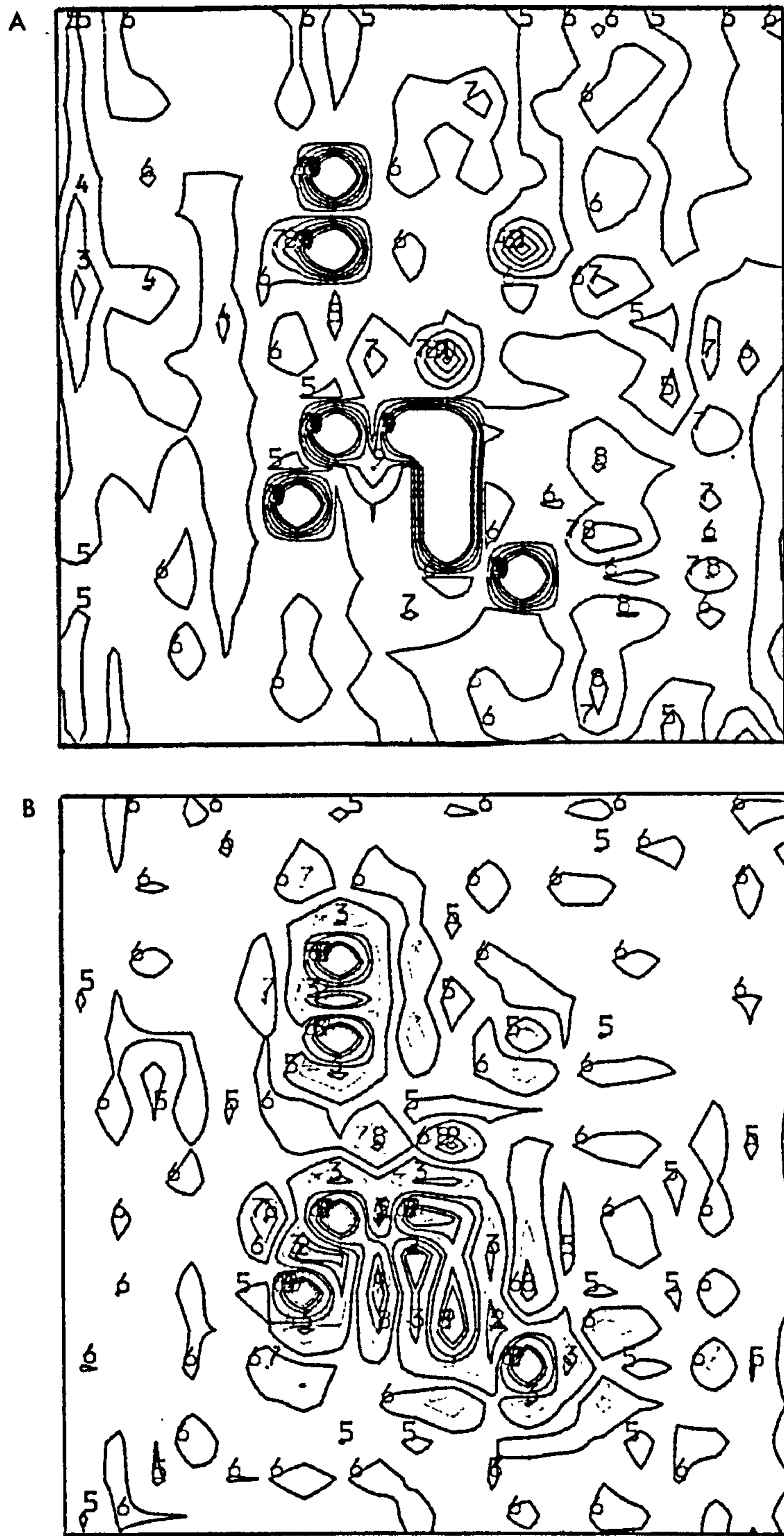


Figure 5.7 Fourier transform filtering.

A) Unfiltered.

B) Filtered.

6.0 SURFACE DISTRIBUTIONS

6.1 CAUSE AND EFFECT

Surface distributions of archaeological material, such as pottery and tile, are often visible in the field. Traditionally significant concentrations of these distributions are judged to be diagnostic of underlying archaeology or of a site (Coles, 1972). However the processes that bring material to the surface and determine their subsequent distribution are not understood and are a source of controversy among many field archaeologists.

The majority of the research into these processes has been empirical and has concentrated on the determination of whether the spatial distribution of surface material relates to buried archaeological features. Redman and Watson (1970), Roper (1976), Ammerman and Feldman (1978), Taylor (1979) and most recently Bradley (1984) have all found a strong spatial association between surface material concentration and buried archaeology, even where disruptive processes such as ploughing have been operational. In contradiction Tolstoy and Fish (1975), Hirth (1978), Bowers, Bonnichsen and Hoch (1983) and Frink (1984) have all found significant alterations to the relationship between surface distributions and buried archaeology due to processes such as ploughing, seasonal rainfall and frost-thaw activity. These processes of disturbance can in some instances have differential effects depending on material size (Baker, 1978) and material weight (Rick, 1976) producing a spatial segregation of the material. In other instances these same

processes produce no apparent segregation at all (Fuchs, Kaufman and Ronen, 1977). Other less obvious processes such as human recreation (Hammond and Hammond, 1981) can also significantly alter the spatial structure of surface distributions.

The problem of defining cause and effect for surface distributions is obviously complex. There are a large number of possible processes operational which seem from the present research not to behave in any consistent and predictable manner. Therefore the formulation of a general model of surface distribution formation applicable to all areas under all possible conditions is impractical, if not impossible. A more pragmatic approach must be taken which involves defining and characterising the processes operational for a particular area under a particular set of conditions.

For the Heslerton study area these conditions are relatively easy to define. Since the study area is flat and level, any sorting or movement of material by weight or size down inclines can be disregarded. The major process of disturbance present is that of seasonal ploughing. This has produced sizable quantities of archaeological material on the surface over recent years. To understand how material is transported to the surface a surface profile model can be constructed, as shown in figure 6.1. Below the plough soil there are the archaeological deposits (AD) which contain the undisturbed archaeological material (except for burial movement processes; see Rolfsen, 1980). Between the archaeological deposits and the plough soil is an archaeological interface (AI). This is where new material is introduced into the plough soil. In a situation where the physical level of the surface is altered due to

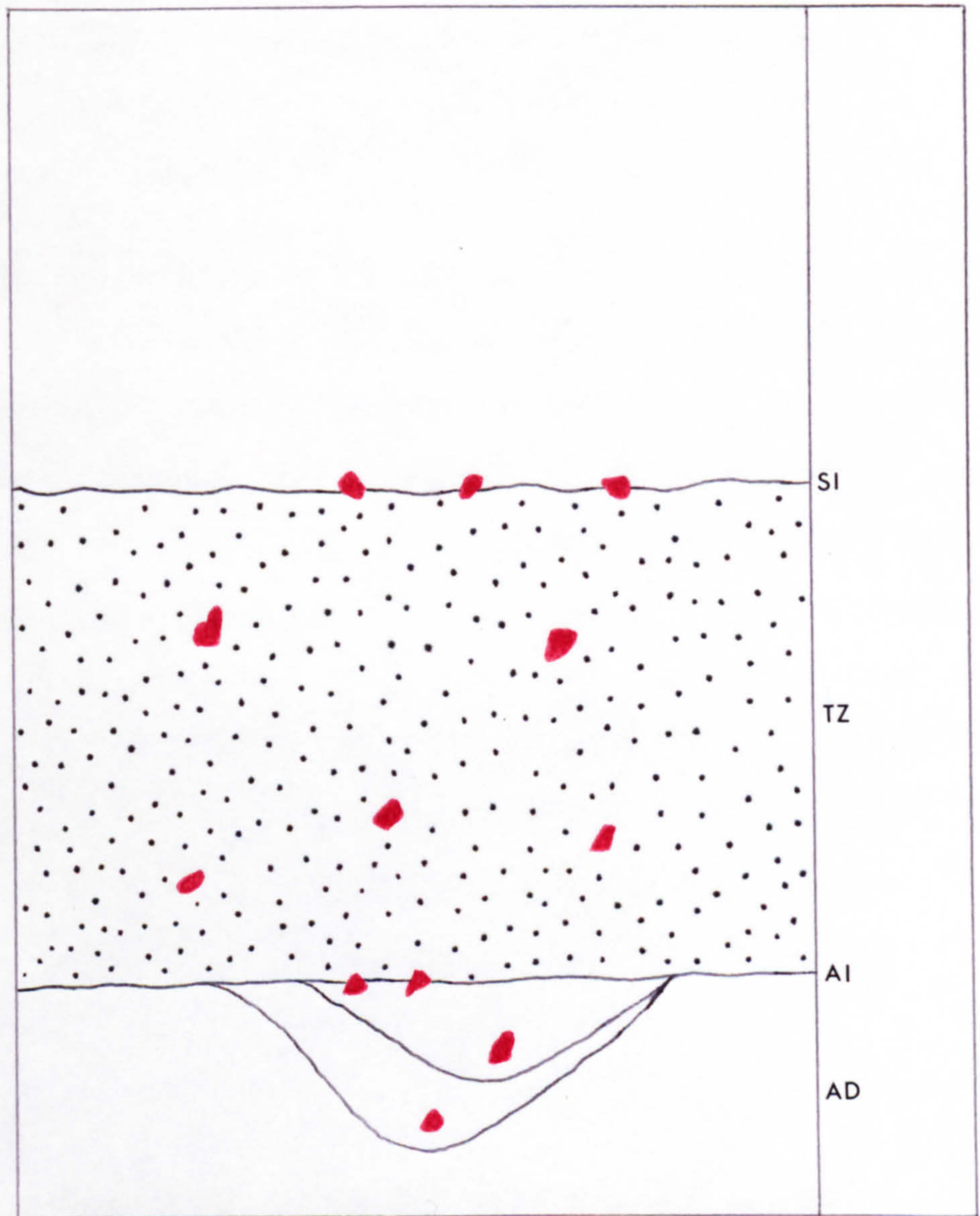


Figure 6.1 Surface profile model.

soil movement by ploughing, this interface will be progressively lowered into the archaeological deposits, as the depth of ploughing is constant. Once material is detached from the archaeological deposits it stays in the plough soil or transport zone (TZ) where it may suffer degradation by chemical, biological and mechanical agents. Material in the transport zone may eventually cross the surface interface (SI) to be deposited for a while as part of a surface distribution.

To study the formation of the surface distributions at Heslerton two test pits were excavated, one in the study area and one in an adjoining field where no surface material was visible but where archaeological features were present as indicated by aerial photography. The test pits consisted of one metre squares which were excavated in 5 cm spits. All pottery sherds within a spit were recovered by sieving and were then individually weighed to the nearest 0.1 gm. The results are shown in tables 6.1 and 6.2, which suggest there may be some segregation of pottery by weight at different depths within the transport zone. However, by assuming that each group of pottery from a spit is a random sample and applying the t-test to the mean pottery weights for each pair of adjacent spits, any segregation in either pit was found not to be statistically significant at the 5% level. These results are obviously not conclusive as the sample sizes are small. The experiment could be improved by enlarging the test pit size or by using multiple test pits. With the latter scheme spatial variations could be studied and large pooled samples achieved. However the results as they stand do seem to indicate that there is no segregation by weight or size (as size is directly proportional to

Level (cms)	number of sherds	total sherd weight	mean sherd weight	standard deviation of sherd weight
0- 5	20	46.9	2.35	2.38
5-10	15	75.9	5.06	7.51
10-15	18	49.1	2.73	2.11

Table 6.1 Study area test pit pottery sherd weights (gms).

Level (cms)	number of sherds	total sherd weight	mean sherd weight	standard deviation of sherd weight
0- 5	4	31.3	7.83	6.24
5-10	2	16.6	8.30	7.78
10-15	5	11.9	2.38	2.33

Table 6.2 Adjoining field test pit pottery sherd weights (gms).

weight) in the transport zone. This in turn tends to support the assertion that for the Heslerton study area the material crossing the surface interface is representative of the material in the transport zone below.

Material when crossing the archaeological and surface interfaces or when resident in the transport zone will be subject to forces causing spatial displacement. This displacement can be resolved into a vertical and lateral vector. The vertical vector is responsible for the presence of surface material, while the lateral vector will determine how far material will travel horizontally from its source point in the archaeological deposits. For the Heslerton study area the effects of lateral vectors are expected to be minimal for two reasons. First the surface distributions have been formed only recently and, even for large lateral vectors, there has not been enough time to accumulate significant horizontal displacements. Secondly the dry sandy soil of the study area has low drag characteristics, thus material or soil after being acted upon by a force, such as ploughing, will tend to settle quickly rather than travel with the acting force.

The surface distributions in the Heslerton study area should therefore probably reflect quite accurately the spatial structure of their source archaeological deposits.

6.2 METHODS OF SURFACE COLLECTION

The three basic approaches to surface collection are sampling, census and selection. In census every piece of archaeological material on the surface is located and its position recorded. The

material is then either removed for later analysis or the attributes of interest are recorded in-situ. The problem with census collection is twofold. It is very difficult to ensure that all material has been located. A large pottery sherd sticking out of the ground is obvious, but a small fragment of pottery the same colour as the soil is much more difficult to locate. Census therefore involves very intensive detailed searching which will be costly in resources. Once located the position of material must be recorded. To do this accurately sophisticated techniques must be used, which again are expensive in terms of resources. The advantage of census collection is that the complete range of types and scales of recordable information are available as the collection unit used is the material itself (Rogers, 1982).

Sampling in archaeology is usually explicitly associated with random or stratified sampling strategies (Mueller, 1974; Wobst, 1983). These sampling schemes although useful for testing or estimating the values of scalar attributes of archaeological entities are not suitable for spatial studies. Only systematic sampling can provide suitable information for reliable spatial analysis (Foard, 1977). In systematic sampling observations are made at regular or systematic places in the sampled population. These observations in surface collection are usually made by defining a quadrat within which all archaeological material is recovered. Quadrats are also usually made contiguous to completely cover the surface area being studied. This total coverage could be argued to give a census rather than a sample. In one sense this is correct, since all the physical material has been totally collected. But information relating to this material, notably its spatial

structure, has only been sampled. For example the density of material is defined as the number of pieces of material per unit area. To measure this we could use a square quadrat, count the number of pieces of material in the quadrat and divide this number by the quadrat area. But what does this measured density apply to, the whole quadrat or a point within the quadrat? To answer this question density should be regarded as a continuous function defined for every point on the surface. By using a quadrat we are simply sampling this function at a set of discrete points, by estimating its mean value over each of the quadrats. This mean value applies to the whole quadrat but for convenience we consider it to belong to a single representative point within the quadrat, usually its centre. By using a contiguous grid of quadrats we are therefore systematically sampling function values within the survey plane. This is the basic data structure defined in 4.1.

The advantages of the sampling approach are mainly in its cost effectiveness. It consumes far less resources than census collection and in many circumstances is the only practical collection method. Its main disadvantage is that the quality and minimum scale of information recoverable is dependant on the size and shape of quadrat used. Much work has been done in plant ecology looking at the characteristics of information returned from random quadrat sampling (e.g. Zahl, 1974; Zahl, 1977; Perry and Mead, 1979) and contiguous quadrats (e.g. Mead, 1974). In all cases the quadrat size and shape determine the scale of spatial structure recoverable. A large quadrat size allows rapid collection but limits the size of any definable spatial structure. Conversely a small quadrat size allows finer spatial structure to be discerned

but involves slower and more expensive collection. A balance must obviously be struck between spatial detail required and resources available. However in reality it is often the resources available that significantly determine this balance.

Even when using sampling collection, biases will be present. These are introduced by the least controllable component of any collection system, the fieldworker. Individuals will introduce recovery biases due to their intrinsic individuality and the physical masking of the material to be collected (Clarke, 1978). A fieldworker on a bad day or with a personal interest in a particular type of material will bias the samples recovered through discrimination and selection. Also differing physical conditions will tend visually both to obscure and to highlight different types and forms of material to the fieldworker. These biases are difficult if not impossible to eliminate. If they can be kept constant or can be measured they can be corrected for. But the determination of biases often presents a set of problems more formidable than the original problems motivating the collection.

The last approach to surface collection is selection in which material and information are recovered on a subjective basis. Selection is the least useful collection method, as although it may seem to satisfy better the aims of collection than the previous approaches it provides no context for further analysis. For example a selection of diagnostic pottery may seem the best method to date a surface site, but it provides no context in terms of the total surface sherd assemblage in which to judge the significance of this dating. Unfortunately although selection is the worst collection method it seems to be the most widely used in archaeology.

6.3 THE HESLERTON SURFACE COLLECTION SCHEME

The collection scheme used to record the surface distributions in the Heselton study area was based on systematic sampling using contiguous quadrats. A preliminary inspection of the surface of the study area gave a subjective idea of the types of material present and their average densities. As these densities were relatively low it was decided to use a contiguous square quadrat of size 5m X 5m, which divided the study area hectare into 400 quadrats. The quadrats were located on the ground by using a collection grid. This consisted of a 20m X 20m square divided into sixteen 5m X 5m cells. It was constructed from nylon twine which allowed it to be secured to the marker pegs for the geophysical grids and rapidly moved from grid to grid. This system is obviously only usable if there is minimal surface vegetation, as was the case in the Heselton study area.

The fieldworkers used for collection were specifically selected to have minimal archaeological fieldwork experience. A selection of the material types present was gathered from the surface adjoining the study area and shown to the fieldworkers. They were instructed to systematically cover their quadrats collecting everything they saw that they considered 'of interest'. They were instructed not to examine any material they considered collectible, but to simply place it in their collection bag, giving the simple rule "everything you touch goes in the bag". To ensure even coverage of quadrats a collection time was set. For the densities of material in the study area a collection time of five minutes was considered satisfactory.

Collection was carried out in the following manner. The collection grid was established in a 20m X 20m square. Empty collection bags were placed in each quadrat. The fieldworkers selected quadrats and were told to start collecting. Five minutes later they were told to stop collecting. They sealed their collection bags, marked their initials on them, left their bag in the quadrat, moved to the next free quadrat and waited to be told to start collecting again. When all the quadrats within the collection grid had been sampled the collection grid was moved to the next 20m X 20m square. The collection bags were then recovered and their locations recorded.

6.4 HESLERTON SURFACE COLLECTION RESULTS

The surface distributions within the Heselton study area were sampled over two weekends, the 22nd to 23rd and 29th to 30th of October 1983. The physical conditions for both weekends were similar and constant; dry and overcast with a surface that had been ploughed then harrowed flat and seeded with winter wheat which had grown to a height of about 5 cm. The collection scheme was successfully implemented without any logistic problems.

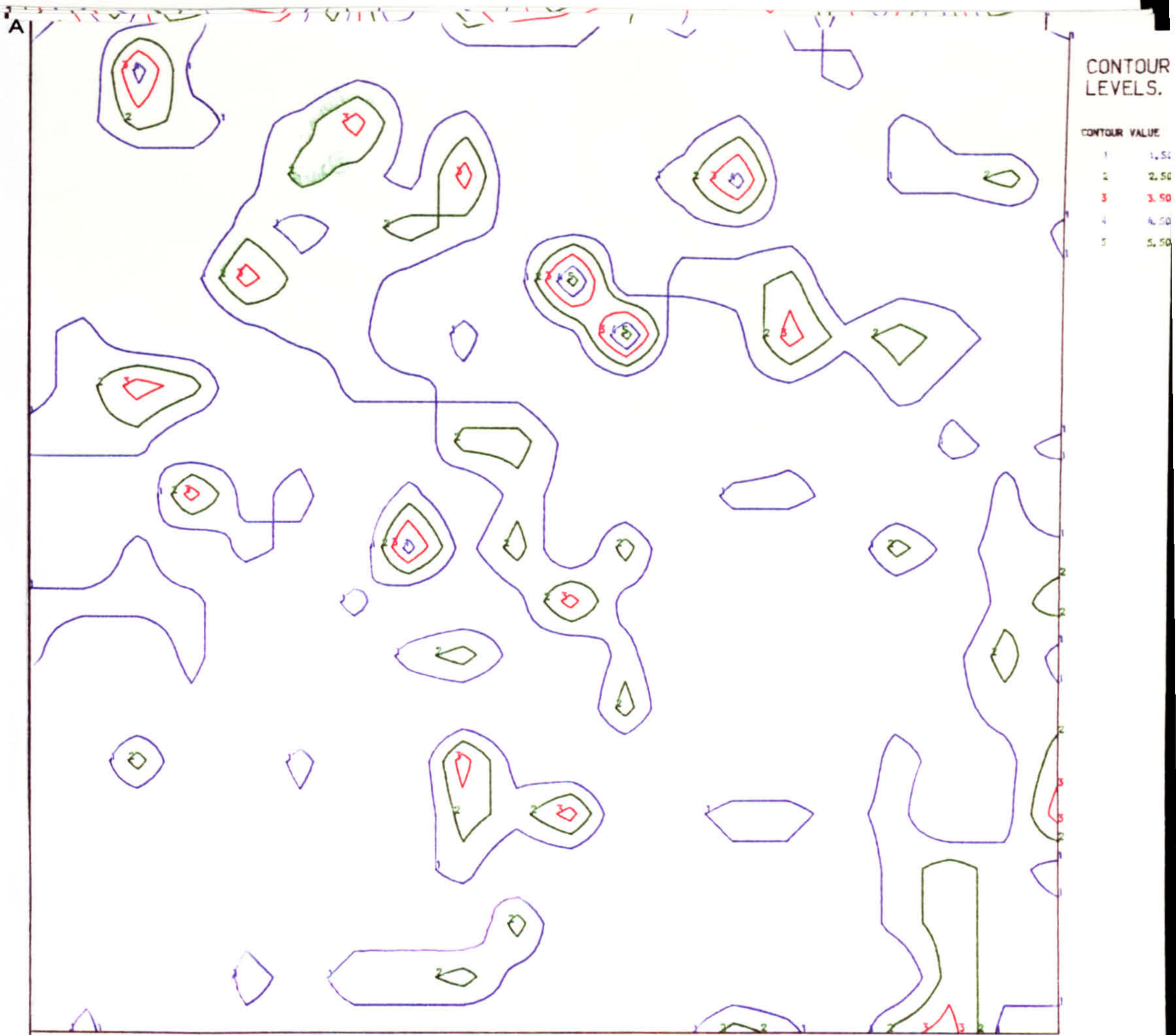
To try and measure any recovery biases for individual fieldworkers, random quadrats were intensively searched after sampling to recover any material not previously collected. However very little if any missed material was found implying that the five minutes collection time was achieving saturation collection. To try and monitor the collection performance of individual fieldworkers the proportion of rubbish (i.e. unwanted material) collected was

measured for all the quadrat samples. Although there was some variation between fieldworkers, individual fieldworkers were surprisingly consistent in their collection patterns. Although recovery biases and variations in collection performances were present, their measured values were too small for meaningful statistical analysis.

The collected material was divided into four classes; modern, cinder, tile and pottery. The modern class consisted of material of post-industrial origin, including; glass, pottery, bone, metal, tile, ceramics and plastic. The majority of this material was of domestic origin and probably introduced into the study area through night soiling or dumping. The cinder class consisted of mostly domestic fire cinders of nineteenth and twentieth century origin. Some slag was also present which was identified as iron slag through X-ray fluorescence (analysis by M. Gill), and was probably also of local nineteenth century origin. The tile class consisted mostly of Medieval and later material with some Roman tile fragments. The pottery class consisted of mainly Romano-British sherds with some later Medieval material. The Romano-British sherds were almost all late fourth to early fifth century Huntcliff type pottery (analysis by J. Evans; Corder and Kirk, 1932). A few sherds of fourth century Crambeck ware (Corder, 1928; Corder and Birley, 1937) were also present.

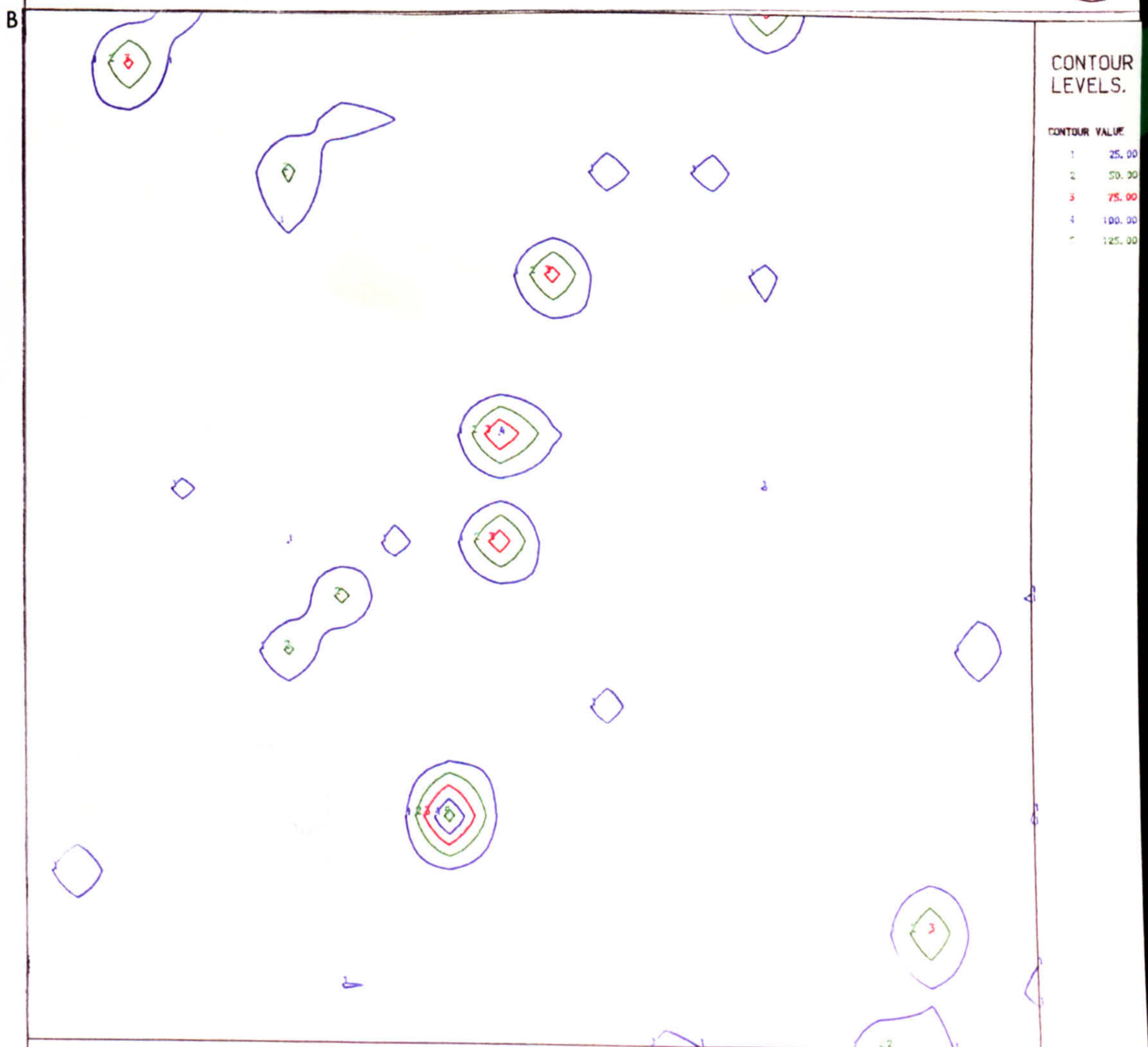
For each class and quadrat the number of pieces of material was counted and the weight of material was measured to the nearest 0.1 gm. This data was then contoured using BAGS. The resultant contour plots for the four classes of material are shown in figures 6.2 to 6.5.

Figure 6.2 Modern material by A) number B) weight.



HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES

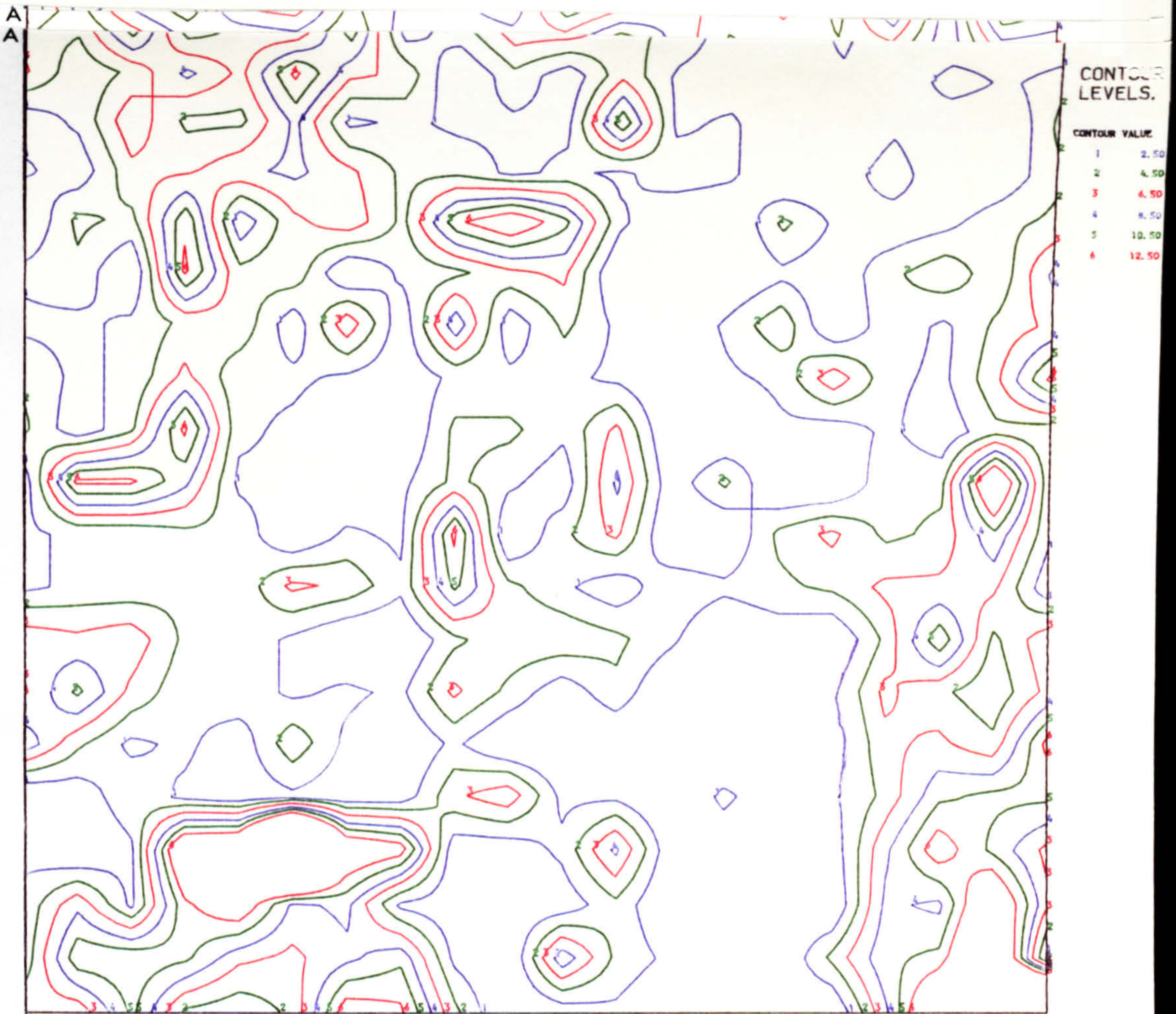


HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES

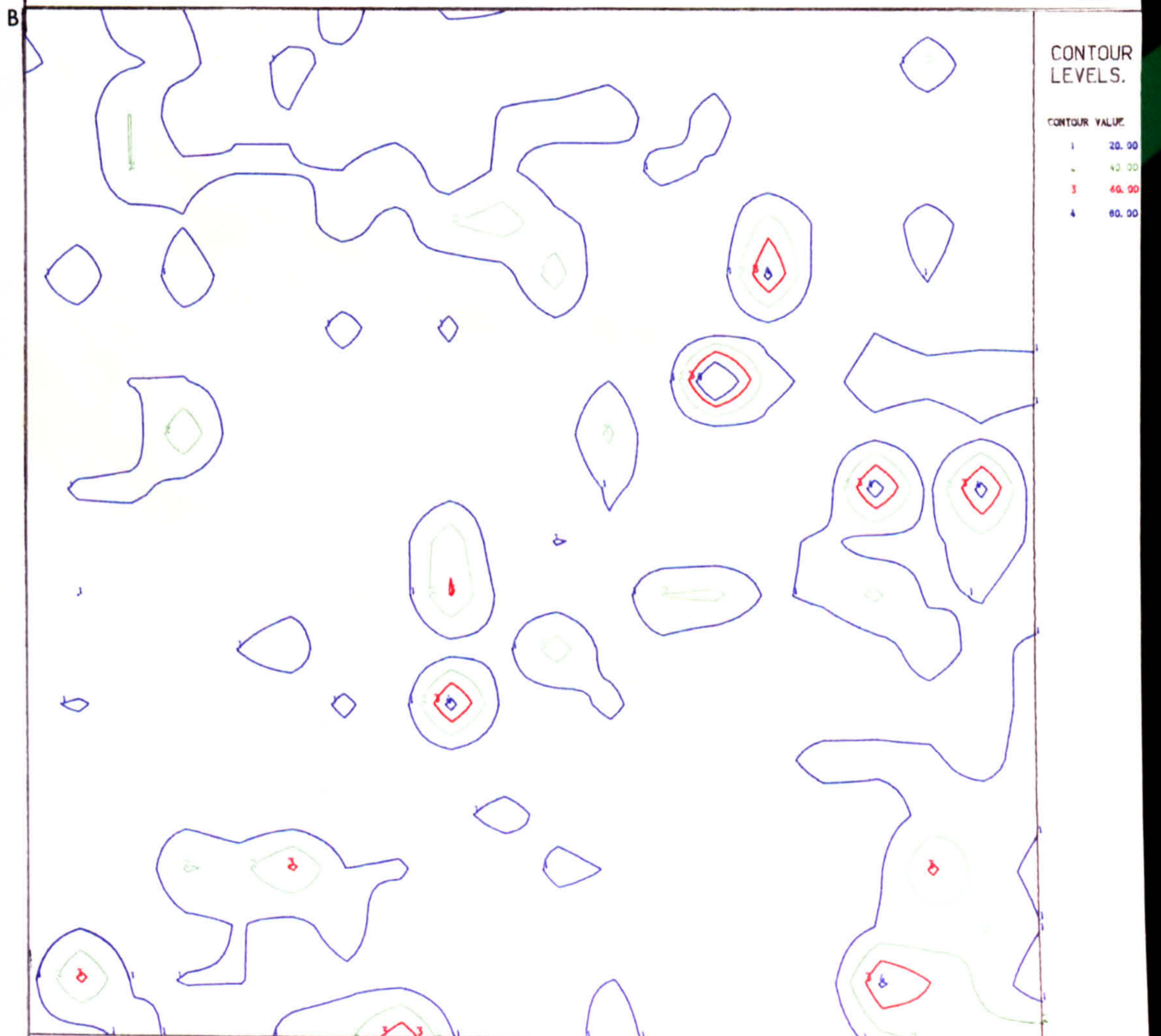


Figure 6.3 Cinder material by A) number B) weight.



HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES

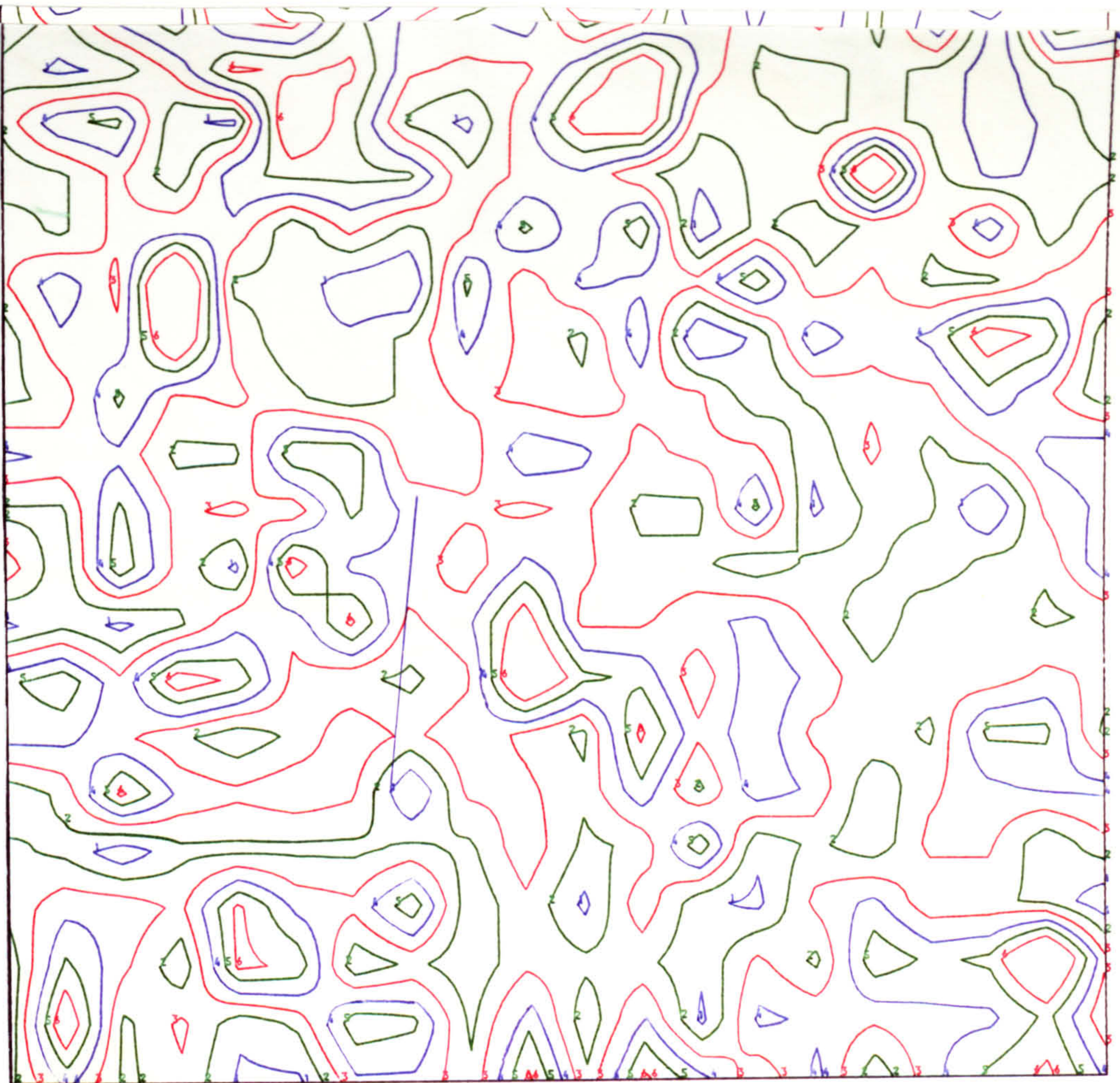


HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES



A



CONTOUR LEVELS.

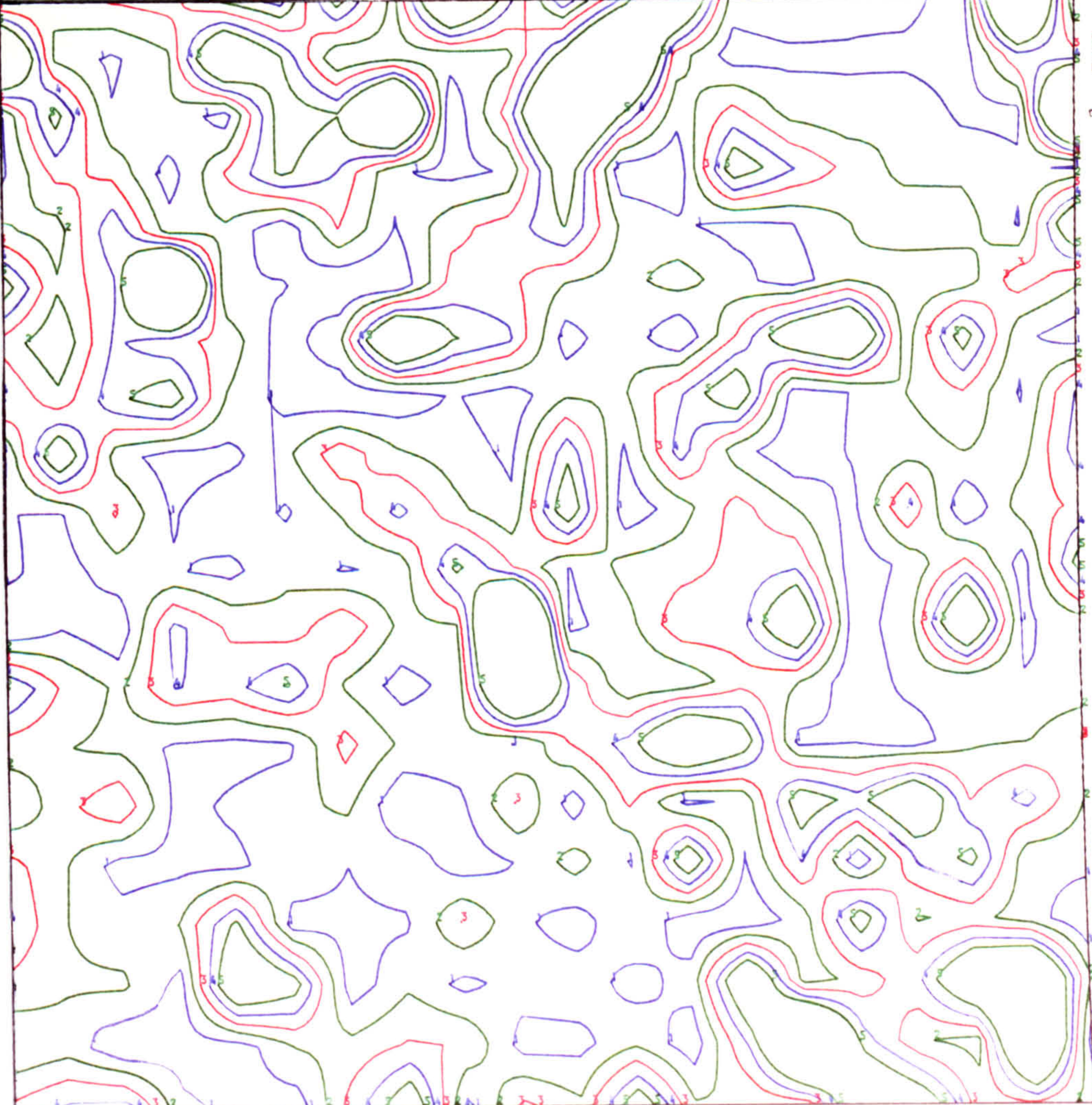
CONTOUR VALUE	1	2.50
2	4.50	
3	4.50	
4	8.50	
5	10.50	
6	12.50	

HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES



B



CONTOUR LEVELS.

CONTOUR VALUE	1	20.00
2	40.00	
3	40.00	
4	80.00	
5	100.00	

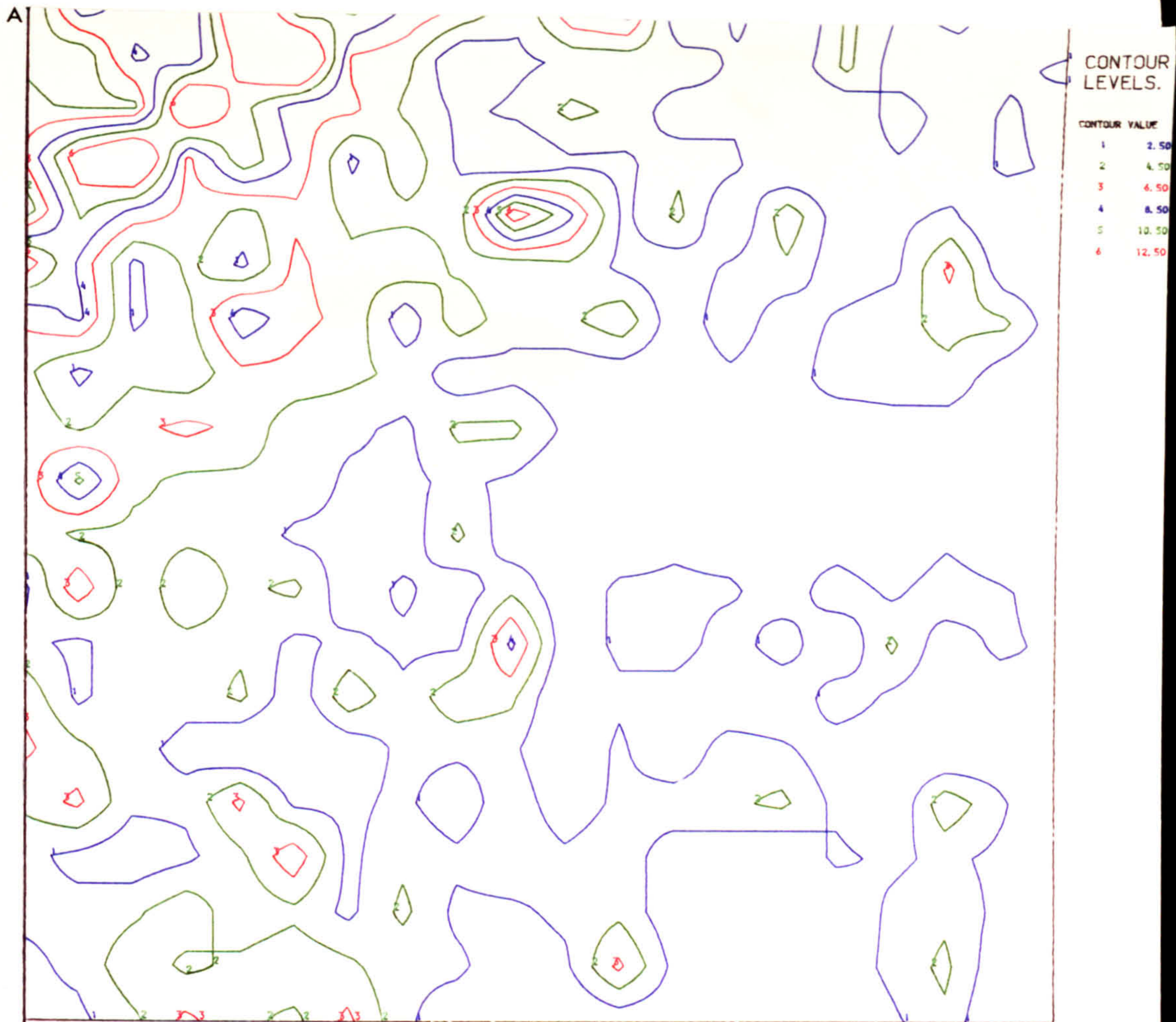
HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES



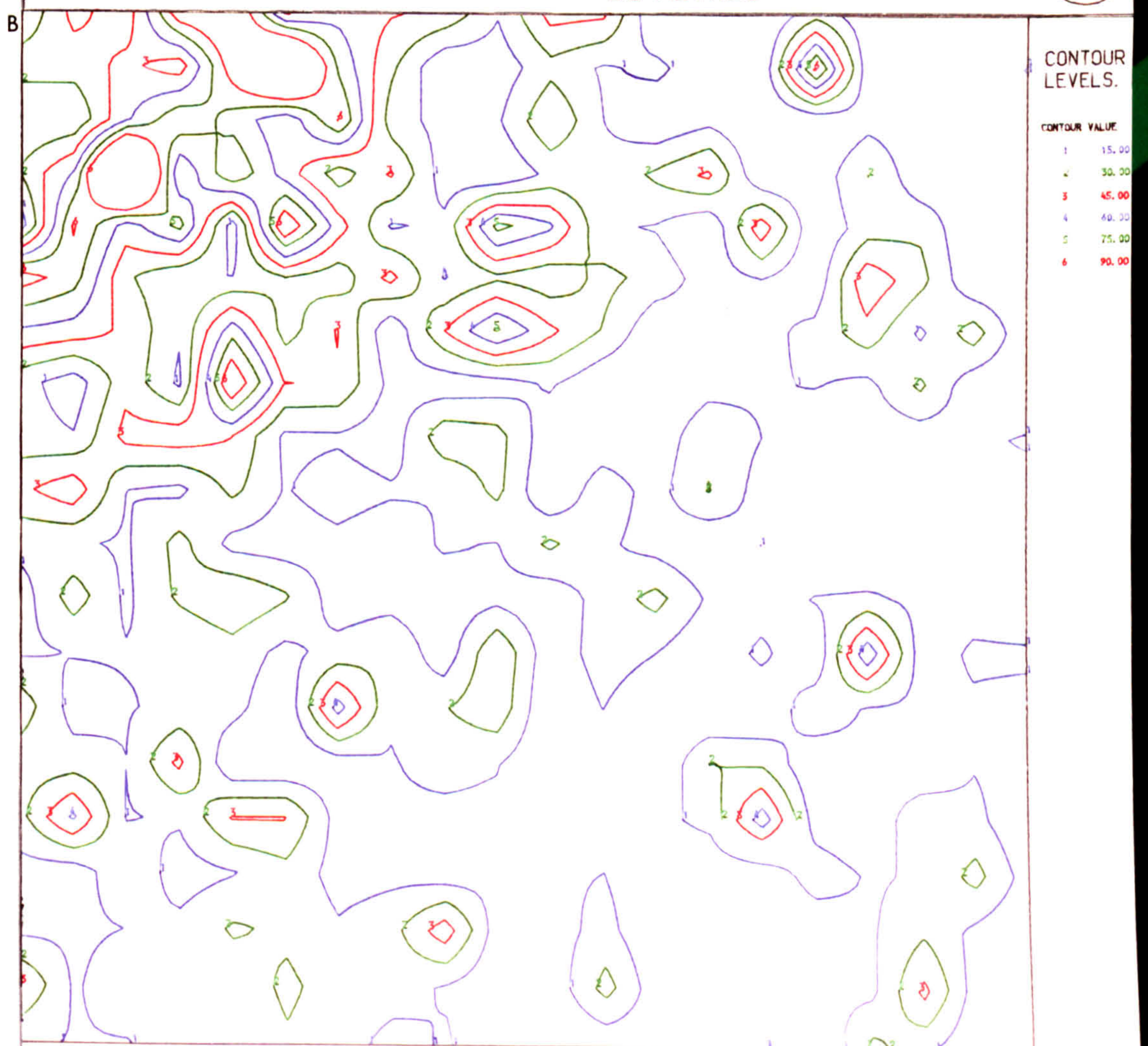
Figure 6.4 Tile by A) number B) weight.

Figure 6.5 Pottery by A) number B) weight.



HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES



HESLERTON SURFACE DISTRIBUTIONS, AUTUMN 1983.

50 METRES



6.5 DISCUSSION

The collected surface material was divided into the four classes of modern, cinder, tile and pottery to allow the modern and cinder classes to measure the effects of lateral vectors within the surface interface. Both types of material were introduced to the study area recently enough to be predominantly effected by modern agricultural practices. The material was probably originally dumped in localised areas rather than being dispersed at random over a large area. However modern agriculture would be expected to disperse these local distributions into a global random distribution. The contour plots of number and weight for the modern material (figure 6.2) shows isolated areas of localised concentration. The same pattern is seen for the cinder material (figure 6.3) as well as a grouping of material towards the east edge of the study area. This is adjacent to a field boundary which was probably used as access to the field for dumping. The surface distributions of modern and cinder material therefore seem little disturbed from their original spatial structure. To confirm this quantitatively, a simple spatial analysis of the material count data was performed, using a variance mean ratio test and a chi-squared test of goodness of fit between the observed frequency distribution and the expected poisson frequency distribution for a random spatial distribution (Fisher, 1925; Greig-Smith, 1952; Greig-Smith, 1964). The results (table 6.3) show that the spatial distributions of both modern and cinder material are significantly clustered. This confirms that modern agriculture has not unduly altered the original structure of the surface distributions of modern and cinder material, and by

Class	V/M	t	d.f.	significant at
modern	1.2677	3.7812	399	0.05%
cinder	4.1093	43.9178	399	0.05%

Class	chi-squared	d.f.	significant at
modern	15.0049	3	1.00%
cinder	576.1312	8	0.10%

Table 6.3 Spatial analysis of modern and cinder material.

implication, of the archaeological distributions of tile and pottery.

The distribution of tile, as shown by the contour plots of its number and weight (figure 6.4), displays no real discernible structure. However there is some localised concentrations that suggest that most of the material was probably dumped in the study area. This is quite a tenable suggestion as the possible date range of the Ladder settlement is earlier than the dating for the vast majority of the tile material. For this reason the tile data will not be used in any further analysis.

In contrast the contour plots of the pottery distribution (figure 6.5) shows significant structure, notably in the north west corner of the study area. However the interpretation of this structure is difficult. Concentrations of pottery sherds may relate to a variety of archaeological features such as middens, pits, ditches and occupation floors. As the only way to associate the pottery distribution structure with such features is in the context of other information, the question of interpretation will be left until part 3.

Beyond the Heslerton study the general phenomenon of surface distributions provides a conflict of interests within archaeology. Such distributions can readily be seen as the visible evidence of the destructive effects of modern arable agriculture on the archaeological heritage (Hinchliffe and Schadla-Hall, 1980). As such, these distributions are a measure of the failure to achieve proper archaeological management. But successful management requires detailed information on the extent and quality of the

archaeological resources. In many instances surface distributions will actually provide this information. Therefore the objectives of preservation and survey seem to be in conflict.

Traditional fieldwalking is essentially a site prospecting technique, in which 'significant' localised concentrations of surface material designate a site. The range and dating of the material found may often provide crude but effective period and function labels for the site. This information constitutes a basic record for any management system. However as this chapter has hopefully illustrated surface distributions may provide a much richer source of information when recorded using surface collection as a survey tool. The more detailed the survey information the more effective the management system, allowing the conflict of preservation and survey to be balanced.

7.0 AERIAL PHOTOGRAPHY

7.1 AERIAL PHOTOGRAPHY IN ARCHAEOLOGY

Aerial photography has been used in archaeology from an early date to discover and document archaeological features in the landscape (Deuel, 1969). These features are usually of a structural kind, such as banks, ditches and pits. They may relate to a wide range of types of archaeological site, from nucleated settlement to dispersed field system.

The features visible on aerial photographs are formed by a variety of processes that can indicate the presence or absence of archaeological features below the ground surface. The best known example is that of crop marks, which are formed by the differential growth of mainly cereal crops. Where the plough soil is enriched by underlying sediments from pits and ditches, crops can be advanced in growth. While where the plough soil is depleted by underlying rubble from walls and banks, crops can be retarded in growth. This differential crop growth can outline buried archaeological features and will be most visible from the air. Other examples of indicators of archaeological features below the ground include soil marks, parch marks and differential frost thawing.

The information obtained from aerial photographs may be put to a variety of archaeological uses. Perhaps the most important generic use is that of landscape survey, which involves compiling a detailed map of the archaeological landscape from the features contained on aerial photographs. Such information is often invaluable to both

the study and management of a region's archaeology. However aerial photography contains inherent problems that makes this survey task difficult.

7.2 PROBLEMS IN AERIAL PHOTOGRAPHY

The processes that make archaeological features visible to aerial photography are very sensitive to the condition of localised factors such as drift geology and soils (Evans and Jones, 1977). The physical constituents of the archaeological features also influences their visibility above ground. This makes it difficult to assess the significance of negative evidence from aerial photography. Even with careful geological and geomorphological examination of an area it can be difficult to decide whether the absence of visible evidence indicates a real absence of archaeological data. Therefore the uncritical use of aerial photography as a basis for landscape survey can produce very biased results (Palmer, 1978).

Many features visible on aerial photographs are not archaeological. For example modern constructions such as drains and pipelines can sometimes appear to be early field boundaries. Such spurious features must obviously be identified and eliminated. This discrimination can often be difficult and usually involves collaborative fieldwork and document research (Wilson, 1982). Even when securely identified as archaeological, many features yield little useful information. Morphological analysis can classify the obvious feature forms, such as barrows and enclosures, but many features are left unclassified often as the ubiquitous 'crop mark complex'.

Most aerial photographs that are used in archaeology are taken at an oblique angle. This means that the form of a feature on a photograph will be distorted with respect to its true form on the ground. Leaving features in their photographic form does not allow the easy comparison of information between different photographs or between photographs and other mapped data. It is therefore of great importance to be able to transform features from their recorded photographic form into their true ground form (Hogg, 1980).

7.3 SOLUTIONS TO PROBLEMS IN AERIAL PHOTOGRAPHY AT HESLERTON

Suitable aerial photographs that contained the Heslerton study area were found by using the North Yorkshire Sites and Monuments Record Database (NYSMRDB). This is a mainframe computer system based at Northallerton that contains detailed information on North Yorkshire's recorded archaeology, including aerial photography data (Griffiths, 1982). Access to the database is open to any archaeologist on an informal basis through the county unit. A radial search facility was used on the NYSMRDB to locate any aerial photographs taken near the Heslerton study area. The search was centred at SE 949 770, a point contiguous to the study area, with a radius of 250 metres. Fourteen photographic records were found, but when consulted only one photograph (reference TP 353/5) showed clear features in the study area. These are shown in figure 7.1.

As the size of the study area is relatively small (one hectare) the problem of differential response of aerial photography can be assumed to be negligible. This assumption is strengthened by the fact that fields adjacent to the study area show a clear

continuation of crop marks defining the latter...
features visible within the study area...
archaeological origin. A careful ground examination...
found no evidence of recent constructions or activities...
have left spurious features. A discussion with...



Figure 7.1 Original aerial photograph (TP 353/5) with crop marks local to Heslerton study area.

continuation of crop marks defining the ladder settlement. All the features visible within the study area seem to have an archaeological origin. A careful ground examination of the area found no evidence of recent constructions or activities that could have left spurious features. A discussion with the landowner confirmed these observations.

The transformation of features from their recorded form on oblique aerial photographs to their true ground form can be achieved by several manual graphical methods (Scollar, 1975). These techniques although relatively easy to use can become inaccurate for fine detail and very slow for large amounts of data. An alternative is to use a computer based system to increase both accuracy and data handling capacity. Such a system based on a microcomputer (Chamberlain and Haigh, 1982) was used to transform the aerial photograph for the Heselton study area.

The system is based on the work of Palmer (1976, 1977). It assumes that the ground form of a feature is defined on a plane, a condition that the Heselton study area approximates well. Equivalent control points are identified on both the aerial photograph and on a map of the area covered by the photograph. The coordinates of these control points are measured using a digitising pad and are then used to calculate a geometrical transformation to transform points from the photograph plane to the map plane. Using this transformation features are transformed by digitising their outlines from the photograph as a sequence of points (figure 7.2) and transforming these onto the map. The resulting transformed features can then be drawn out using a plotter.

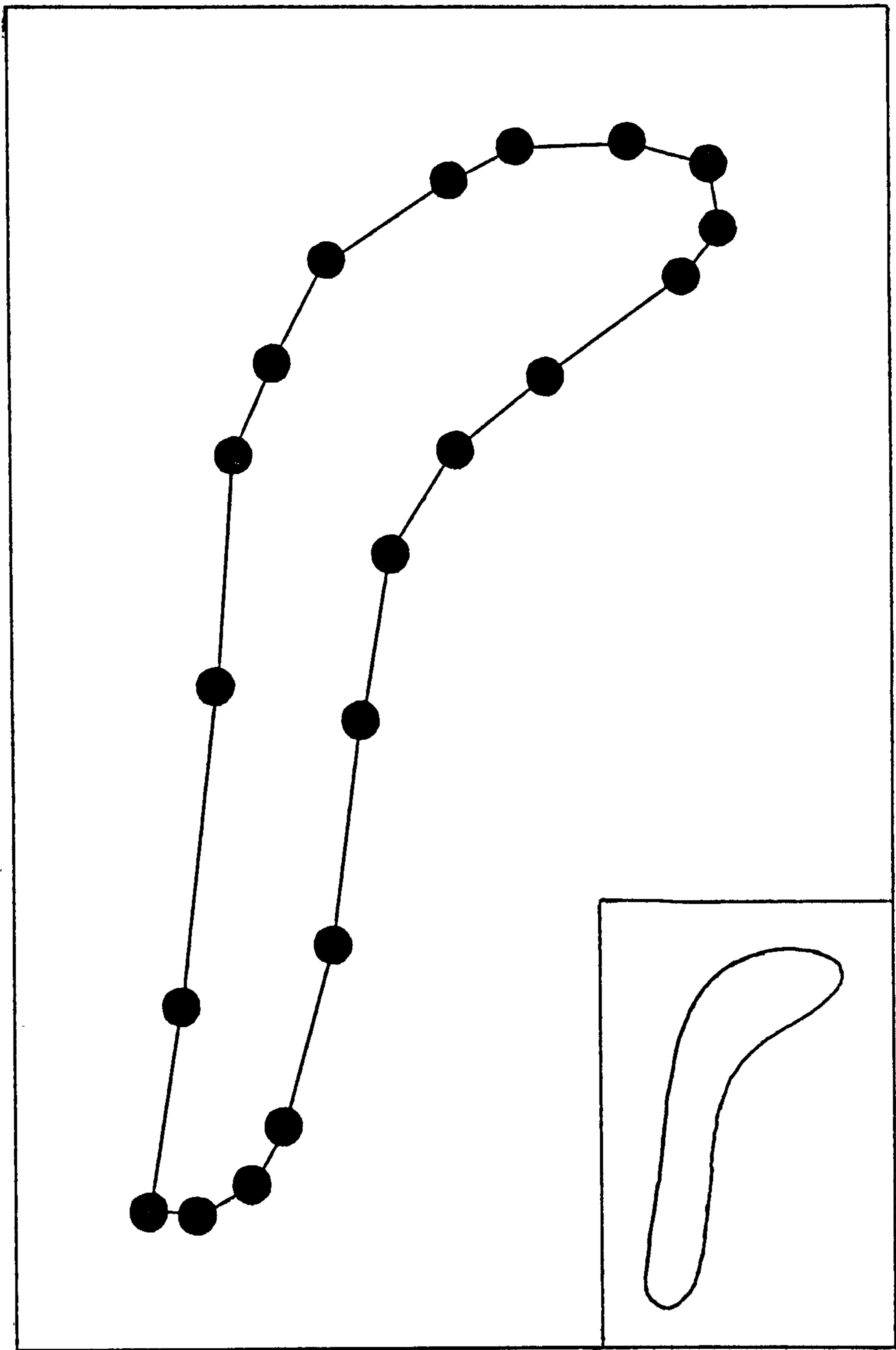


Figure 7.2 Digitisation of a feature.

The accuracy of the transformed data depends on three factors:

- 1) accurate co-location of control points
- 2) ground equivalence to a plane
- 3) accurate recording of data from the aerial photograph

At best the system should transform points to an accuracy of one metre. In practise a working tolerance of two to three metres is more realistic. More powerful computer based systems are emerging that use direct video input of aerial photographs (e.g. Scollar, 1978; O'Brien et al, 1982). These can give greater accuracy and reliability as features can be defined more objectively through the techniques of image processing. However these systems are still under development and are not generally available.

The features defined in the aerial photograph for the Heslerton study area were transformed into their map form using the microcomputer system with the following qualifications. All features were defined as solid shapes rather than simple lines so they had an area. A feature consisted of a sequence of coordinate points such that:

- 1) When each point was joined to its successor by a line segment and the last point was joined to the first point by a line segment, a closed polygon was formed.
- 2) For the vector from any point to its successor (including the last and first points), the inside of the feature lay to the right of the vector and the outside of the feature lay to the left of the vector.

The explanation for these qualifications is left until 8.2. The transformed photograph for the Heslerton study area is shown in figure 7.3.

7.4 DISCUSSION

From an archaeological viewpoint the crop marks within the Heslerton study area do not define many outstanding features. To the east of the study area (figure 7.3) there is a feature that is possibly part of a minor trackway running west to east. This interpretation is strengthened by the fact that other minor trackways with a similar orientation are clearly visible in surrounding fields. In the adjoining eastern field there is a major trackway running north to south with coincident pit alignments. A minor trackway running west to east joins it at its southern end. In two fields to the west there are clear crop marks indicating minor trackways running again west to east, one of which passes to the south of the study area.

The most interesting crop mark lies in the north west corner of the study area, it is a semi-circular feature with a possible internal pit. This could possibly represent a hut circle or even a ploughed out barrow. The remaining crop marks are less distinct and many could represent the edges of enclosures.

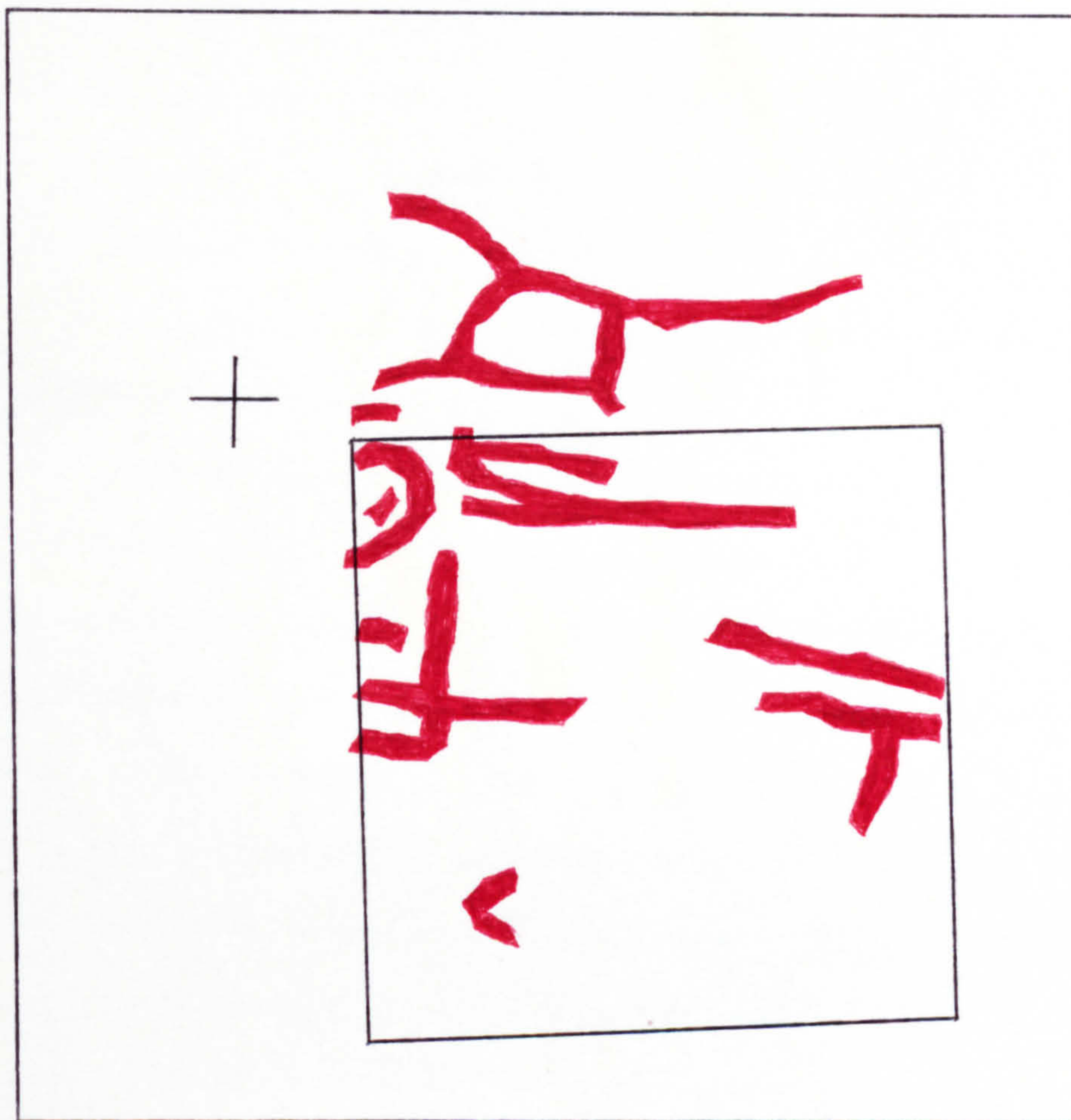


Figure 7.3 Transformed aerial photograph (TP 353/5) with
Heslerton study area outlined.
Scale 1:1500.
Reference cross at SE 94645 77505.

PART 3

8.0 SITE ASSOCIATION OPERANDS

8.1 OPERANDS AND OPERATORS

Archaeology can be viewed as a collection of data with a set of procedures that act upon this data. For example in part 2 several physical procedures for the collection of geophysical, surface distribution and aerial photography data were described. Other more intuitive procedures of interpretation were also employed. Most archaeological procedures produce data for other procedures to act upon. The relationships thus formed between procedures define a complex structure, within which the flow of data from procedure to procedure is very similar to a control system (Barnett, 1975). Procedures themselves may often be defined in terms of simpler more fundamental sub-procedures. For example the procedure of geophysical data collection of 5.0 can be defined as a series of four sub-procedures; marking the sampling point on the ground, positioning the instrument onto the sampling point, measuring the earth resistance, and recording the measurement. To emphasise this structuring of procedures, data can be regarded as operands and procedures as operators. Simple operands and operators can then be combined to form the more complex data structures and procedures of archaeology.

Association analysis is a procedure and can therefore be thought of as an operator. All association operators have one major characteristic in common, they all operate on more than one operand. When the operands are of the same type it is relatively easy to

apply an association operator. But with, 'mixed' operands simple association operators cannot be used. However the use of more complex operators is nearly always avoided and naive 'eyeball' techniques are used instead (Clark, 1983), so that the results may be expected to be very subjective and often trivial. An alternative approach is to convert or transform all the operands into a compatible type. A simple association operator can then be applied.

8.2 TRANSFORMS OF OPERANDS

For the Heslerton study the geophysical and surface distribution data have equivalent data structures, a continuous function sampled on a systematic grid (see section 4.1). However the aerial photography information of 7.0 is defined in terms of features and therefore needs to be transformed into an equivalent gridded form if a simple association operator is to be applied. These features obviously have a complex data structure that cannot be transformed in its entirety. Therefore a set of simplified aspects that can be realistically transformed and which are archaeologically relevant need to be chosen and explicitly defined.

The first intuitive aspect of the aerial photography information that could be measured, is the amount or density of information present, as some areas will obviously have a higher density of archaeological features than other areas. The density (as in surface distributions) will be a continuous function, that can be systematically sampled to provide a gridded data set. But how is this density to be defined? It could be defined as the number of features per unit area. However this is somewhat problematic as the

separation of the information into individual features is usually arbitrary. For example a small linear segment, or several linear segments that form a rectangle, or even several contiguous rectangles enclosed by a perimeter could all be classed as features.

A more objective measure of density is one defined on some global physical aspect of the features, such as their area. Given a sampling quadrat (figure 8.1.a) the density of features within it, based on area, will be defined as:

$$\frac{\text{area of features}}{\text{area of quadrat}}$$

This density will theoretically range between zero, when there are no features in the quadrat, and one, when the quadrat is totally enclosed by a feature. For this density measure to be applicable the restriction must be imposed that all features have an area and are not simply defined as lines. This is the reason for the qualifications added to the digitisation of the aerial photography features in section 7.3. All the recorded features have an area and have a data structure that explicitly defines the inside and outside of a feature.

Most features will lie in several quadrats (figure 8.1.b), so that a method is required to select those parts of a feature in a particular quadrat. As the features are defined as directional polygons a simple polygon clipping algorithm will do this (Sutherland and Hodgman, 1974). The algorithm recursively clips each line segment in a polygon to the four sides of a quadrat, producing a clipped polygon completely contained by the quadrat.

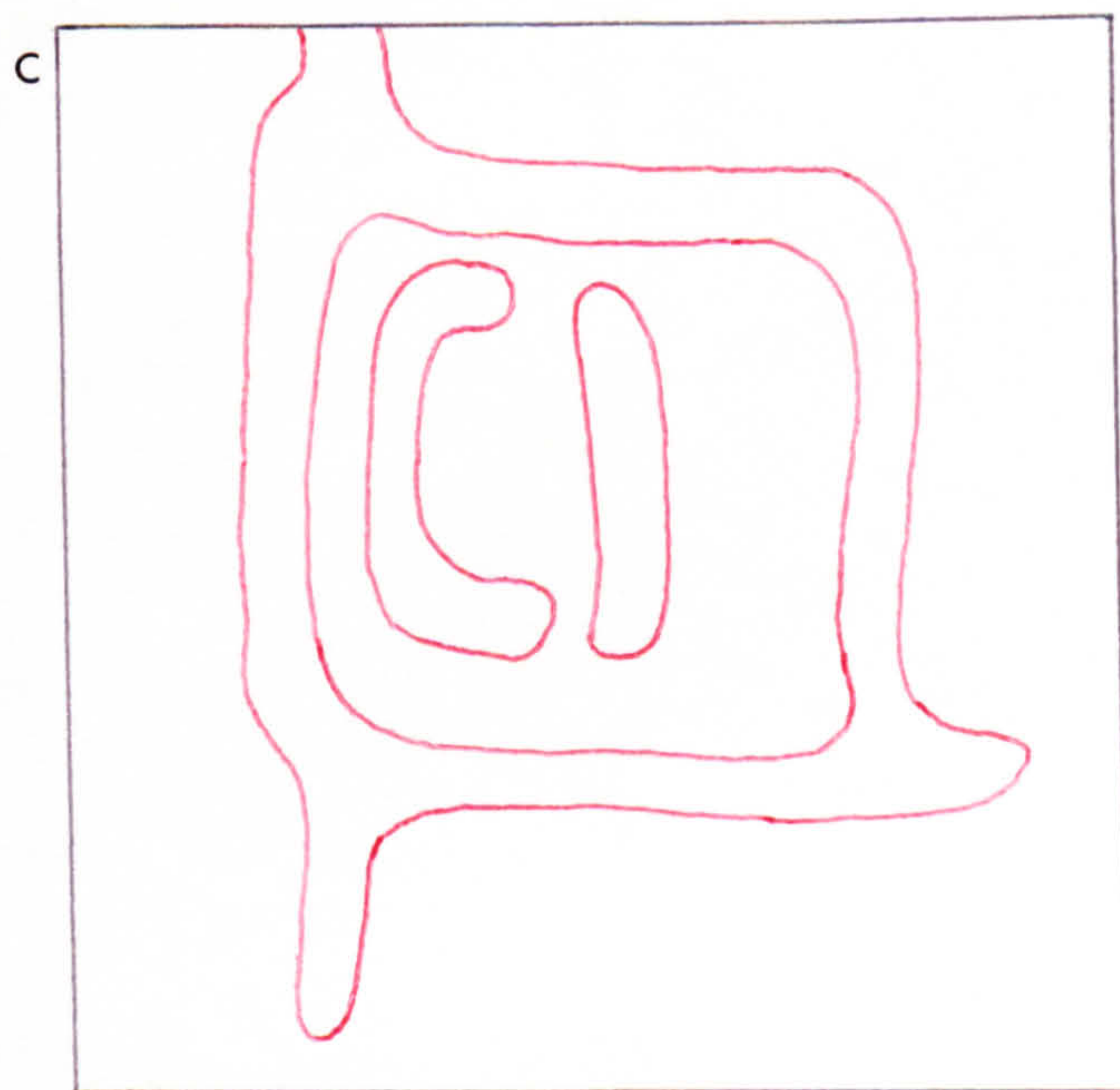
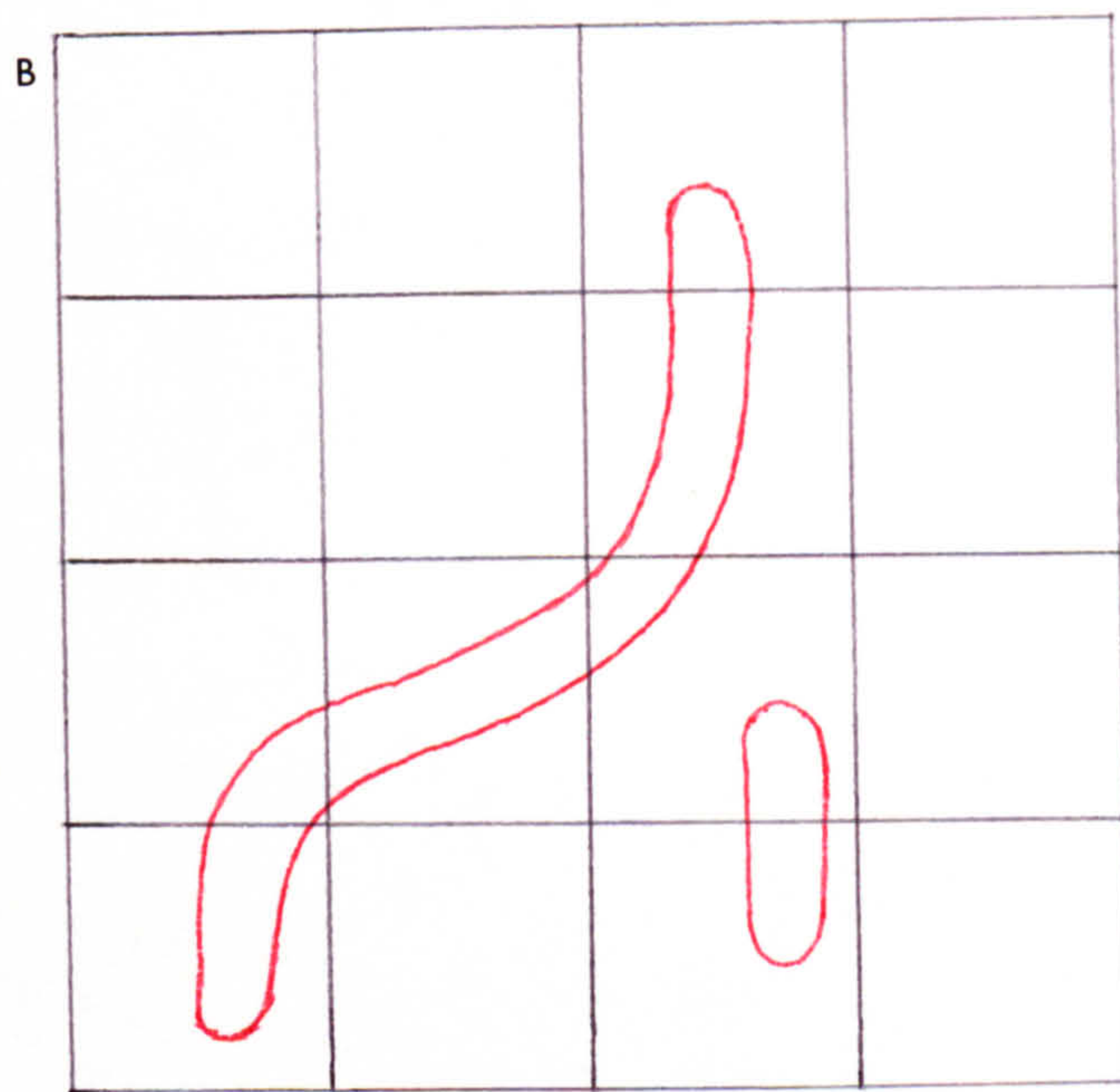
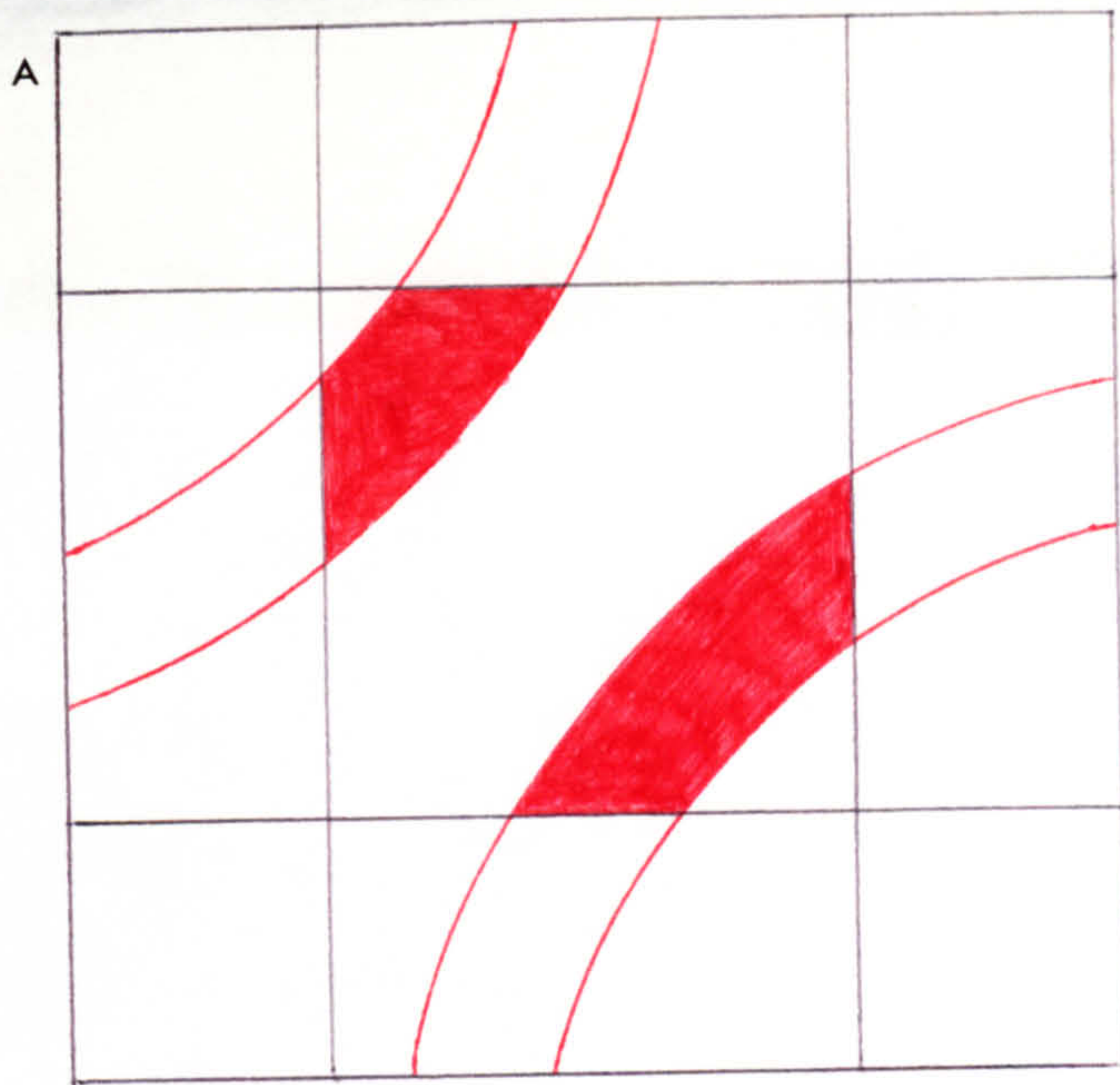
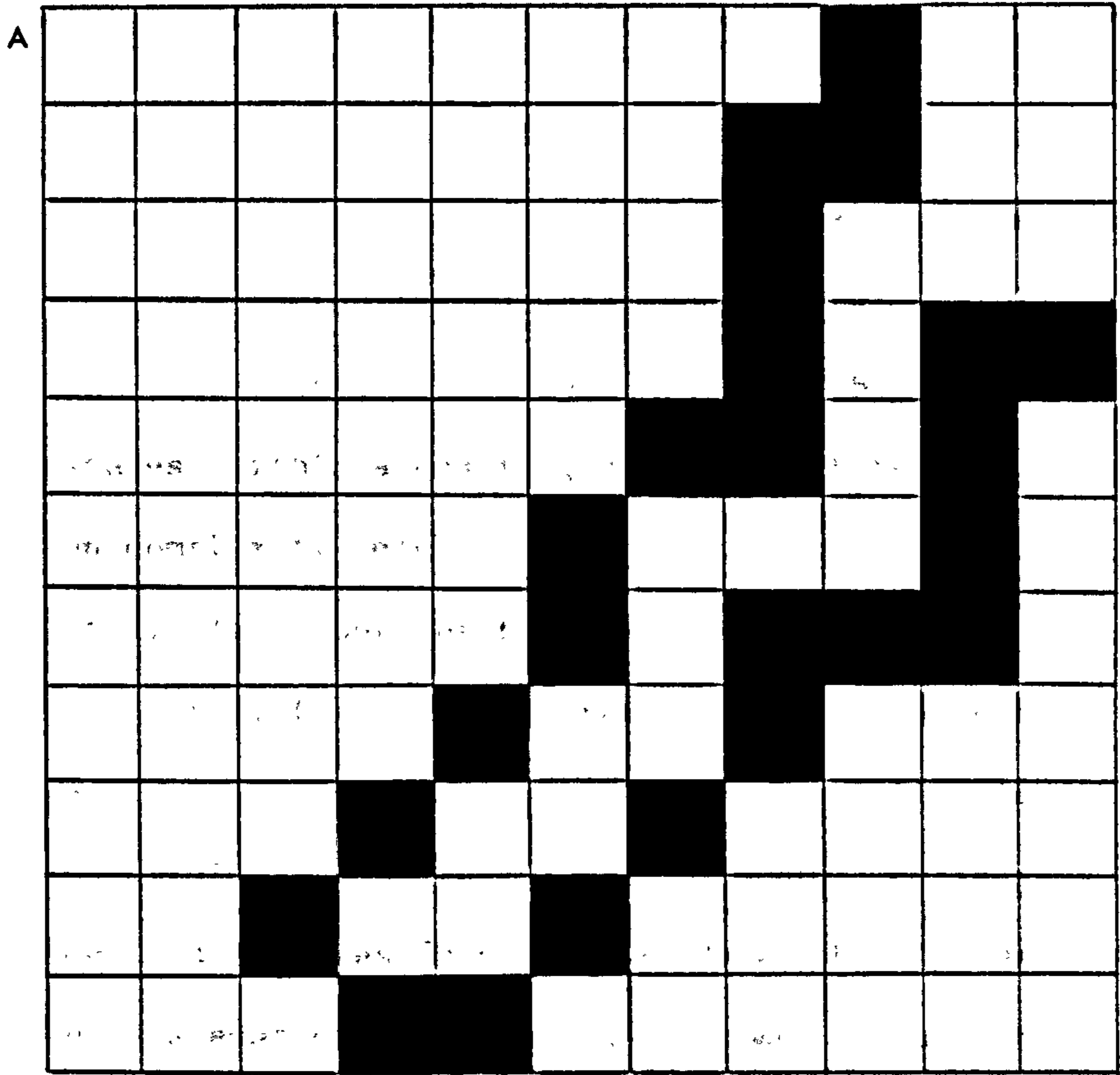


Figure 8.1 Feature density. A) Feature area. B) Feature extent. C) Nested features.

Having clipped the features within a quadrat their areas now need to be determined. Several methods exist for the measurement of a polygon's area given a point inside the polygon. However although these methods are generally satisfactory, even for re-entrant polygons, they are difficult to implement where several polygons are nested (figure 8.1.c). An alternative approximate method that does work with nested polygons, and has an advantage for other measurements as shown later, is based on the concept of raster computer graphics. The quadrat is divided into a large number of cells or pixels. The clipped polygons representing features within the quadrat are converted to pixels (figure 8.2.a) using a simple digital differential analyser (Newman and Sproull, 1981). Seed points lying within the features are calculated from the directional information for the polygons and are recorded as pixel coordinates. When all the feature outlines have been converted to pixels these seed pixels are used to initiate a flood fill algorithm (figure 8.2.b) for a boundary defined four connect region (Pavlidis, 1982). The density measure is then simply defined as:

$$\frac{\text{number of pixels set}}{\text{total number of pixels in quadrat}} \cdot$$

Problems can arise with this approximate method. The pixel size must be small enough to distinguish the minimum distance between any two feature outlines. In practice given the maximum resolution of the aerial photography transformation of section 7.3, the pixel size can be set to an order of magnitude lower. The method will not work if a quadrat is totally enclosed by a feature, but as most features are linear in form a reasonably large quadrat size should be sufficient to overcome this problem.



Approximate feature density. A) Pixel outline. B) Pixel area.

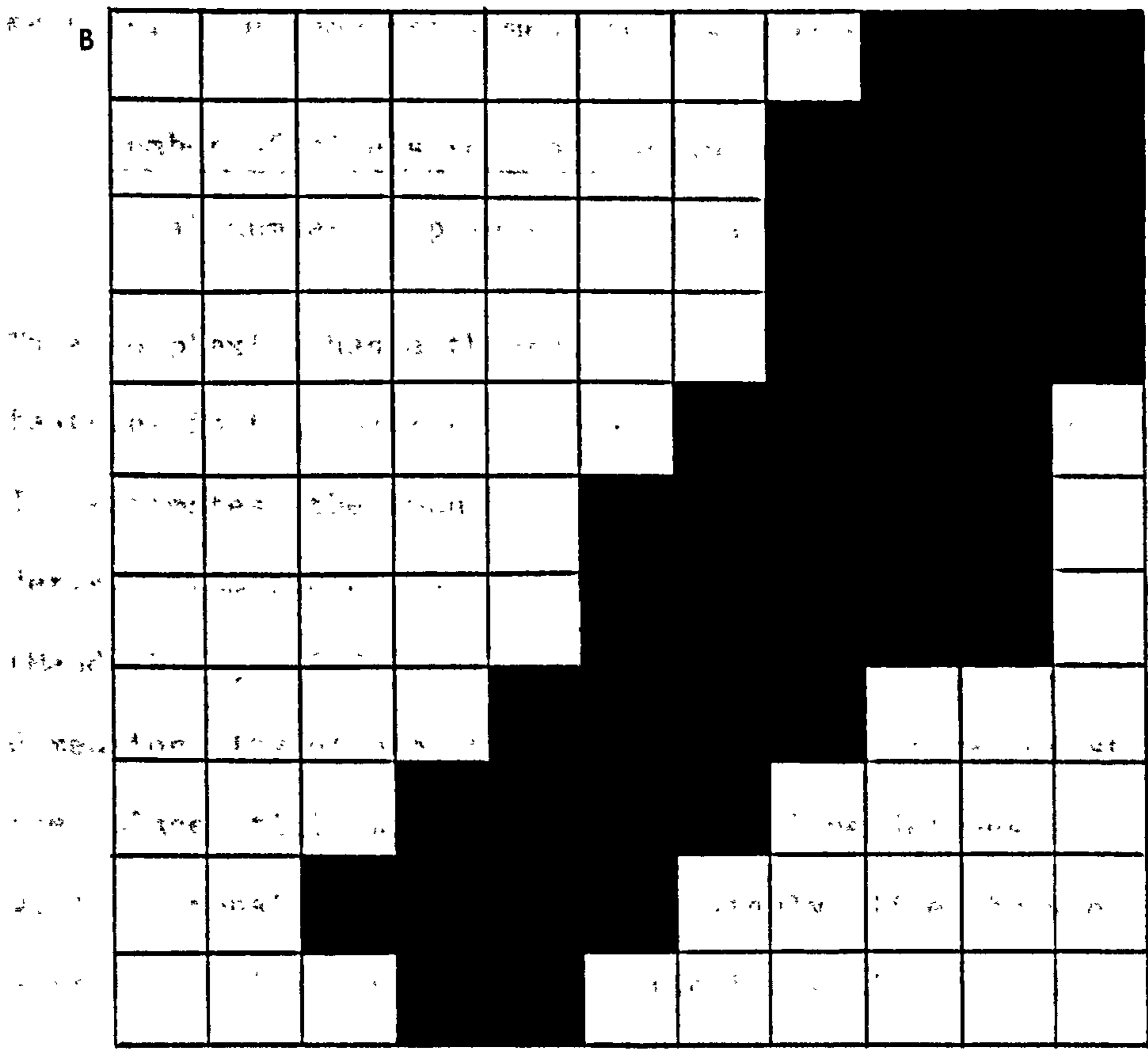


Figure 8.2 Approximate feature density.
 A) Pixel outline. B) Pixel area.

The situation could arise where two quadrats have the same density but have a very different pattern of features within them. Another measure would therefore be desirable to differentiate them. This measure could record the complexity of features within a quadrat. Complexity, as for density, could be defined as the number of features within a quadrat, or different types of features could be given complexity weightings and a weighted average could then be calculated. Again both of these measures are problematic due to the arbitrary definition of features. A more objective complexity measure could be defined as the total outline length of all the features within a quadrat. This measure defines a continuous function that ranges from zero to theoretically infinity and which can be systematically sampled. However a more manageable approximate complexity measure is one based on the pixel representation of the feature outlines as described for the density measure. The complexity measure now equals:

$$\frac{\text{number of pixels set for outlines}}{\text{total number of pixels in quadrat}}$$

This complexity has a theoretical range from zero, when there are no features in the quadrat, to one, with infinitely complex features. It estimates the amount of feature outline wiggle in a quadrat and derives from the concept of fractals or fractional dimensionality (Mandelbrot, 1982; Sorensen, 1984). Here a straight line has a dimensionality of one, but if you wiggle the line you start to fill the plane till an infinitely wiggly line defines a complete plane with a dimensionality of two. A wiggly line between these two extremes will therefore have a fractional dimensionality between one and two.

Other measures could be defined for the aerial photography data such as the density of linear elements, or circular elements, or even feature intersections. But the density and complexity measures seem to define the two most important aspects of the aerial photography data; how much information is there and how structured it is.

These measures of density and complexity have been implemented in a Pascal program `apgrid`, whose structure is outlined in appendix A.4. The program reads point coordinate data, as produced by the digitising system, that has been structured as indicated in section 7.3. The quadrat grid is then defined and the density and complexity measures calculated. These are then outputted as matrices of values.

For the Heselton study area an aligned grid of 20 X 20 quadrats was used, each quadrat being 5 metres square. Quadrats were then divided into a mesh of 20 X 20 pixels, each 25 centimetres square. The resulting density and complexity matrices are presented as contour plots in figures 8.3 and 8.4. They both define three areas of crop marks within the study area; a large area to the north west, a medium area to the east and a small area towards the south west. Both the density and complexity contour plots are very similar in appearance and structure, mainly because in this particular instance the feature forms are very simple. For this reason only the density measure will be used in any further association analysis.

Having achieved a compatible operand type for all three data sets in the Heselton study, only the problem of different sampling densities remains. Both the surface distribution and aerial

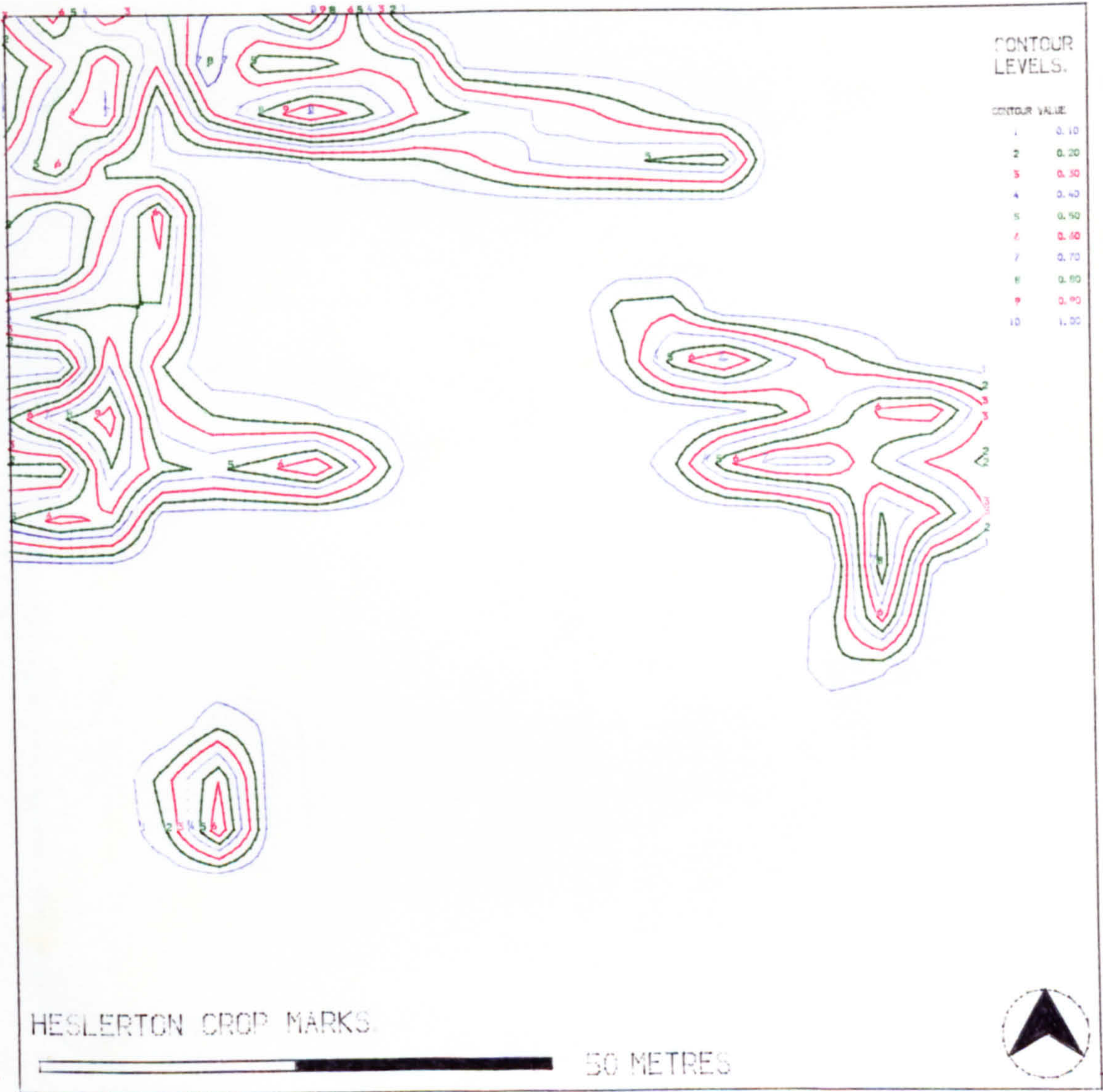


Figure 8.3 Feature density.

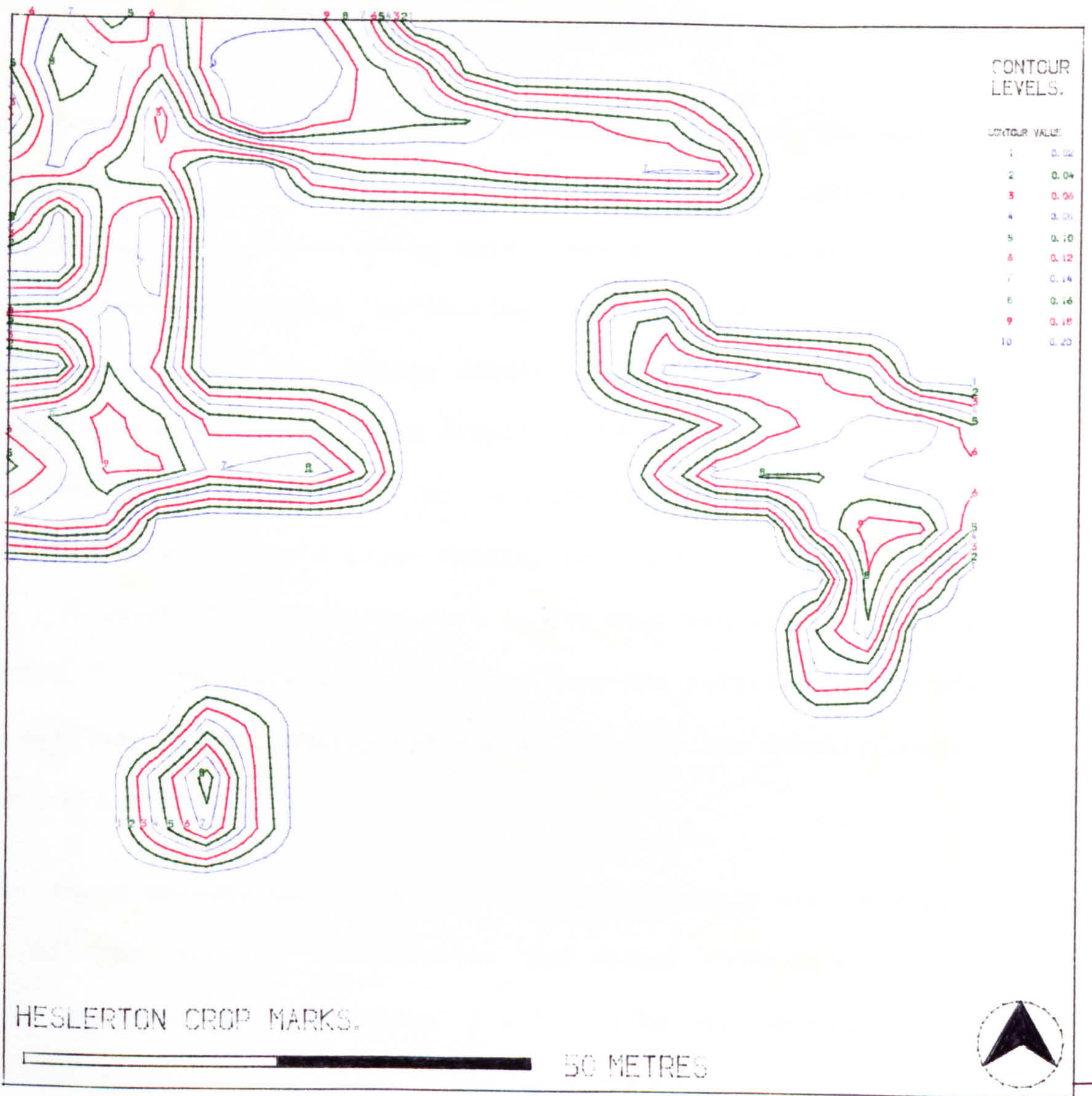


Figure 8.4 Feature complexity.

photography data have been sampled at 5 metre intervals, while the geophysical data has been sampled at the more intensive 1 metre interval. The sampling densities of the operands need however to be the same if any simple association operator is to be applied. This leaves two alternatives, convert the operands either to the highest sampling density or to the lowest sampling density.

In converting to the highest sampling density, new sampling points must be established for the low density operands by interpolating from the existing data. Any form of interpolation is liable to produce spurious patterning at the greater levels of detail created by the higher density. In contrast by converting operands to the lowest sampling density only details of the existing patterning will be lost. The amalgamation of contiguous sample point values to produce a lower density can be done in a variety of ways. However the simplest, such as the mean and total value, are probably the most satisfactory as they keep the relationship between the original high density operand and the new low density operand clear and concise.

For these reasons the geophysical sampling density was reduced to that of the surface distribution and aerial photography data by taking the mean of contiguous 5 X 5 blocks of readings. The resultant geophysical data are shown as contours in figure 8.5., and will be used in any further association analysis.



Figure 8.5 Low density resistance data.

8.3 INVERSE TRANSFORMS OF OPERANDS

The ability to transform operands of different structure to a common form is, as already discussed, extremely useful if not essential for association analysis. But this also raises the interesting question as to whether it is possible to define an inverse transform that will change an operand from a common form back to its original structure. To try and answer this question the geophysical data of 5.0 can be considered. Here the original operand has the basic data structure of section 4.1, but a transform can be envisaged to define features as is intuitively done in interpretation (5.7). This transform can be regarded as an inverse transform as the change from features to systematically sampled function was used for the aerial photography data in the previous section. However the use of the term inverse transform is challengeable, as in mathematics a transformation has a precise definition, with a transform and its inverse transform having a defined relationship. In the transform from aerial photography features to density measure, and as will be shown, in the inverse transform from geophysical readings to features, no such relationship exists. Given one of the transforms you cannot deduce the form of the other transform. However even though the archaeological use of the term is not precisely correct, it is still useful in conveying the nature of the transformation process. Therefore the inverse transform will be referred to as a quasi-inverse transform.

The spatial operations of feature extraction and classification in image processing offer a good starting point for the definition

of a suitable quasi-inverse transform. In one form of classification known as pattern recognition, defined feature forms are searched for within the data. For geophysical data this is an inappropriate technique as the universe of possible feature forms is so large that searching for each within the data, even using heuristic methods (Rich, 1983), would be impossible. A more suitable technique lies in feature extraction, where the data are partitioned into significant divisions which can then be equated with geophysical features. However the determination of significance is complex as feature extraction is usually based on a model-guided partition procedure (Rosenfeld and Kak, 1982).

A simple model could divide the data into partitions that belong either to high resistance features (H), undisturbed ground (U), or low resistance features (L); each of which is assumed to be represented by a single unique earth resistance value, respectively h , u and l . This model will be called the simple HUL model. Data values obtained from surveying would be expected to equal h , u or l , in reality they do not as random variations in measurements will give a distribution of readings around these values. If the 'law of errors' is assumed to apply then these distributions will be Normal (Mood, Graybill and Boes, 1974). Thus:

$$H = N(\mu_H, \delta H),$$

$$U = N(\mu_U, \delta U),$$

$$L = N(\mu_L, \delta L),$$

with the means of these distributions equal to h , u and l .

The station survey points, as indicated in section 5.2, form a systematic sample. In another sense however they form a random

sample, as repeated resurvey of a single station will give varying measurements due to changes in soil conditions and errors in relocating the probes. Thus the H, U and L distributions could be viewed as populations from which samples of station readings have been randomly drawn. The problem of feature extraction can now be seen as the definition of groups of station readings that can be shown to be samples from H, U or L.

If the only population parameter known or estimated is μ_U then the following algorithm can be used to extract features. A statistical significance level (α) and a minimum seed configuration of stations is decided upon. The data is then searched for significant seeds for H. For each seed found, contiguous station values are added recursively until the group reaches a level of non-significance for H. The algorithm is then repeated for L.

As the seed configuration is assumed to be a random sample from a Normal population with known mean μ_U and unknown variance, the t-test is applicable for determining significance (Dixon and Massey, 1983). If a seed configuration contains N stations and has a mean \bar{x} and standard deviation s, then:

$$t = \frac{(\bar{x} - \mu_U)}{\frac{s}{\sqrt{N}}},$$

will have Student's t-distribution with N-1 degrees of freedom. For H if:

$$t > t_{1-\alpha}(N-1),$$

then statistically at the 100 α % significance level this sample

belongs to H. For L if:

$$t < t_{\alpha}(N-1),$$

then statistically at the 100 α % significance level this sample belongs to L. The same test procedure can be used to test the significance of the recursively formed station value groups.

Using a minimum seed configuration of two adjacent stations the simple HUL model for feature extraction has been implemented as a Pascal program `f_extract`. An outline of the program structure is given in appendix A.5. Figure 8.6 shows the results of running `f_extract` on the Heslerton survey data. The significance level used was 5% and μ_U was estimated as equal to the mean of all the resistance readings for the survey. The majority of the features extracted seem to relate to the ploughing striations but some seem to have a possible archaeological interpretation. There also seems to be very little undisturbed ground. This is probably because the contiguous grouping of station values does not in fact form a proper random sample because of the non-independence of contiguous readings. The measured variance is therefore an under-estimate of the true variance which has led to the dichotomy effect shown in figure 8.6.

However these results should not be considered too critically as the simple HUL model used is indeed very simple and has mainly been used to illustrate a more general approach. In reality high and low resistance features will not be characterised by single values but by a range of values. Individual features themselves may show considerable intrinsic variations in value. A more realistic model would therefore need to be much more complex, but would be

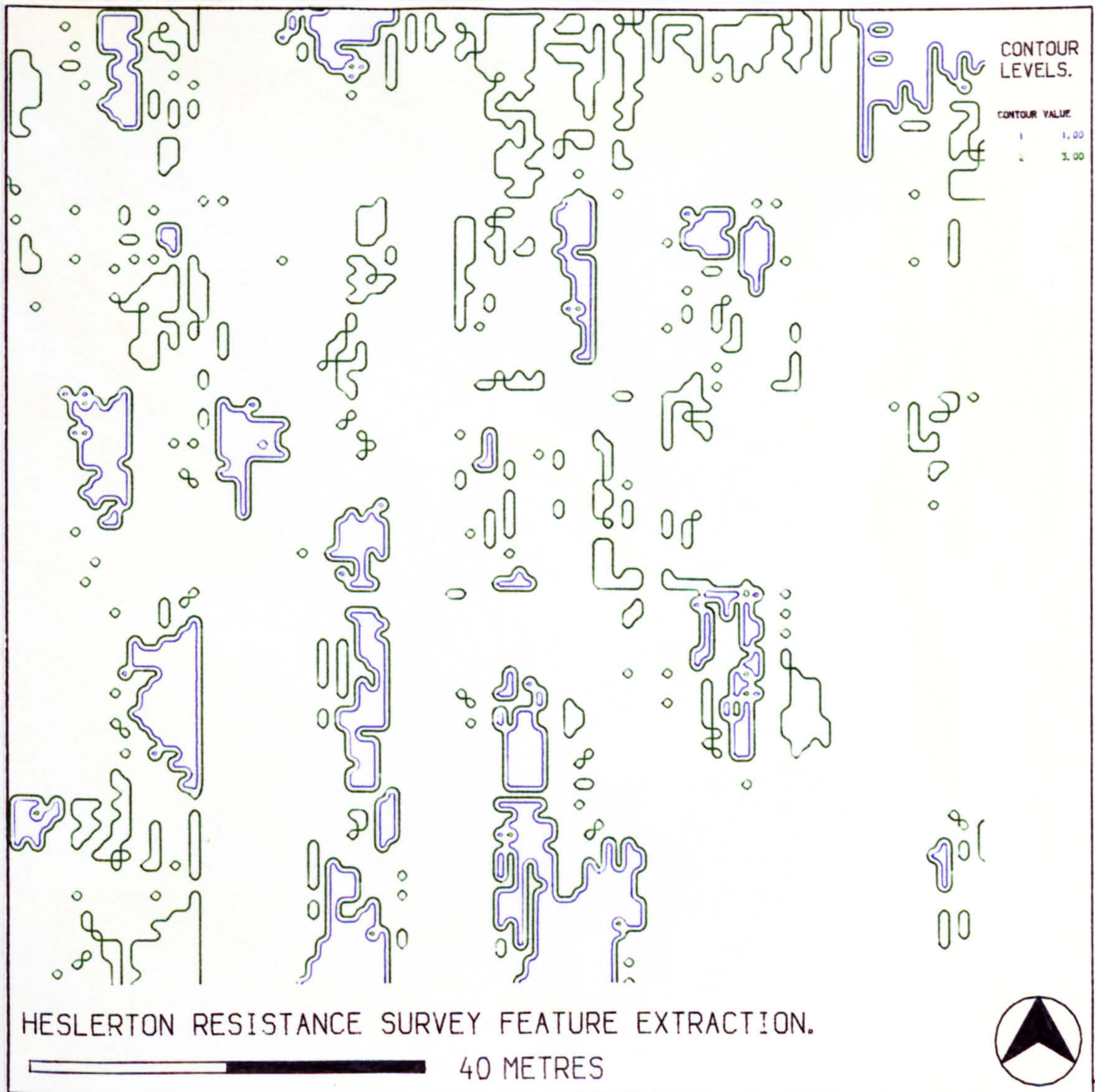


Figure 8.6 Resistance survey feature extraction.
 Low resistance features < 1, high resistance features > 3.

profitable as it would allow existing mathematical models of survey response to defined feature forms (Lynam, 1970; Heathcote, 1983; Houlder, 1983) to be effectively applied to the features identified by feature extraction.

8.4 MEASURING THE ARCHAEOLOGICAL SIGNIFICANCE OF OPERANDS

The preceding sections have presented some aspects of archaeological association in terms of operands and operators. This approach has implicitly tried to define some form of objective framework within which to operate. The idea of formalized frameworks within archaeology is important, as they offer the opportunity of handling in a rational manner not just practical problems in archaeology, but more abstract ideas that are usually the provenance of so called 'archaeological theory'. However one of the major problems with any theoretical approach to archaeology, is the potential loss of archaeological integrity. A formalized framework may force problems to fit the framework and lose sight of true archaeological goals, although the definition of these goals is itself controversial (Hawkes, 1968).

These problems of determining archaeological significance extend even to the basic archaeological data types, such as pottery sherds recovered through excavation or surface collection. Orton (1975) has argued that conventional measures of sherd count, sherd weight and vessel equivalents are inadequate archaeological quantifications of pottery. Instead a realistic and workable stochastic model for the break-up of a vessel into sherds is needed to allow an objective measure of vessels represented. Orton (1982a) has proposed such a

breakage model, termed a Kirby process, which he has attempted to apply using recursive sampling. A similar problem of quantification exists for the Heslerton surface distribution of pottery sherds. Here the simple measures of sherd count and sherd weight used in section 6.4 only indicate the density of surface pottery. More detailed information, potentially of greater archaeological significance, is carried by the actual frequency distribution of sherd attributes.

For example given a sample of recently broken vessels we could represent the frequency distribution of their sherd weights by a Normal distribution. Over time these sherds would continue to be broken and intuitively we would expect them to exhibit a change in their weight distribution pattern as shown in figure 8.7.a. To model this behaviour quantitatively we need to find a suitable statistical distribution. A good candidate, from purely empirical observations, is the Gamma distribution, which is defined as:

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^{\alpha} \Gamma(\alpha)} I_{[\alpha, \infty)}(x),$$

where $\alpha > 0$, $\beta > 0$.

The Gamma distribution has a positive skew which becomes less pronounced when its shape parameter α is large. In fact for sufficiently large α the Gamma distribution approximates a Normal distribution. When α is set to one, the Gamma distribution specialises to the exponential distribution. The behaviour of the Gamma distribution for different values of α is shown in figure 8.7.b. It can be seen that it closely resembles the behaviour of our intuitive model. The mean and variance of the Gamma

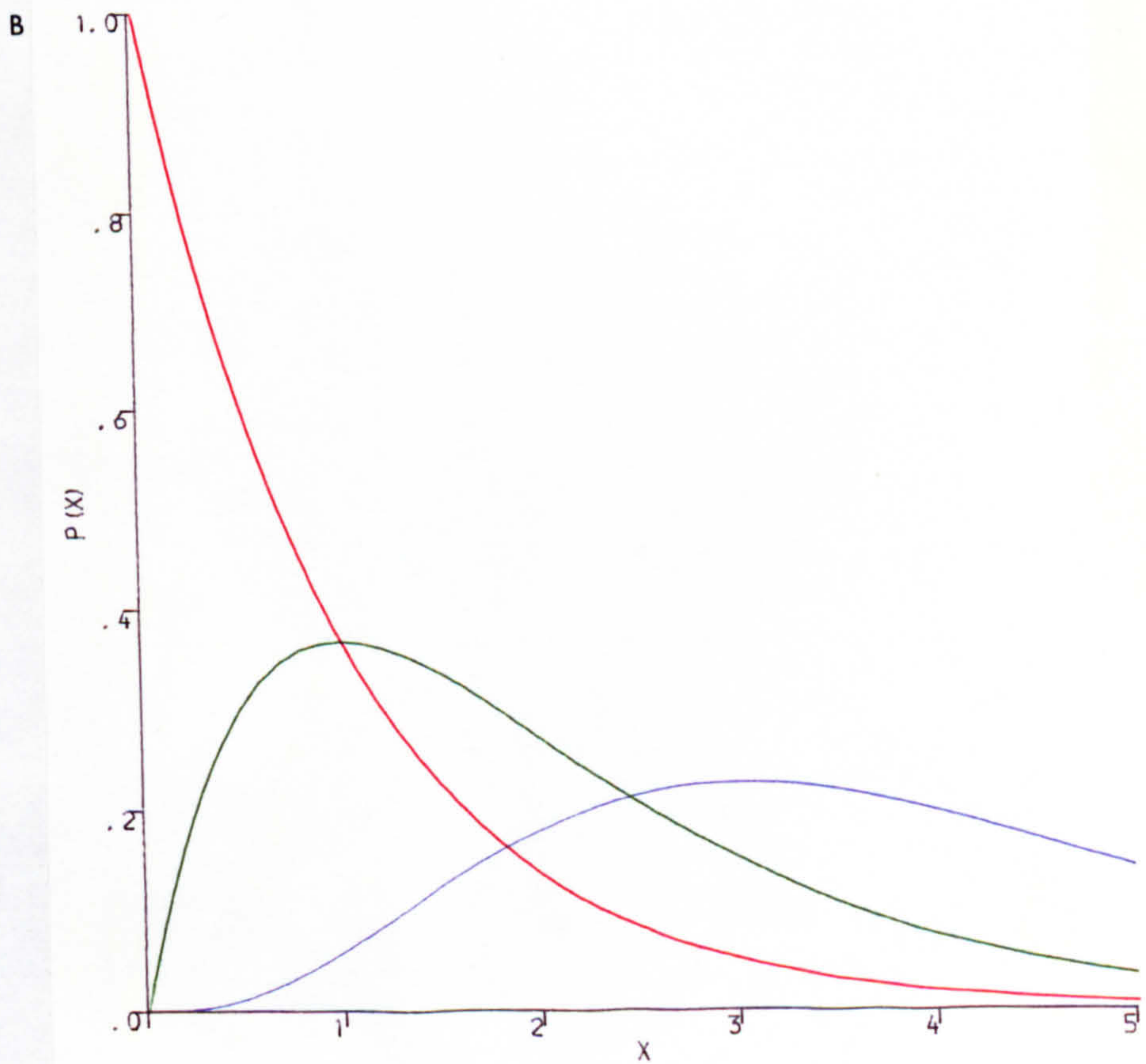
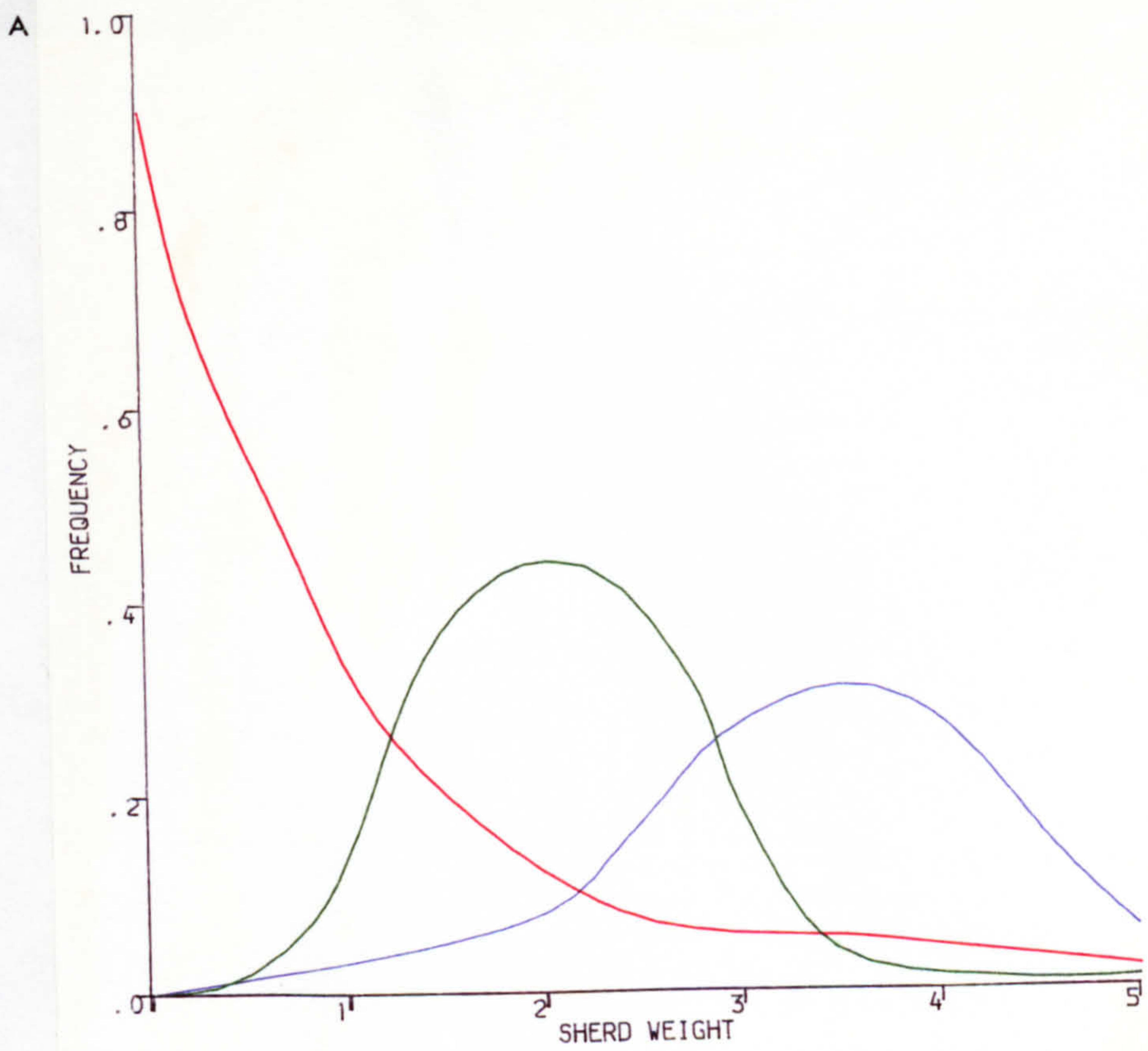


Figure 8.7 Break-up of pottery.
 A) Intuitive model. B) Gamma distribution.

distribution are defined by its two parameters:

$$E[x] = \alpha\beta,$$

$$\text{Var}[x] = \alpha\beta^2.$$

Maximum likelihood estimates of α and β (Johnson and Kotz, 1970)

are:

$$\ln \left[\frac{\text{arithmetic mean}}{\text{geometric mean}} \right] = \ln \hat{\alpha} - \psi(\hat{\alpha}),$$

$$\hat{\beta} = \frac{\bar{x}}{\hat{\alpha}},$$

where $\psi(\cdot)$ is the psi function.

More convenient approximate estimates (Greenwood and Durand, 1960)

are:

For $0 < Y \leq 0.5772$

$$\hat{\alpha} \doteq Y^{-1}(0.5000876 + 0.1648852Y - 0.0544274Y^2),$$

For $0.5772 < Y \leq 17$

$$\hat{\alpha} \doteq Y^{-1}(17.79728 + 11.968477Y + Y^2)^{-1}(8.898919 + 9.059950Y + 0.9775373Y^2)$$

with $\hat{\beta}$ calculated as before,

where $Y = \ln \left[\frac{\text{arithmetic mean}}{\text{geometric mean}} \right]$.

The error on these approximations of α is estimated by Bowman and Shenton (1968) to be less than 0.0088% for the first approximation and less than 0.0054% for the second approximation.

To apply the Gamma distribution model to the Heslerton pottery data it was first necessary to group contiguous samples into larger

blocks to give workable sample sizes (figure 8.8). The summary statistics for these new blocks are given in table 8.1 and histograms of the frequency distributions of their weights are shown in figures 8.9 to 8.13. The maximum likelihood estimates for α and β were then calculated using the approximations and are shown in table 8.2. To avoid complicating the interpretation of these results no errors of estimate were calculated. However all the estimates based on samples of less than eight have been precluded from the following considerations, as these seem to be the most erroneous.

The maximum likelihood estimates of α are shown in figure 8.14. They display an interesting valley like structure which contrasts with the sherd weight contour plot of figure 6.5.b. The α parameter can be taken as a simple indication of degree of brokenness. However the exact archaeological significance of brokenness is difficult to define. When a pot is broken and discarded it does not immediately enter the depositional environment. As Binford (1978) has shown, site rubbish may undergo substantial modifications before finally being sealed within a deposit. Thus pottery sherds within a high activity zone would probably become more broken than sherds in a low activity zone. This pattern does seem to fit the structure of figure 8.14, where the evidence indicates an occupation zone in the north of the study area. Although these results are encouraging the sampling density of the pottery blocks is very low because of the need to achieve adequate sample sizes. For this reason the use of the results in any further association analysis is not really feasible. Instead the pottery sherd weights will be used as a default measure of pottery quantification.

POT1	POT2	POT3	POT4	POT5
POT6	POT7	POT8	POT9	POT10
POT11	POT12	POT13	POT14	POT15
POT16	POT17	POT18	POT19	POT20
POT21	POT22	POT23	POT24	POT25

Figure 8.8 Pottery blocks.

Block	number of sherds	total sherd weight	mean sherd weight	standard deviation of sherd weight
POT1	66	539.9	8.18	8.36
POT2	68	464.8	6.84	6.96
POT3	20	141.7	7.09	6.90
POT4	9	140.3	15.59	16.60
POT5	4	12.2	3.05	2.51
POT6	52	405.9	7.81	8.36
POT7	38	321.8	8.47	12.00
POT8	16	97.5	6.09	6.66
POT9	5	39.6	7.92	7.52
POT10	3	44.4	14.80	5.73
POT11	30	135.0	4.50	5.41
POT12	9	68.5	7.61	9.32
POT13	6	25.2	4.20	1.67
POT14	9	86.0	9.56	7.27
POT15	2	16.4	8.20	0.42
POT16	21	161.1	7.67	9.05
POT17	6	34.7	5.78	3.20
POT18	8	41.6	5.13	5.18
POT19	5	53.5	10.72	8.36
POT20	2	5.0	2.50	0.42
POT21	13	51.6	3.97	1.83
POT22	13	50.8	3.91	2.43
POT23	5	19.8	3.96	2.07
POT24	0	-	-	-
POT25	0	-	-	-

Table 8.1 Summary statistics of pottery blocks (weight in gms).

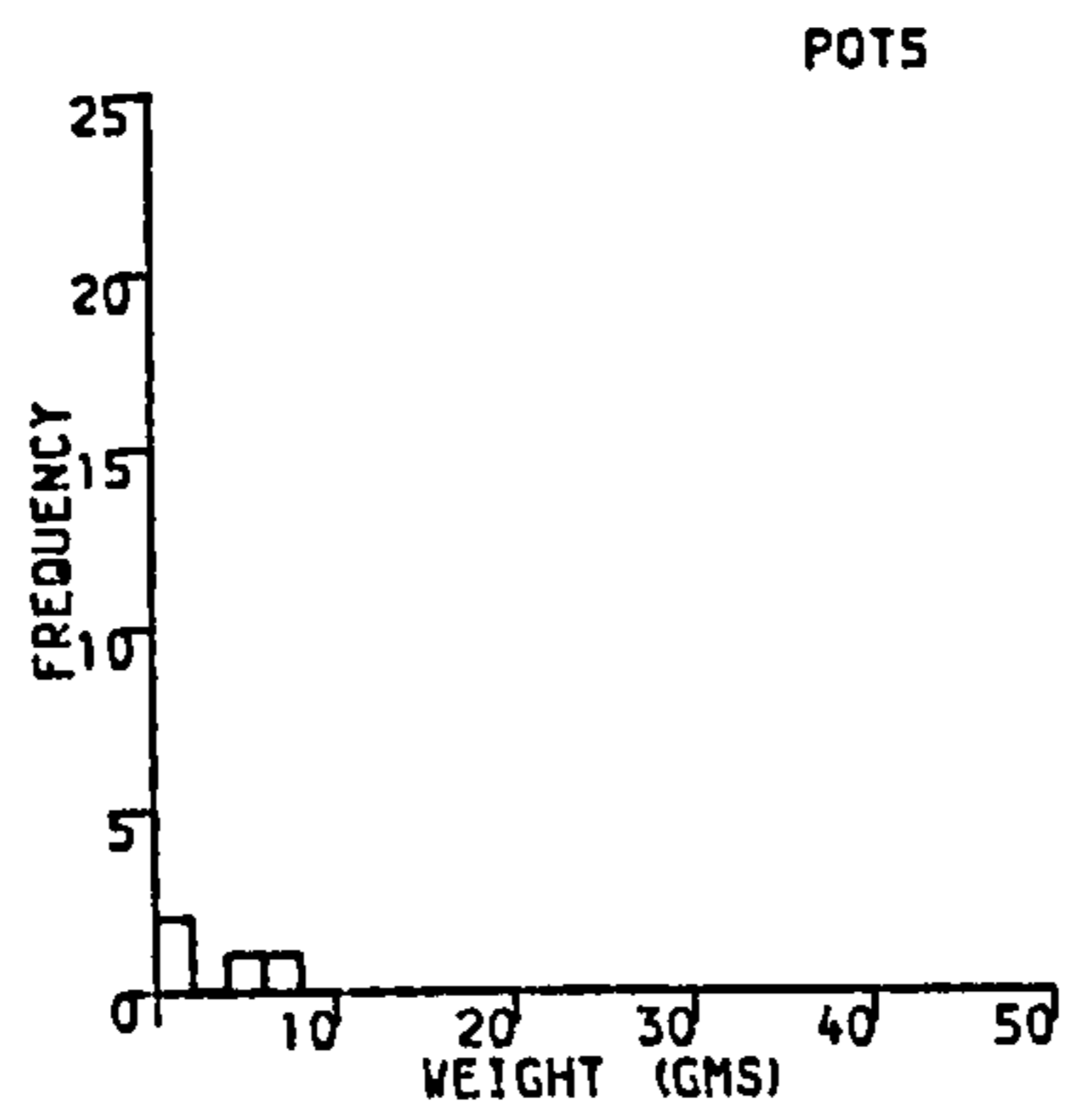
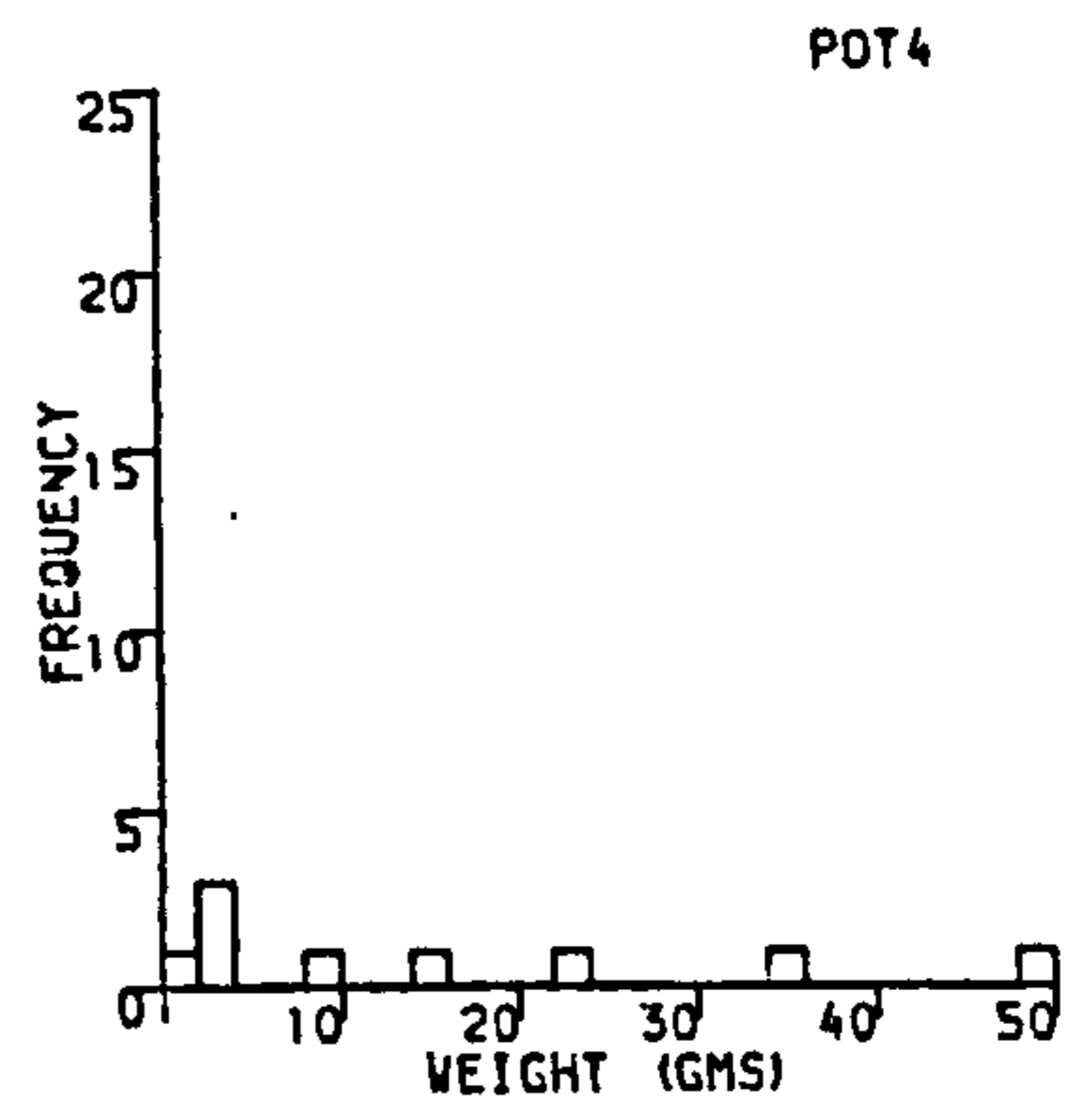
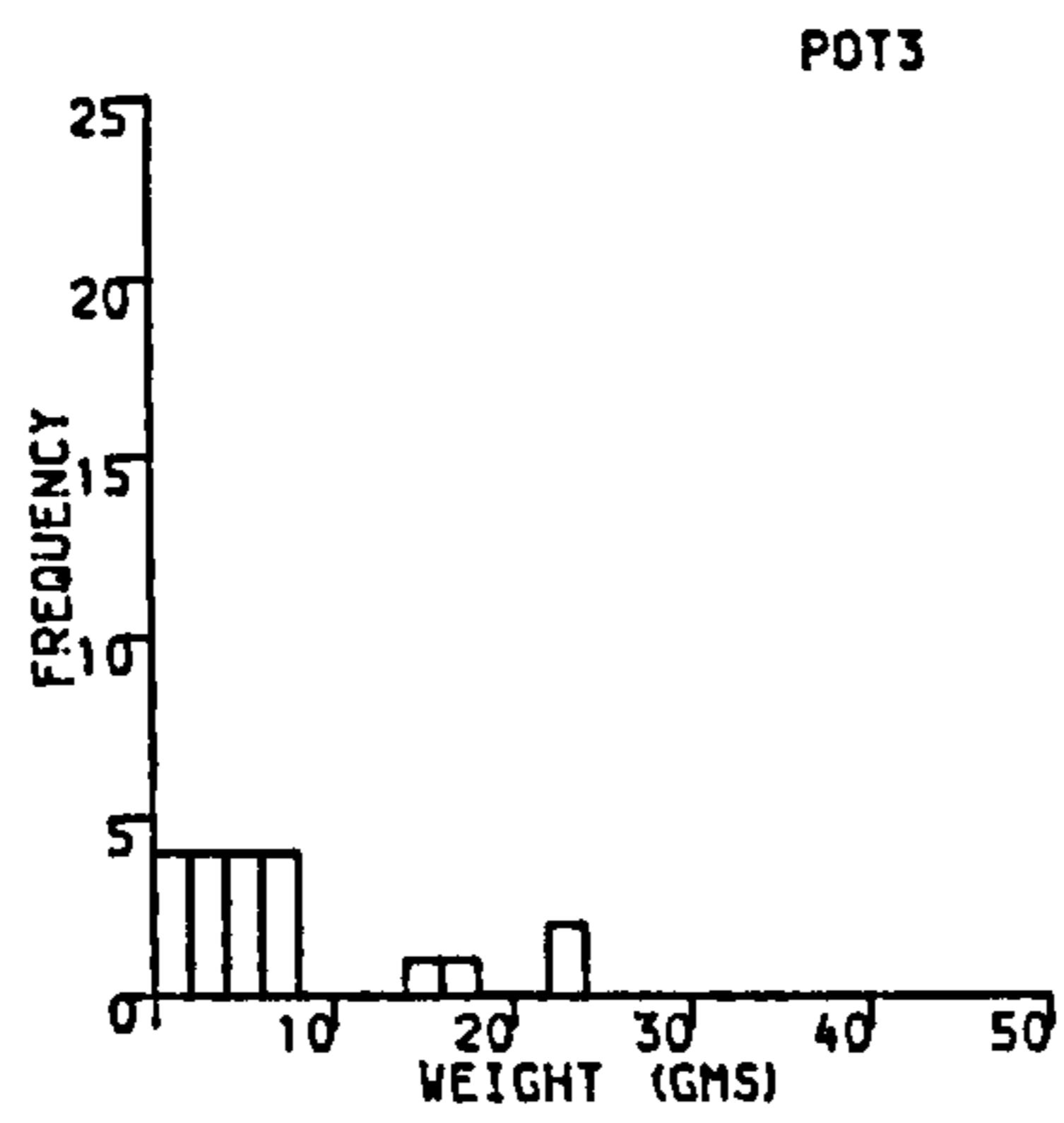
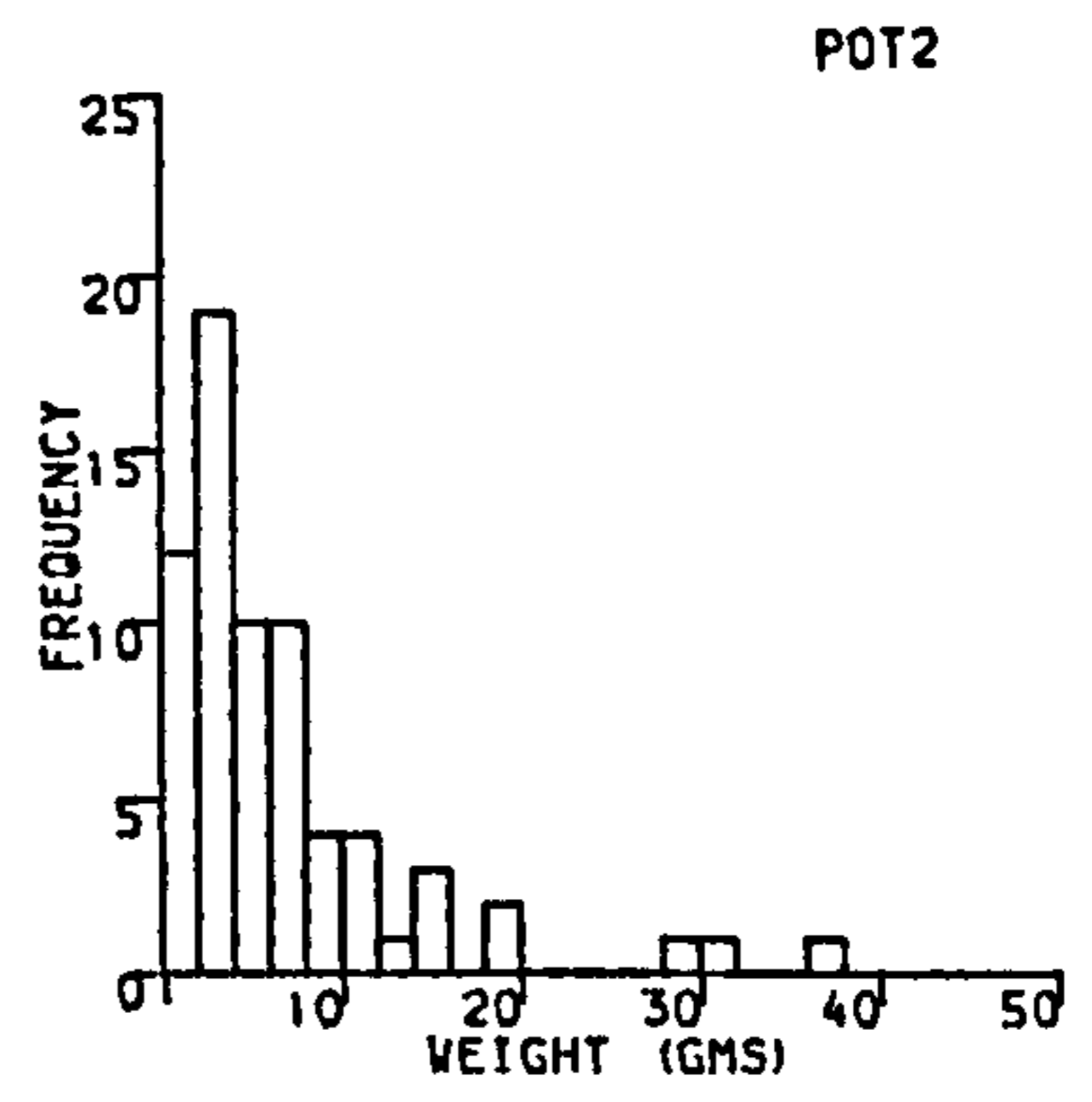
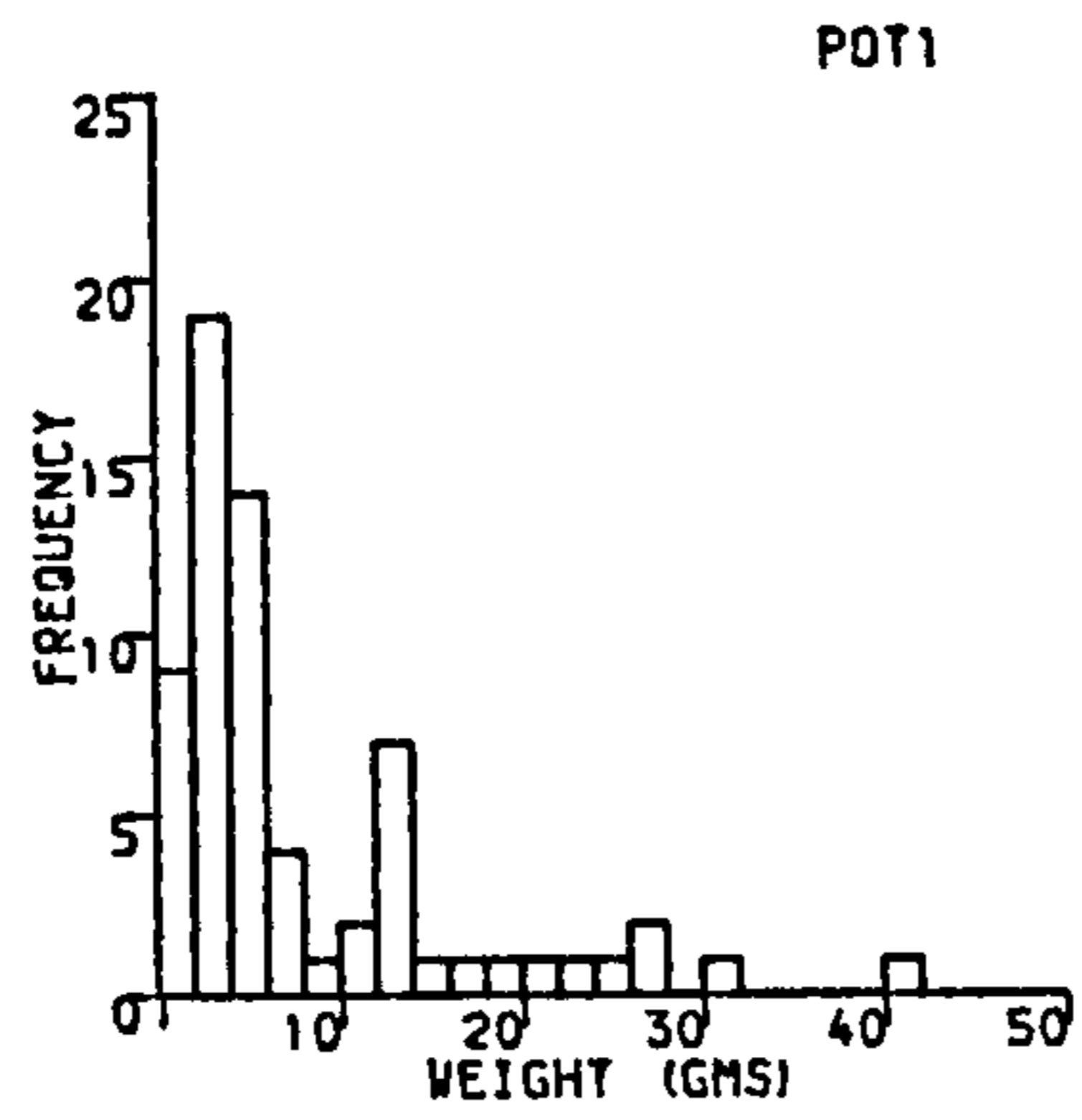


Figure 8.9 POT1 to POT5 weight histograms.

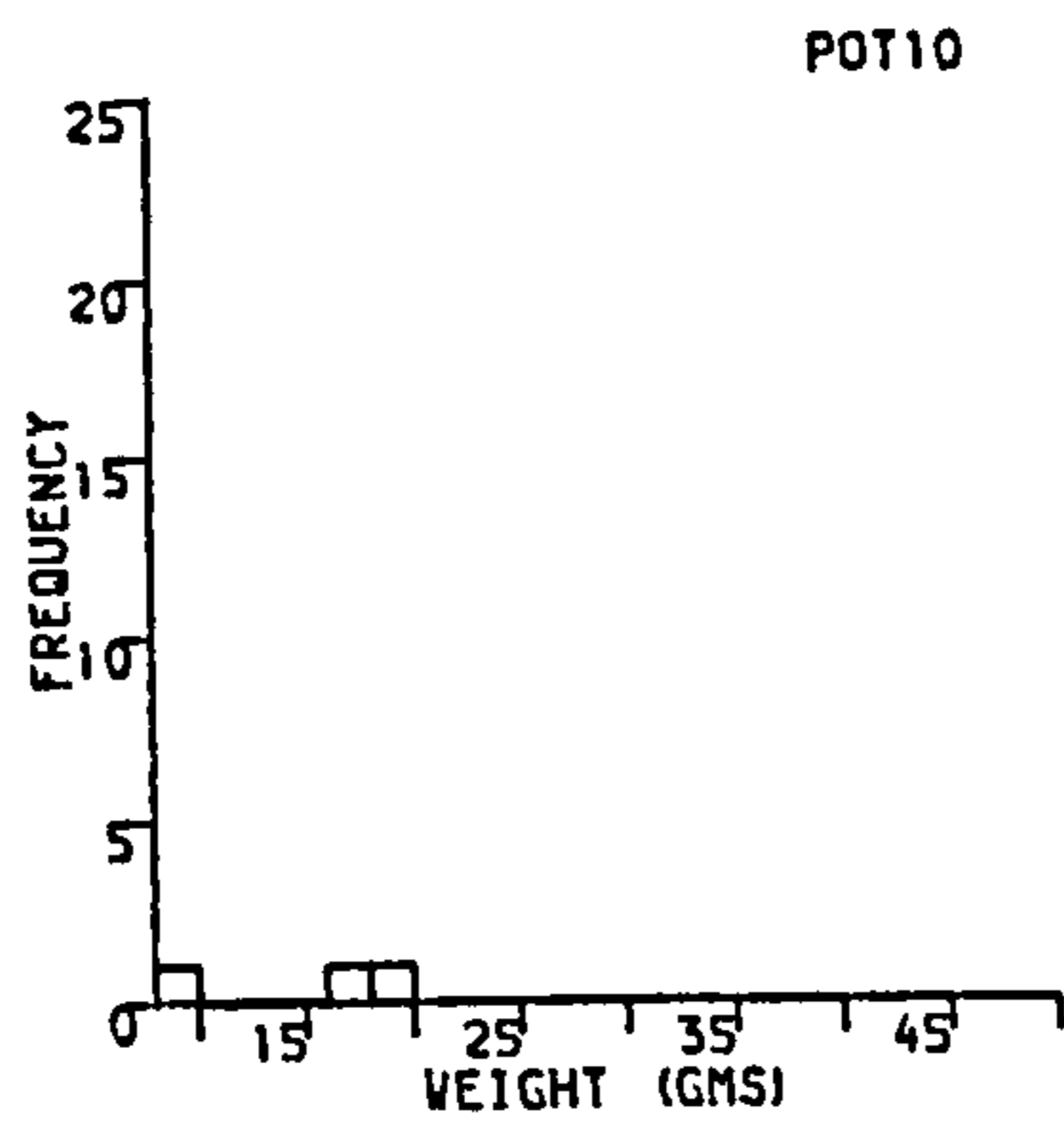
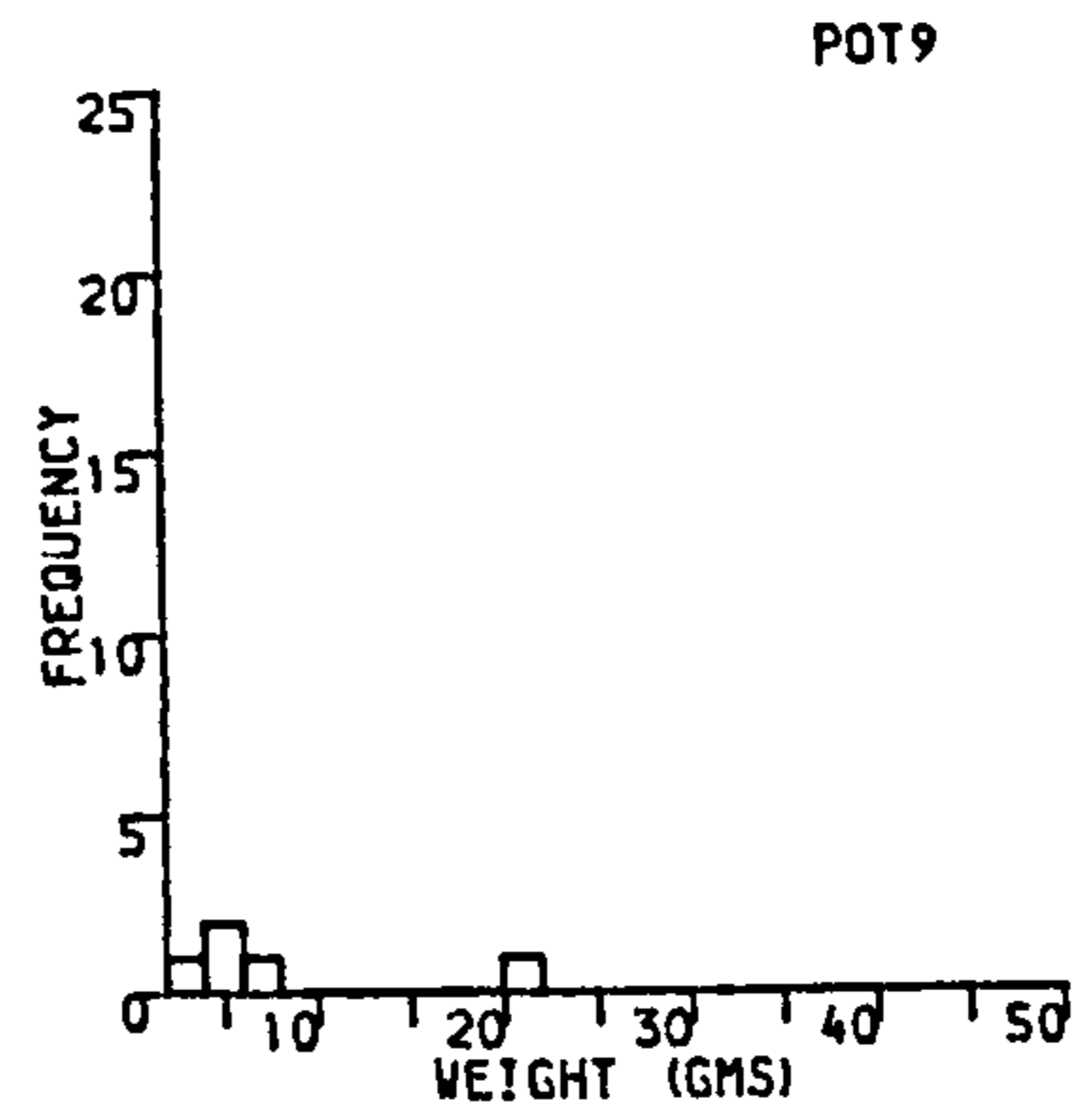
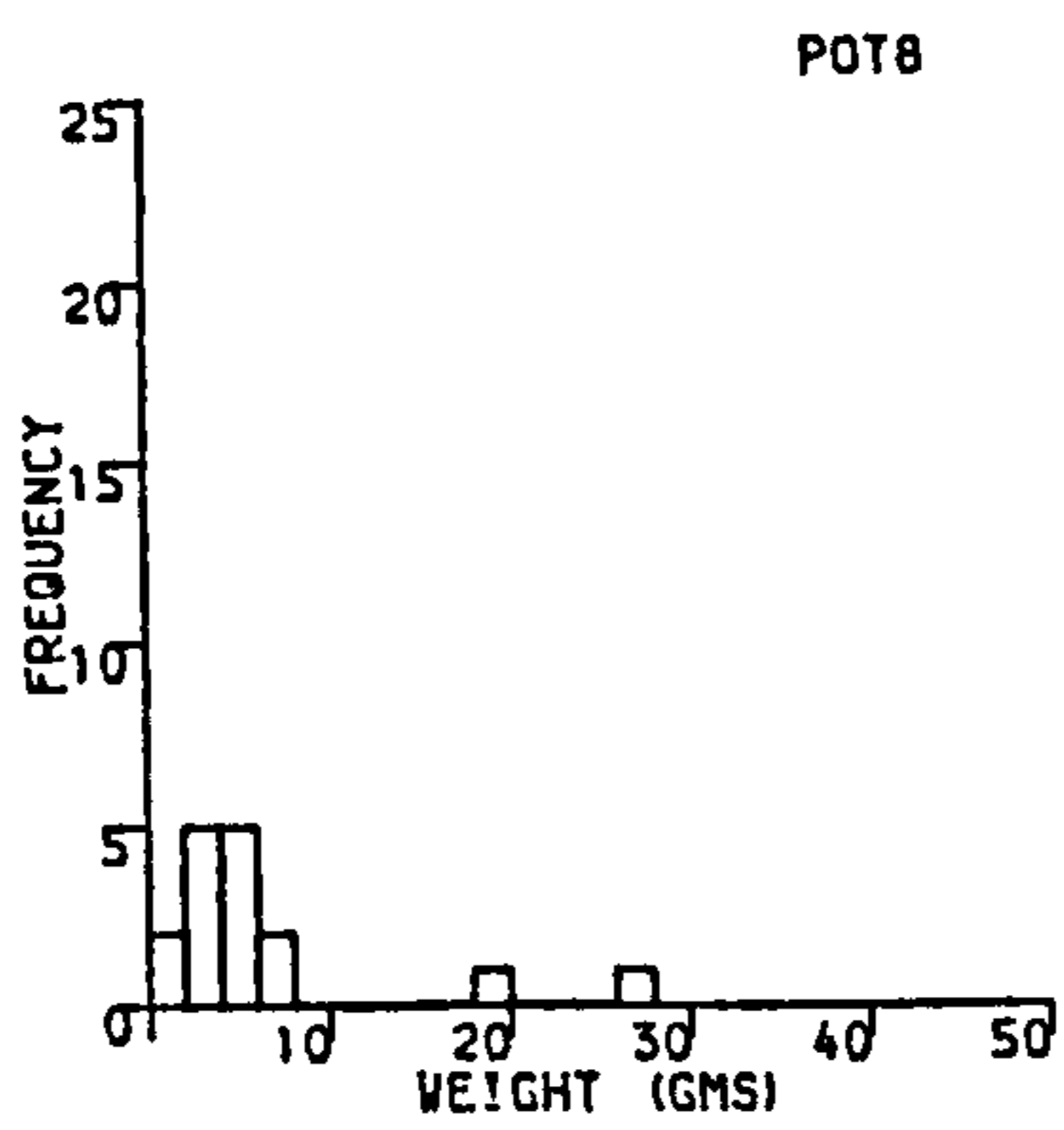
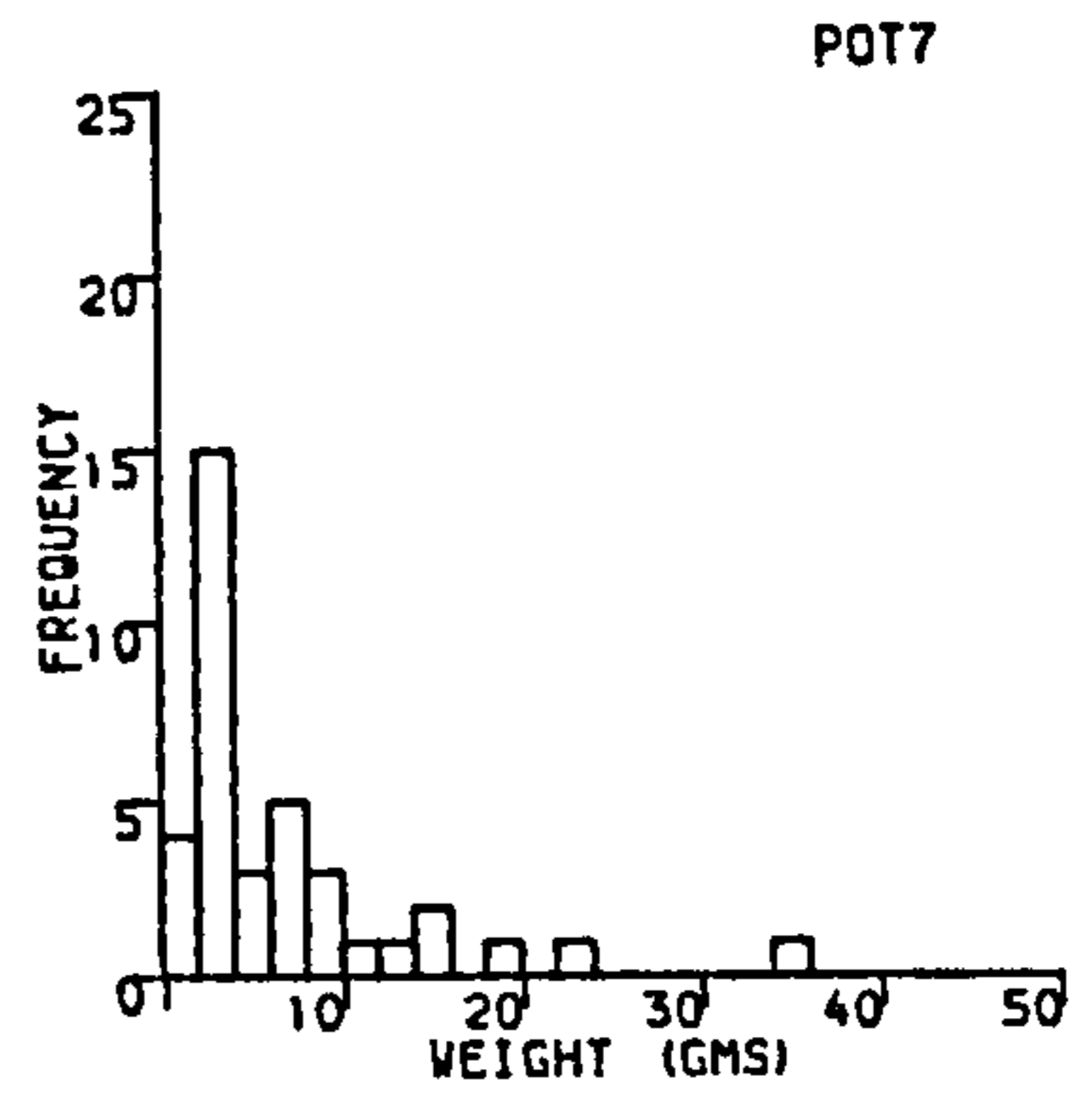
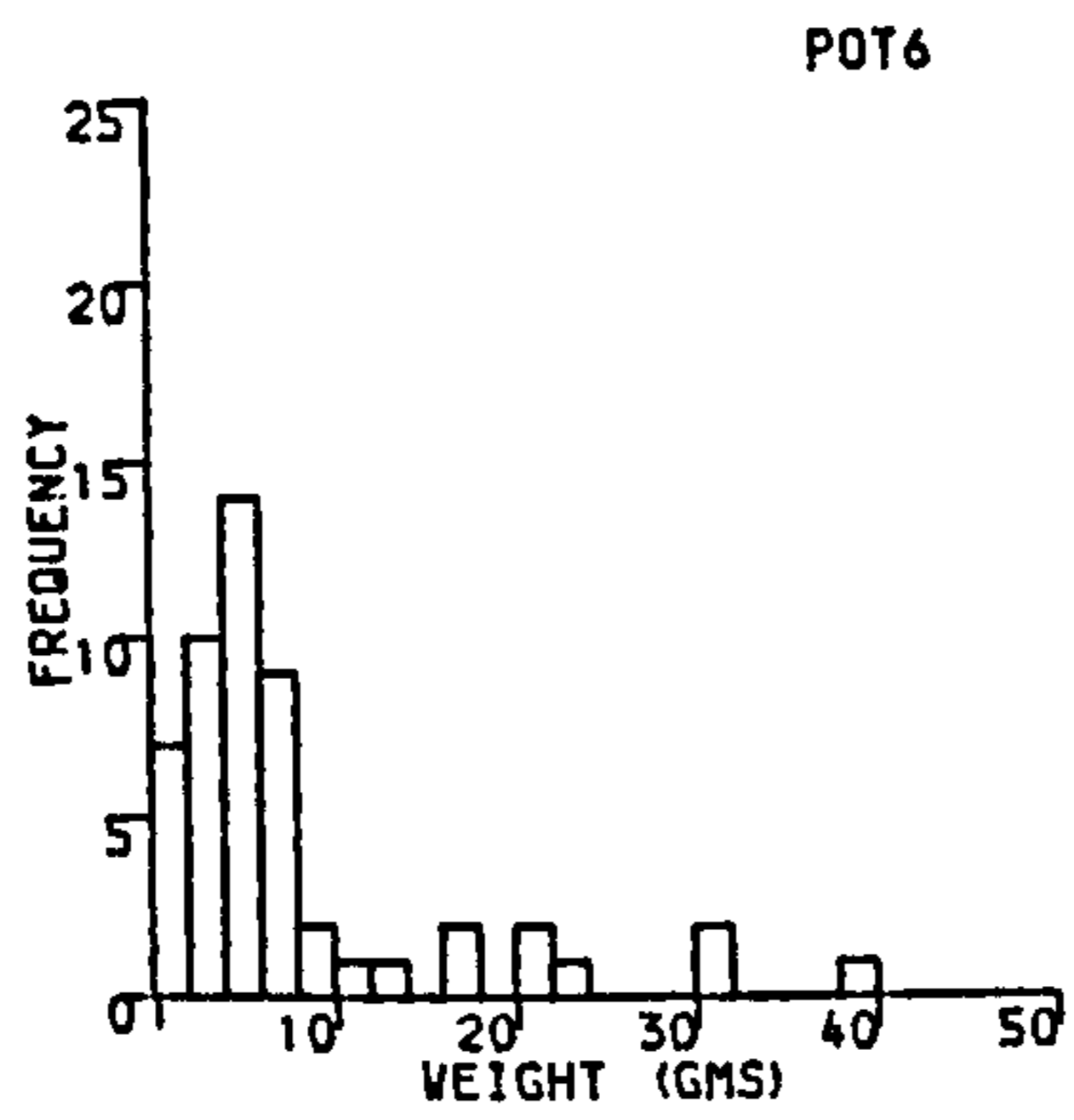


Figure 8.10 POT6 to POT10 weight histograms.

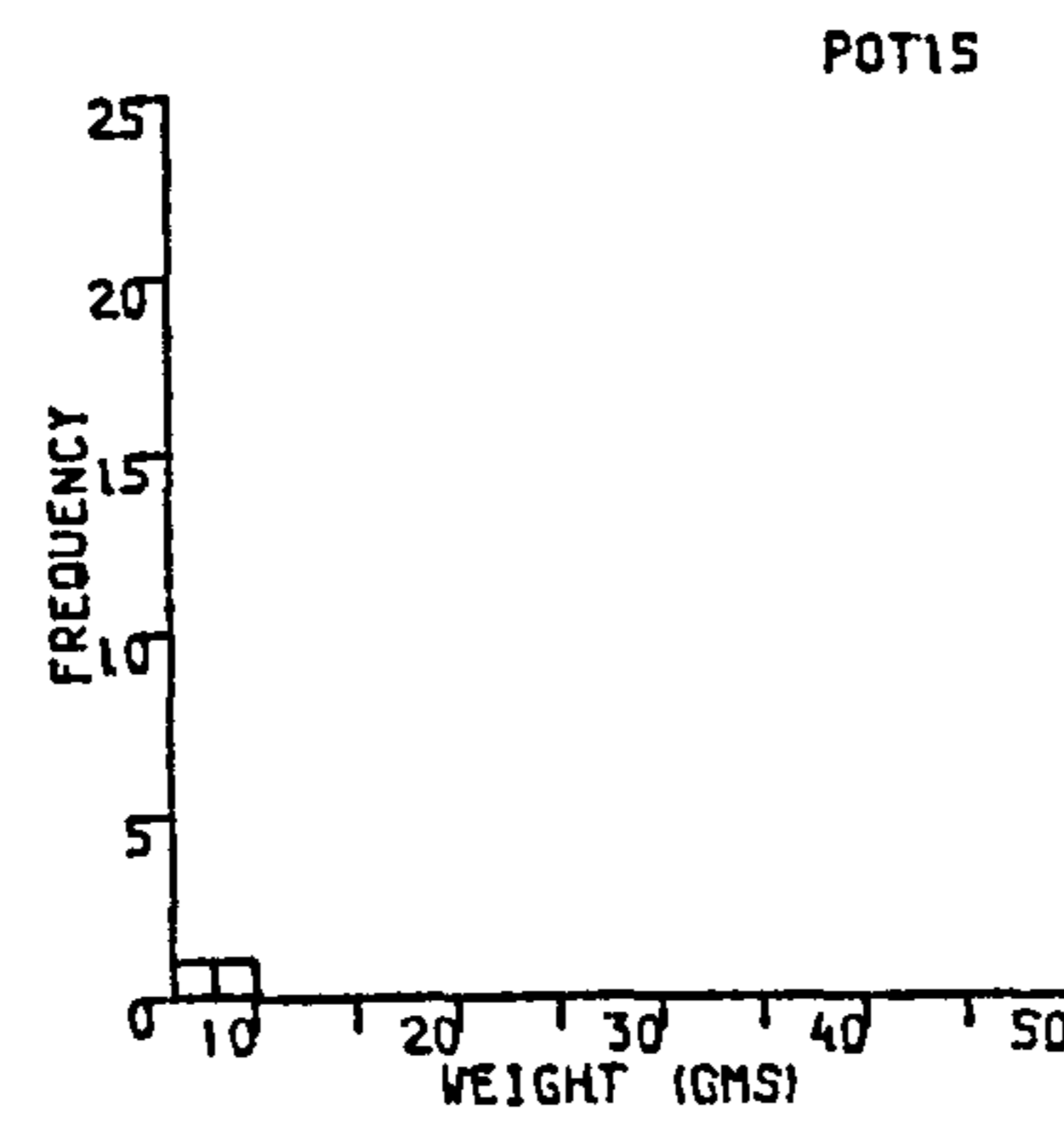
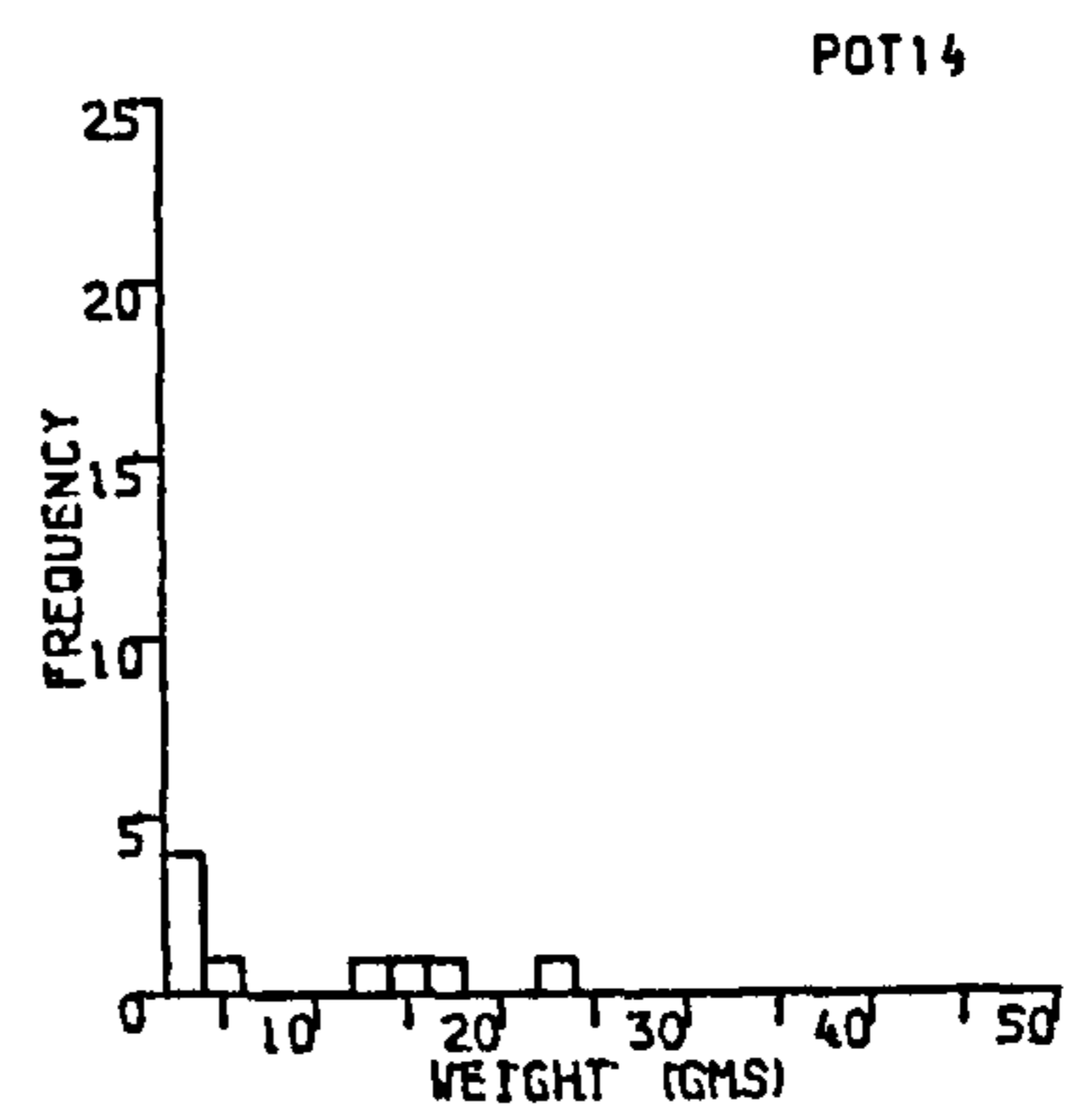
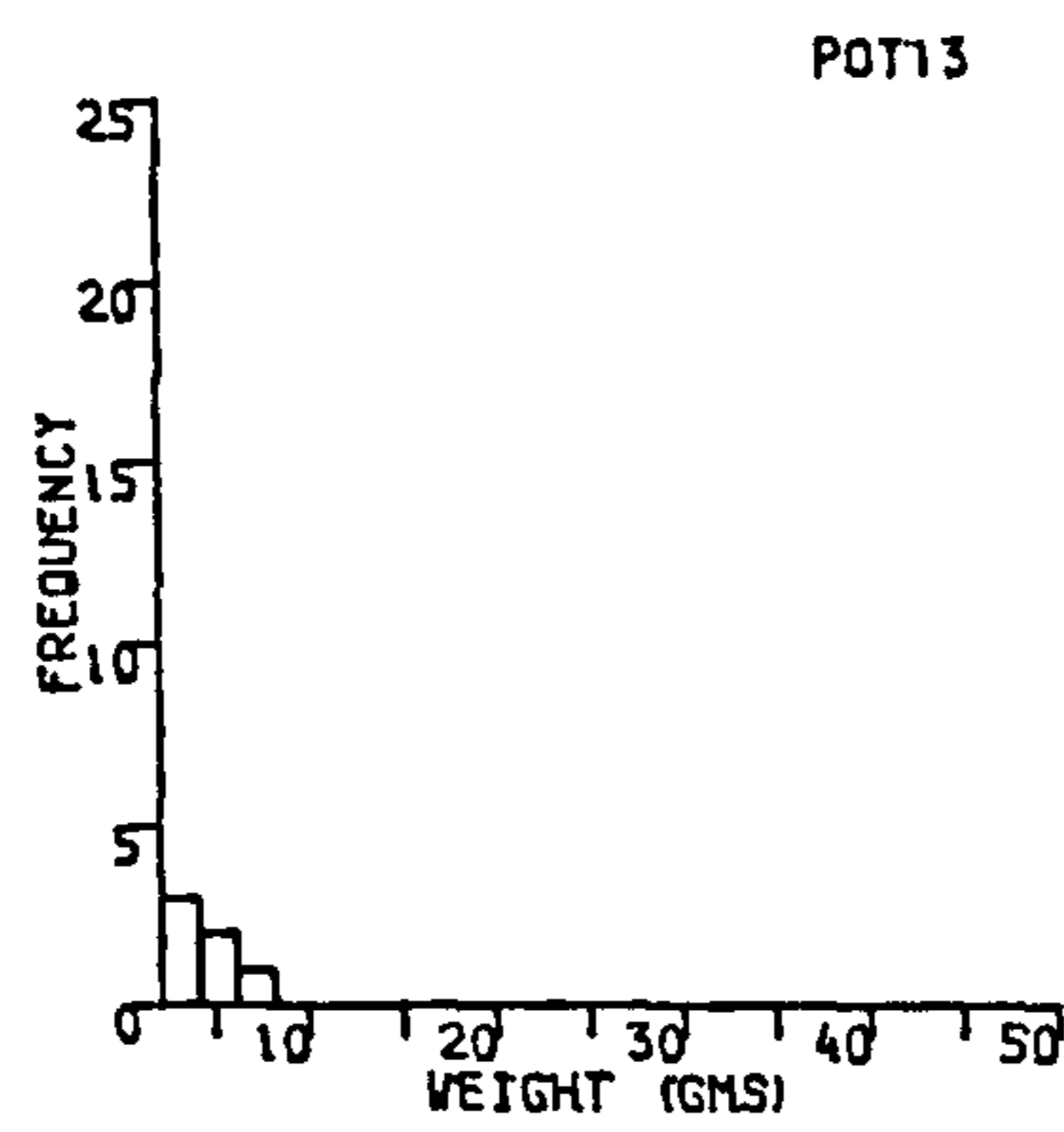
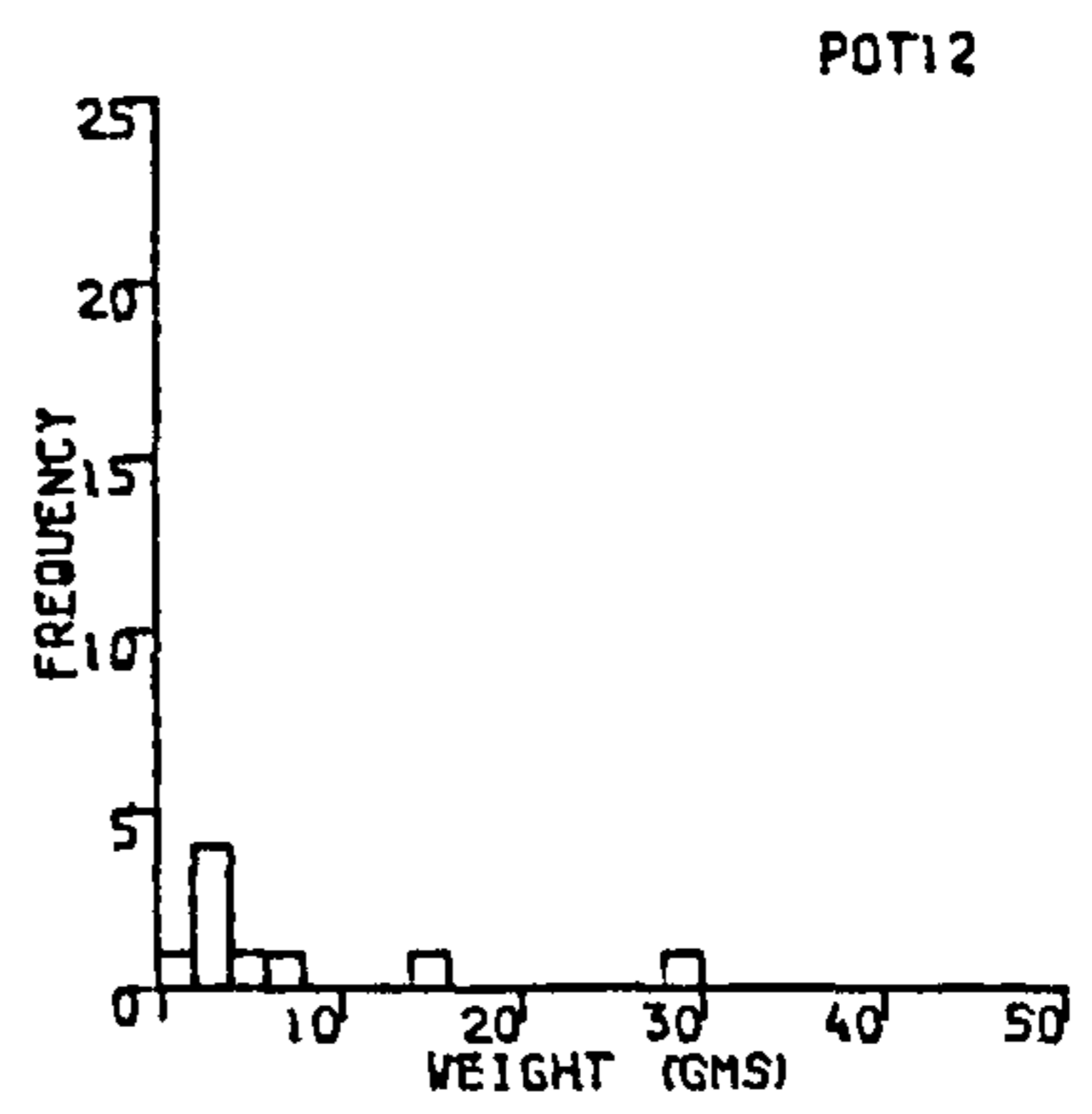
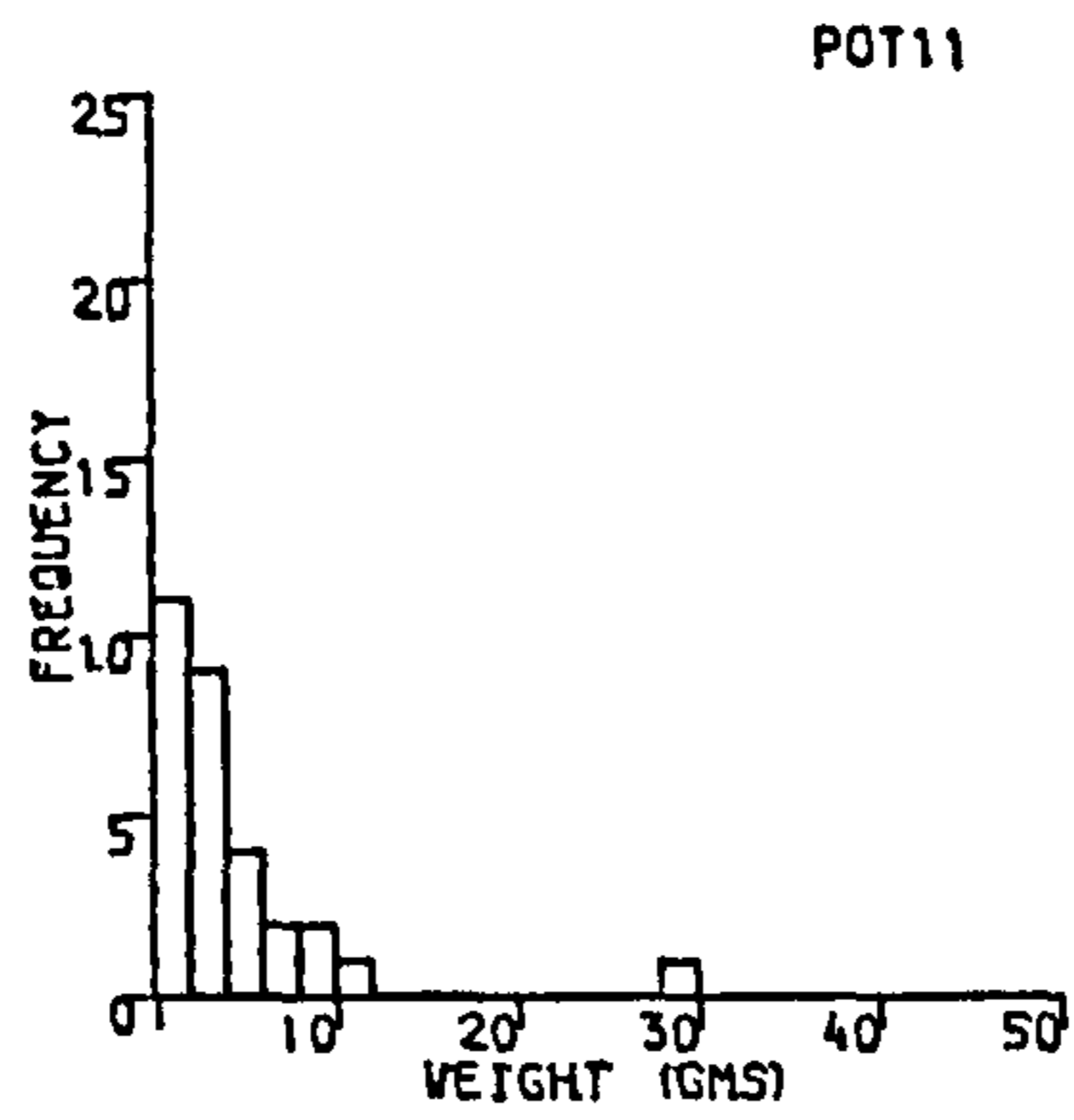


Figure 8.11 POT11 to POT15 weight histograms.

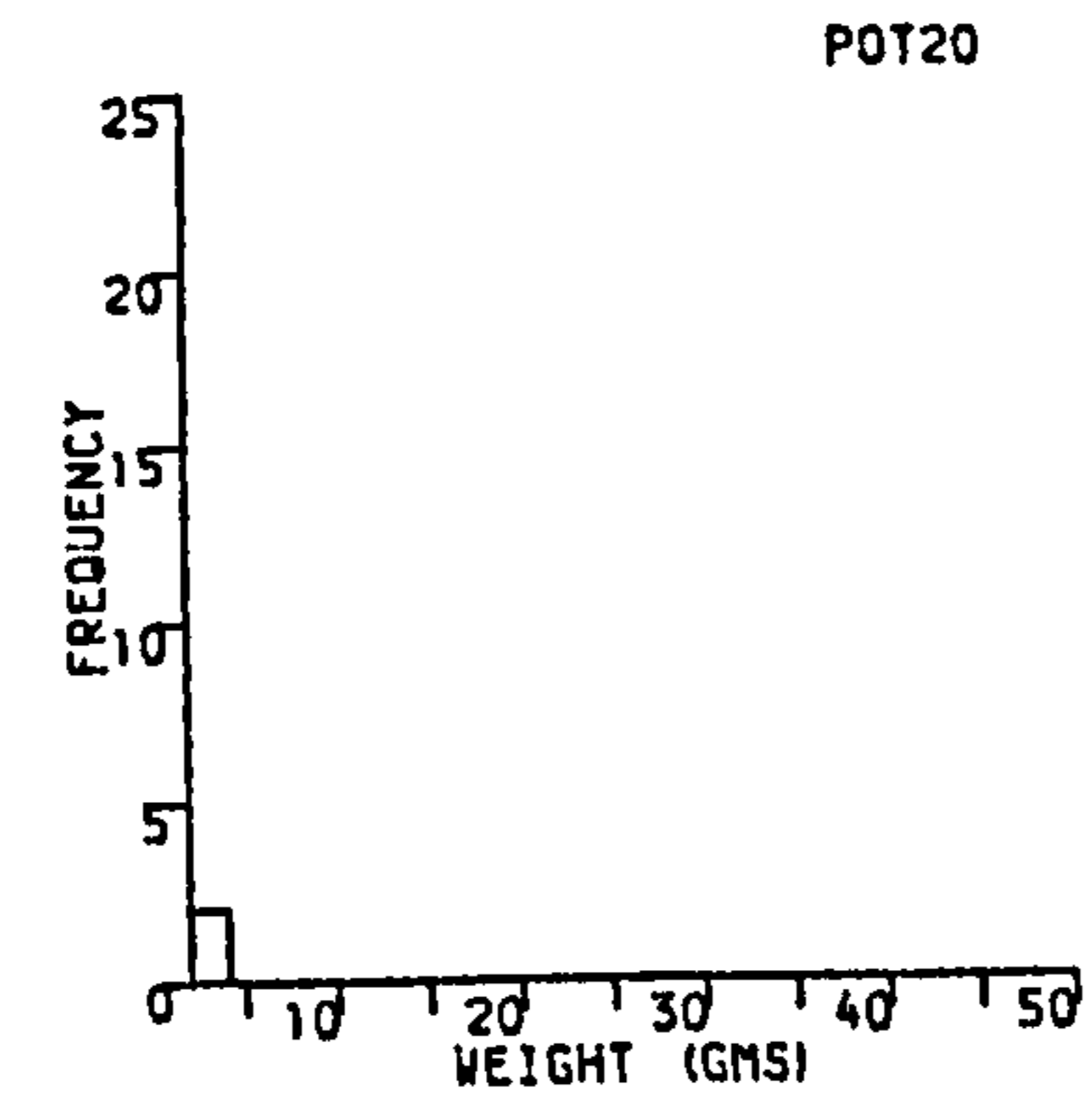
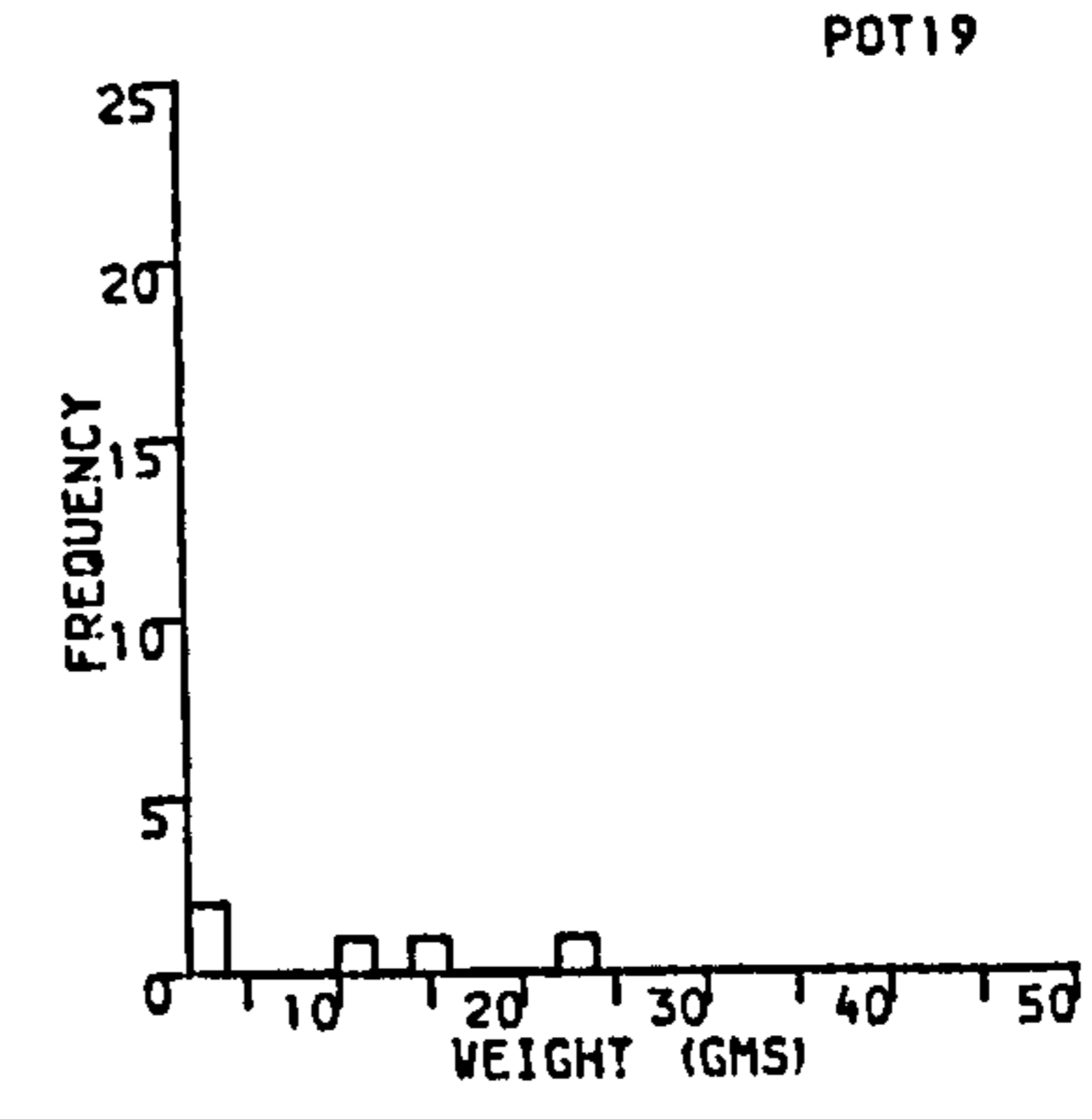
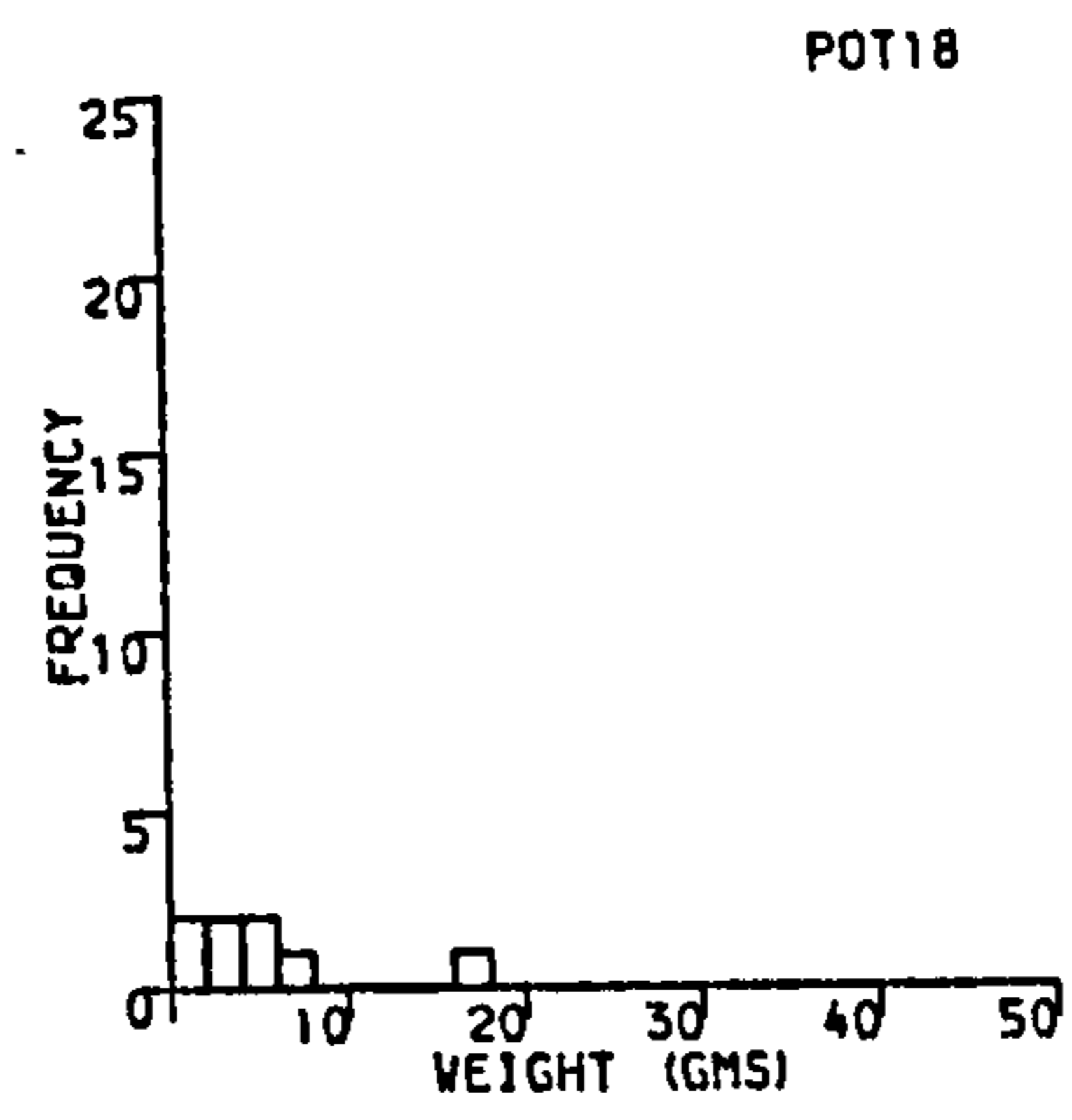
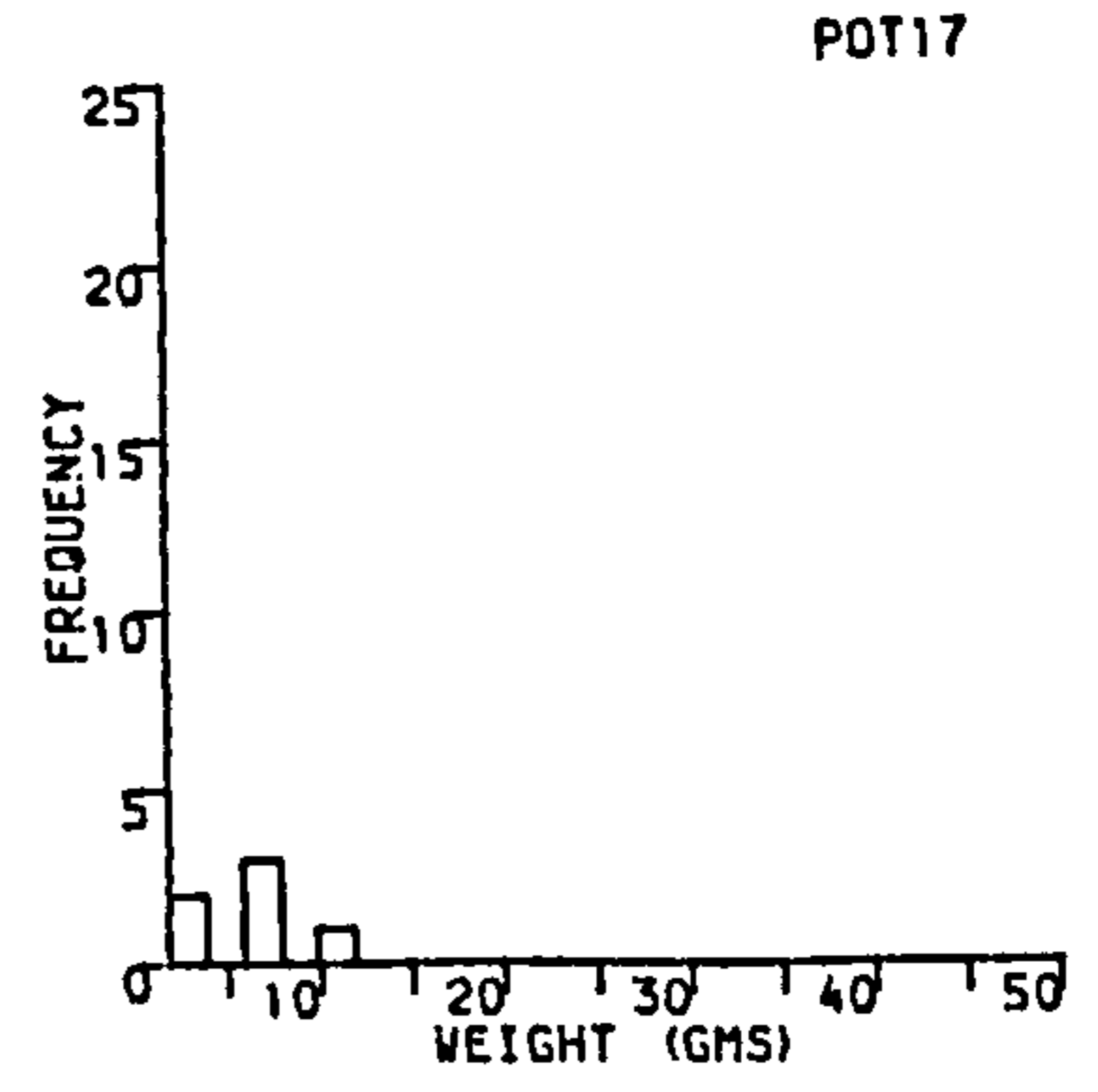
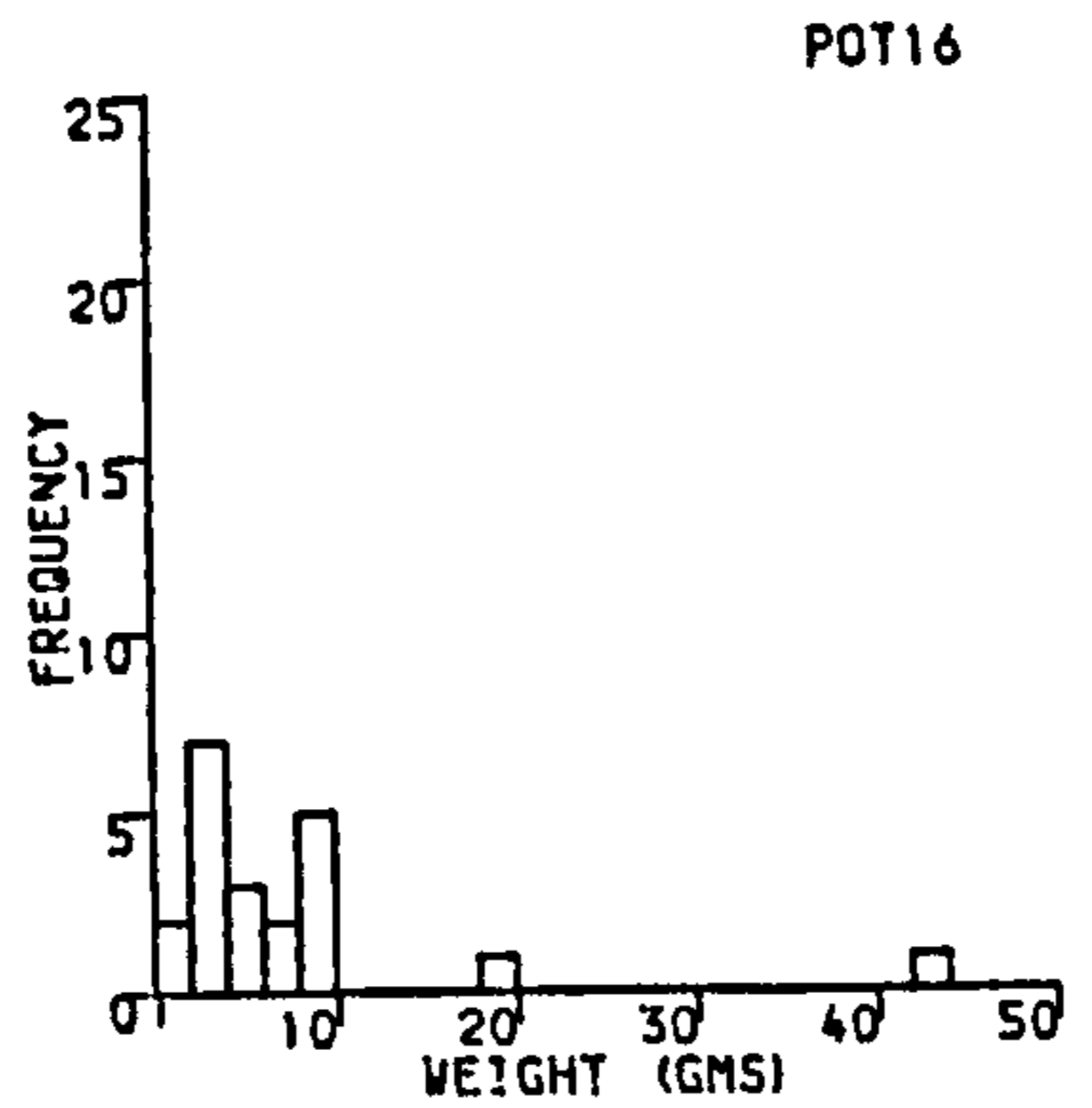


Figure 8.12 POT16 to POT20 weight histograms.

POT 1
 POT 2
 POT 3
 POT 4
 POT 5
 POT 6
 POT 7
 POT 8
 POT 9
 POT 10
 POT 11
 POT 12
 POT 13
 POT 14
 POT 15
 POT 16
 POT 17
 POT 18
 POT 19
 POT 20
 POT 21
 POT 22
 POT 23
 POT 24
 POT 25

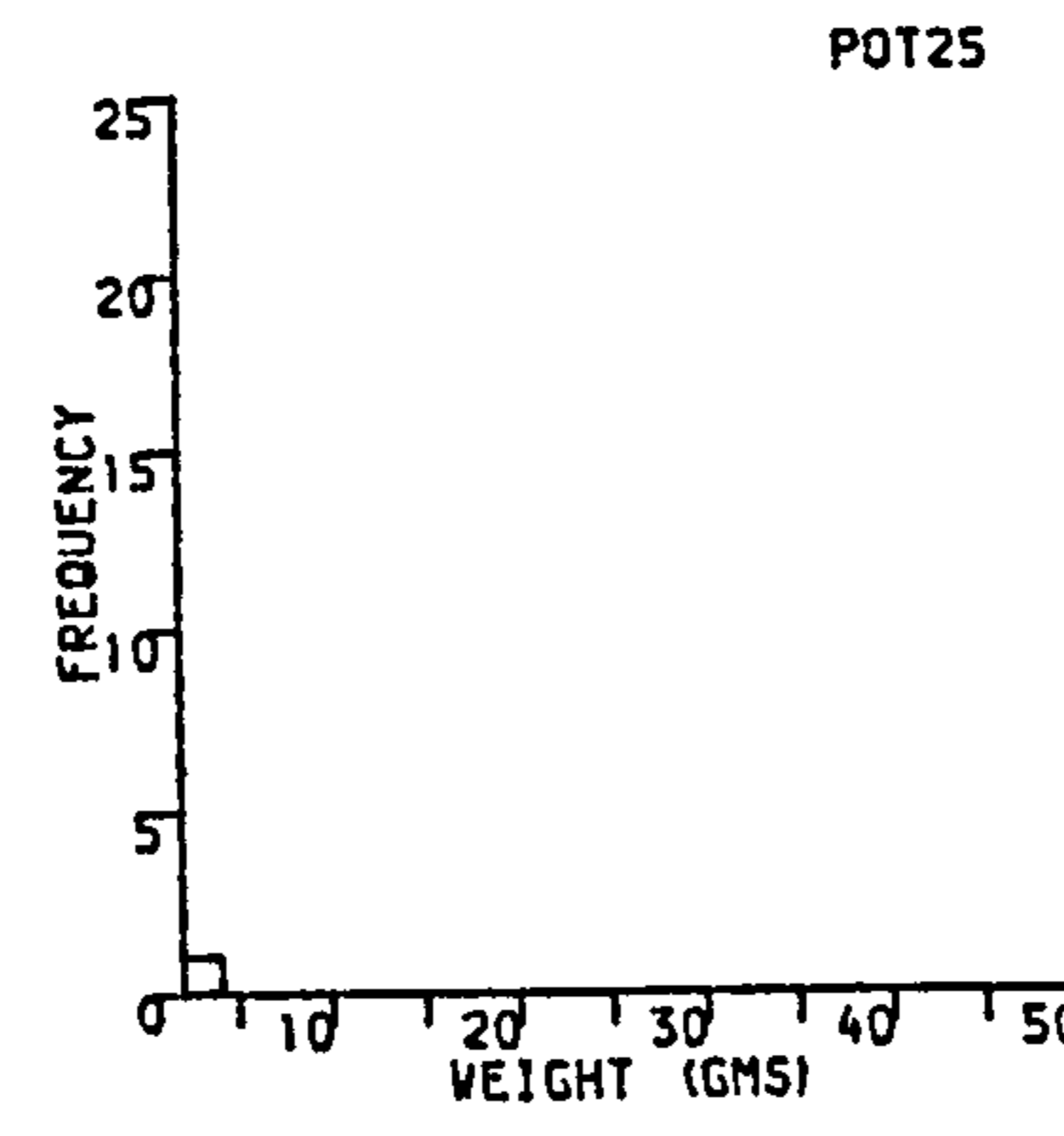
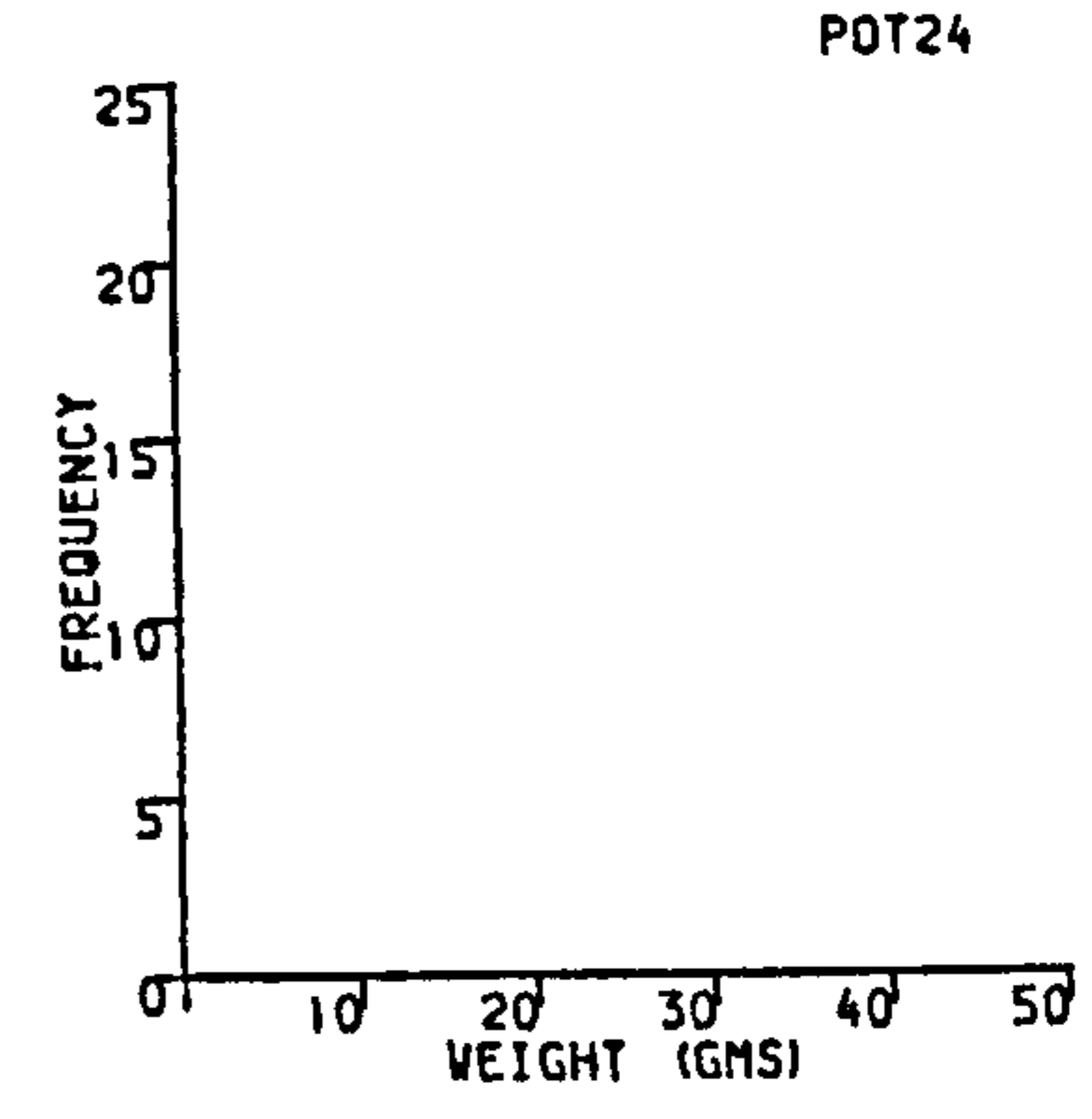
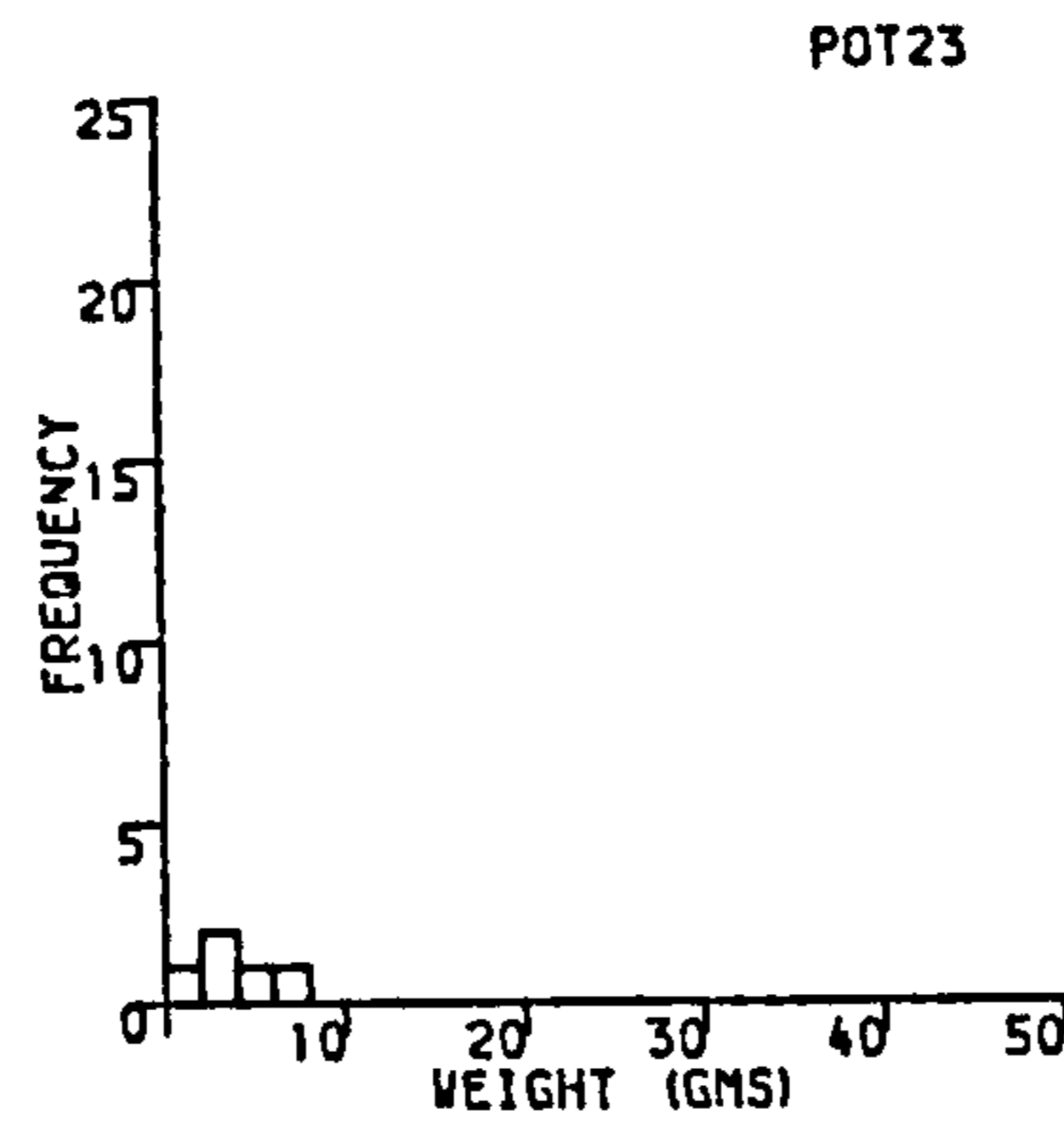
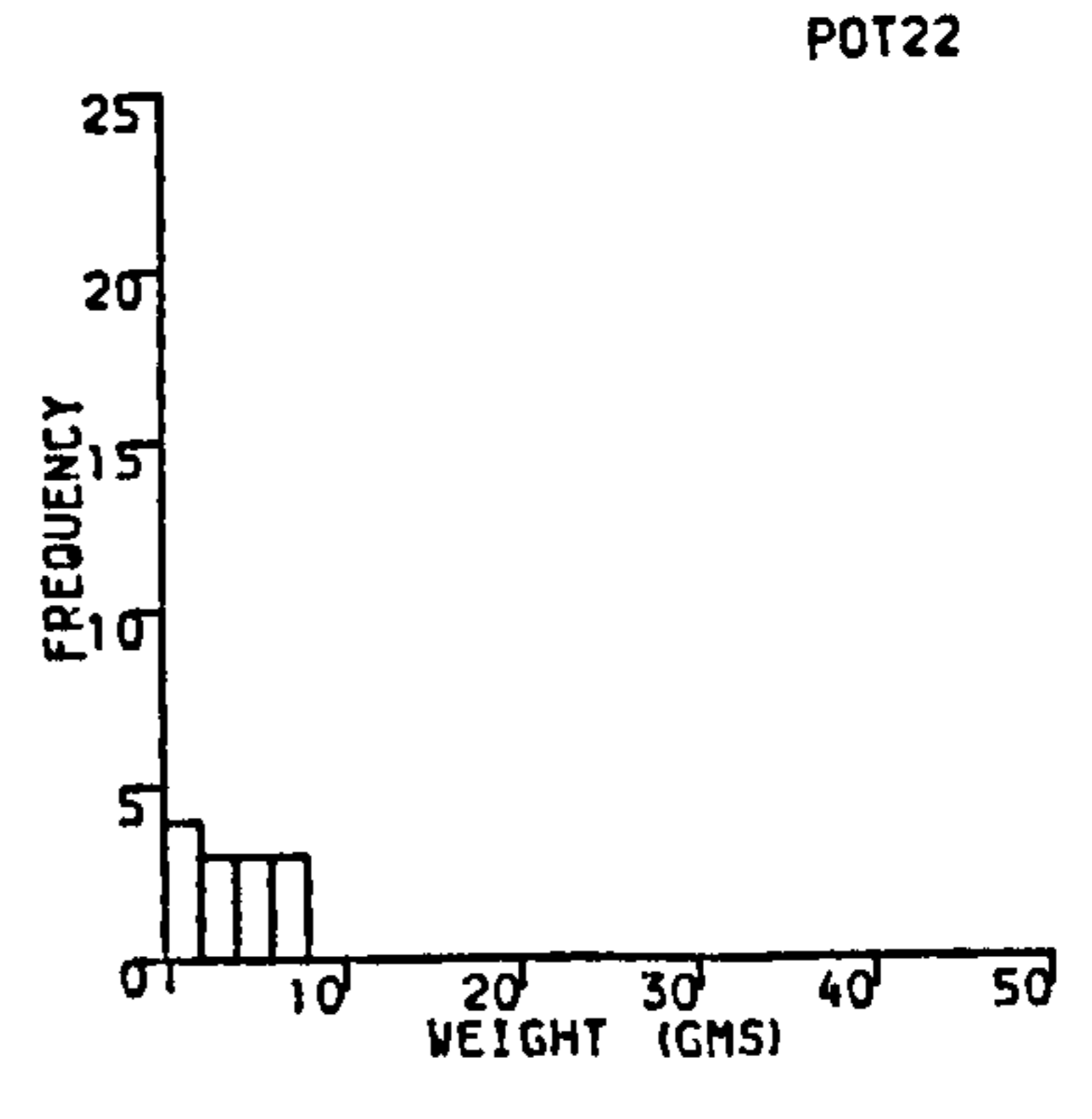
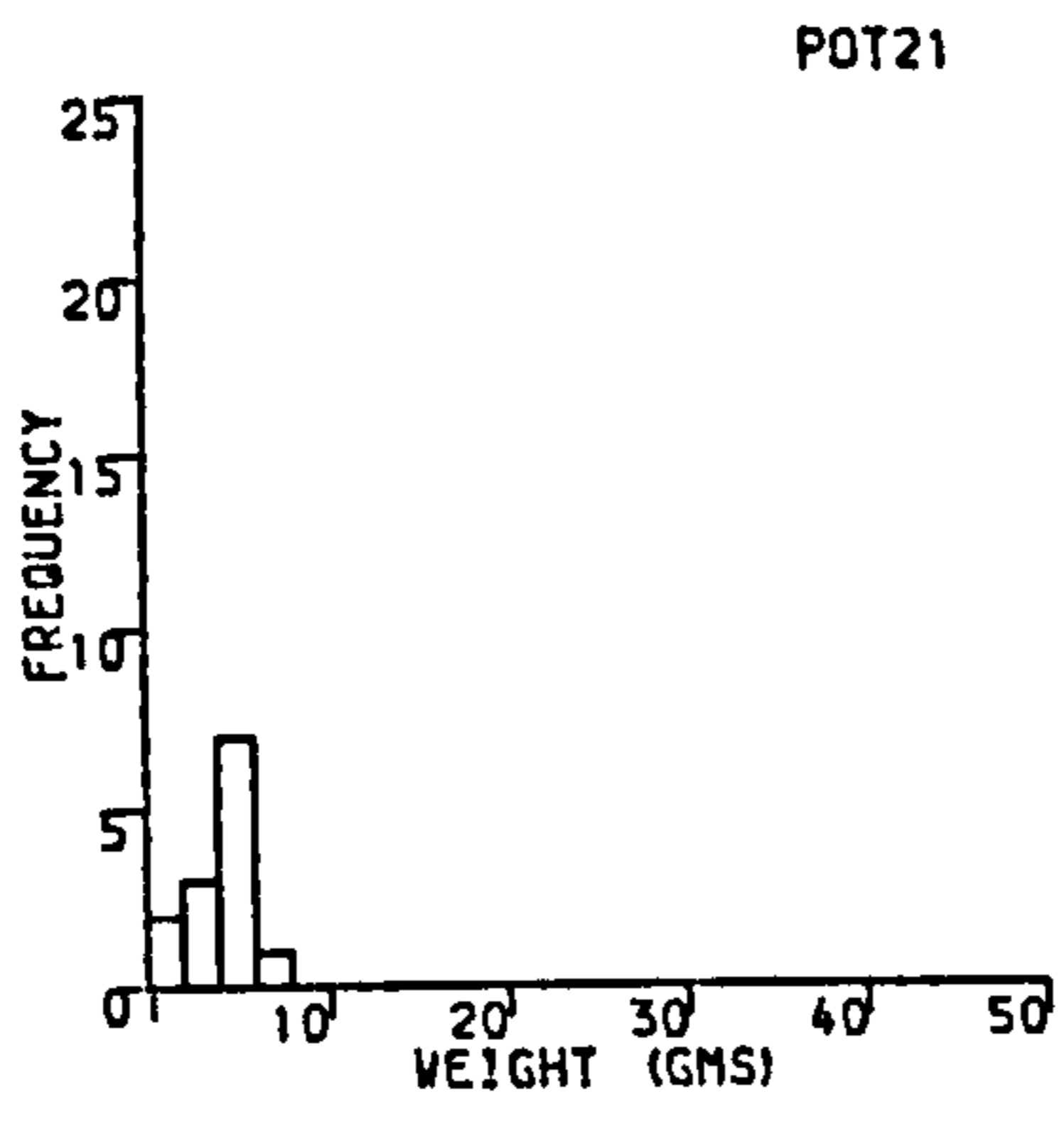


Figure 8.13 POT21 to POT25 weight histograms.

Block	alpha	beta
POT1	1.2669	6.4569
POT2	1.4631	4.6716
POT3	1.2702	5.5780
POT4	1.3295	11.7251
POT5	1.7076	1.7861
POT6	1.4014	5.5699
POT7	1.1448	7.3973
POT8	1.5244	3.9976
POT9	1.9191	4.1269
POT10	8.1241	1.8217
POT11	1.4993	3.0013
POT12	1.1414	6.6683
POT13	8.2290	0.5104
POT14	1.9535	4.8916
POT15	746.9067	0.0110
POT16	1.5102	5.0796
POT17	3.4605	1.6712
POT18	1.6863	3.0391
POT19	1.6912	6.3387
POT20	69.1198	0.0362
POT21	4.4348	0.8950
POT22	2.3370	1.6721
POT23	4.7048	0.8417
POT24	-	-
POT25	-	-

Table 8.2 Maximum likelihood estimates of Gamma distribution parameters.

1.27	1.46	1.27	1.33	-
1.40	1.14	1.52	-	-
1.50	1.14	-	1.95	-
1.51	-	1.69	-	-
4.43	2.34	-	-	-

Figure 8.14 parameter estimates.

8.5 DISCUSSION

This chapter has presented site association in terms of operands and operators. The use of this terminology helps to emphasise the way association, and archaeological procedures in general, are often structured. To apply simple association operators, operands need to be of the same type. For mixed operand types this can be achieved through their transformation to compatible data structures. Section 8.3 has also shown that quasi-inverse transforms of operands can be defined, which in themselves can have useful applications. Finally the difficulty of defining archaeologically relevant operands has been illustrated using the Heslerton surface pottery distribution data.

The idea of being able to manipulate data structures, as in the transform and quasi-inverse transform of operands, is important within archaeology. This same idea is found in many areas of mathematics and statistics where it is used to move data from one frame of reference to another. For example the aerial photography transformation of section 7.3 is a geometrical transform from the photograph plane to the map plane. The Fourier transform used in section 5.7 changes the frame of reference from a complex waveform to a summation series of simple sine or cosine functions. Although as pointed out in section 8.3 the use of the terms transform and inverse transform are not mathematically correct, they are still very useful in emphasising the manipulative potential of operands.

The manipulation of operands can sometimes present problems. Most operands will have associated estimates of error. When

transforming operands these error estimates should also be transformed. If they are not, then computations carried out in the new frame of reference may be completely fallacious. It is therefore probably better to record errors in a relative rather than absolute form, as this simplifies their transformation.

In general the more fundamental the data structure of an operand the greater the variety of potential operators that can be applied. However many areas of archaeology find it difficult to separate data from interpretation, and even when successful cannot decide at what level to record data. In fact this is the same problem of measuring archaeological significance discussed in section 8.4. For example much of so called 'problem orientated' archaeology is itself a problem. Very specific data is often collected to answer proposed problems. In the future these problems may be deemed irrelevant and the collected data may be too specialised to be applied to other problems. It is therefore of great importance to view data in terms of operands. The value of data can then be seen not as conventional measures of rarity but in terms of the range of transforms or manipulations that can be applied.

9.0 SITE ASSOCIATION OPERATORS

9.1 TYPES OF OPERATORS

To apply simple operators to site association, operands need to be of the same type. Operand types can be classified as either discrete or continuous, and this will determine the type of association operators that can be used. In archaeology discrete operators have tended to predominate spatial association, with techniques such as the analysis of contingency tables (Everitt, 1977) having wide spread use (e.g. Dacey, 1973; Rackerby, 1973; Clark, 1976). However the Heselton study operands are systematic samples from continuous functions (4.1) and therefore continuous association operators are needed.

Archaeologically association operators can be applied at two levels; global and local (Kelly and Haigh, 1983). At the global level the association between operands over the complete analysis area is measured. At the local level the association between operands is measured over sub-regions within the analysis area. Global association is the most commonly used technique and is primarily useful in synthesis, where the general association characteristics of sites are compared. Local association is much more useful for the detailed analysis of the intra-site patterning of operands. However local association operators have not yet been effectively developed within archaeology.

These global and local approaches to association analysis will be developed in the context of the Heselton study operands in the

following sections. Here the operands used will be abbreviated to the following:

GEO - the low density geophysical data of section 8.2.

POT - the pottery sherd weights of section 6.4.

AP - the aerial photography density measure of section 8.2.

9.2 GLOBAL OPERATORS

All global association operators are based on a statistical model. The simplest model is the linear one, where the association between two variables can be described by a straight line (figure 9.1.a). Other more complex models, such as the quadratic, cubic and quartic (figure 9.1.b to 9.1.d), can also be used to describe association. However, as in trend surface analysis (4.3), a polynomial of high enough degree could in theory be found to produce an apparent complete association between any two arbitrary operands. It is therefore important to select a suitable model based on criteria relating to the association problem and not to the desired degree of association. For the Heslerton study, the simple linear model seems the most suitable choice. Features, such as ditches, should have a high earth resistance (5.6), a high aerial photography density (8.2) and probably a high sherd weight density (6.4). The subsequent association measures will test these assumptions. This is a more objective approach than finding the model that gives the best association fit.

The simplest global operator is a correlation coefficient which measures the amount of association or correlation between operands, given a statistical model. This type of operator has limited use as

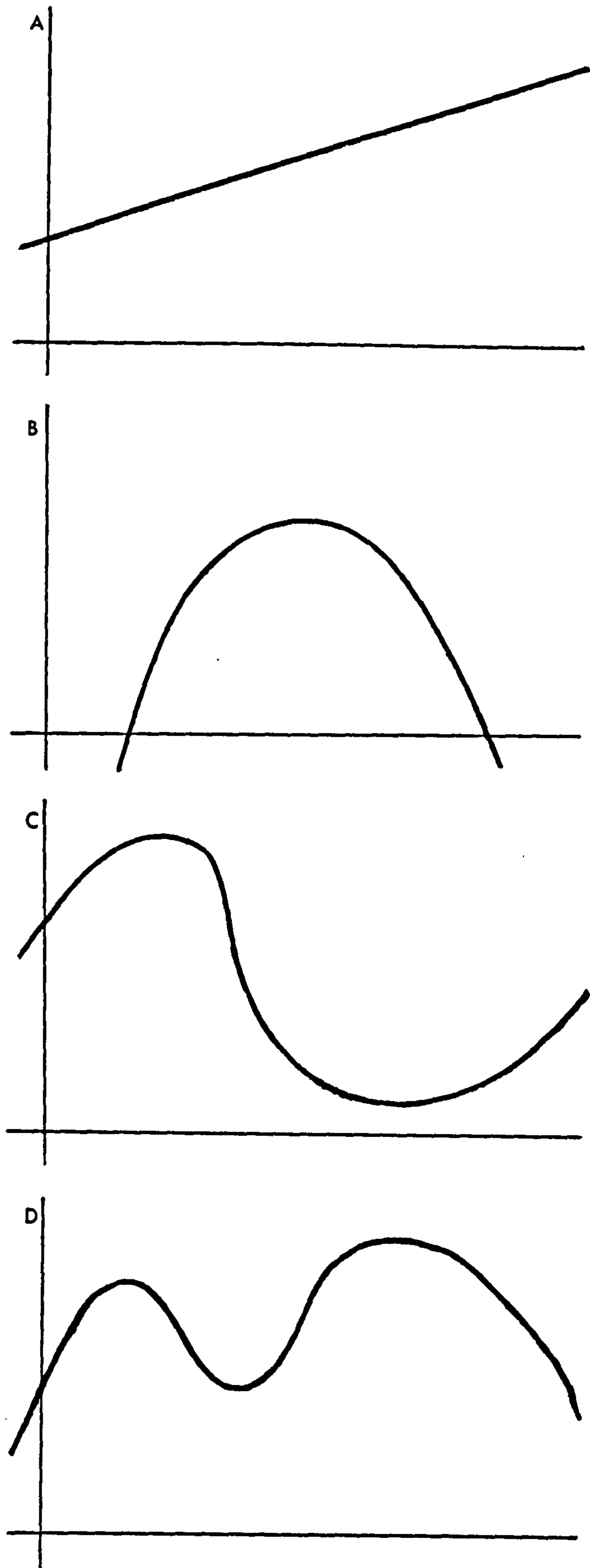


Figure 9.1 Association models. A) Linear
B) Quadratic C) Cubic D) Quartic.

although it can quantify overall association for a particular model, it can not define the parameters of that model. With regression analysis the statistical model parameters can be estimated, which then allows quite powerful comparison of association results. However neither of these operators takes account of the spatial locations of the data. This means that a given measure of association may reflect the degree of correspondence over the entire analysis area, or may be the result of a large deviation in a small sub-region. Using global operators these two possible situations are not easily distinguishable.

9.2.1 BIVARIATE CORRELATION

The Pearson product-moment correlation coefficient (r) is probably the most widely used measure of correlation between two continuous variates. If two variables x and y are assumed to have a bivariate Normal distribution, then the sample correlation coefficient r is defined as:

$$r = \frac{SXY}{\sqrt{SSX \cdot SSY}},$$

$$\text{where } SSX = \frac{\sum x^2 - (\sum x)^2}{n},$$

$$SSY = \frac{\sum y^2 - (\sum y)^2}{n},$$

$$SXY = \frac{\sum xy - \sum x \sum y}{n},$$

$n = \text{sample size.}$

The sample correlation coefficient will vary between -1 (perfect

negative correlation), 0 (no correlation) and 1 (perfect positive correlation). The sample correlation coefficients for the Heslerton study operands are shown in table 9.1. They seem to indicate that none of the operands are significantly correlated. This can be statistically tested by seeing how significantly different from 0 each r value actually is.

The sample correlation coefficient is assumed to be an estimate of the population correlation parameter ρ . The null and alternative hypotheses:

$$H_0: \rho = 0,$$

$$H_1: \rho \neq 0,$$

can be formulated. The Fisher-Z transformation:

$$Z_f = (1.1513) \cdot \log \left(\frac{1+r}{1-r} \right),$$

produces a variable Z_f that is approximately normal with mean and variance:

$$\mu_z = (1.1513) \cdot \log \left(\frac{1+\rho}{1-\rho} \right),$$

$$\sigma_z^2 = \frac{1}{n-3}.$$

This provides a test statistic:

$$Z = \frac{Z_f - \mu_z}{\sqrt{\sigma_z^2}},$$

which can be compared to standard Normal tables for a two tailed

	AP	GEO	POT
AP	-	-0.1036	0.2183
GEO	-	-	0.0090
POT	-	-	-

Table 9.1 Sample correlation coefficients.

	AP	GEO	POT
AP	-	-0.1040	0.2219
	-	-2.0700	4.4200
GEO	-	-	0.0090
	-	-	0.1800
POT	-	-	-
	-	-	-

Table 9.2 Correlation test statistics; top Zf, bottom Z.

Z	X	Y	t	df
AP	GEO	POT	-4.6959	397
GEO	AP	POT	-1.8050	397
POT	GEO	AP	-2.8779	397

Table 9.3 Prediction test statistics.

test. The test statistics for the Heslerton study operands are shown in table 9.2. These results show that the GEO and POT operands are not significantly correlated at the 5% significance level. The GEO and AP operands are significantly correlated at the 5% significance level but not at the 1% significance level. The AP and POT operands are significantly correlated at the 1% significance level. The GEO operand is associated with both the non-significant correlations, it is therefore useful to test to see if this association is itself significant.

Hotelling (1940) has devised a test to see if a variable z is predicted better by a variable x or y . The null and alternative hypotheses:

$$H_0: \rho_{xz} = \rho_{yz},$$

$$H_1: \rho_{xz} \neq \rho_{yz},$$

are formulated. Under the assumption that the conditional distributions of z are Normal with the same variance for all values of x and y , the test statistic:

$$t = \frac{(rxz - ryz) \sqrt{(n-3)(1+rxy)}}{\sqrt{2(1-rxy^2 - rxz^2 - ryz^2 + 2rxy \, rxz \, ryz)}}$$

has a t distribution with $n-3$ degrees of freedom. The test statistics for the Heslerton study operands are shown in table 9.3. They show that for the AP operand there is a significant difference in prediction using the GEO and POT operands at the 1% significance level. For the POT operand there is also a significant difference in prediction using the GEO and AP operands at the 1% significance level. While for the GEO operand there is a significant difference

in prediction using the AP and POT operands at the 5% significance level, but not at the 1% significance level. This confirms the conclusions of the previous test results, that the AP and POT operands show a degree of global association but seem to be unassociated with the GEO operand.

9.2.2 MULTIVARIATE CORRELATION

The Pearson product-moment correlation coefficient can be extended to the multivariate case (Lindeman, Merenda and Gold, 1980). Here the multivariate correlation is based on the same n individuals. Given a multivariate Normal distribution of variates:

$$y_1, x_1, x_2, \dots, x_k,$$

a sample multivariate correlation coefficient:

$$R_{y,1,2,\dots,k},$$

can be calculated as an estimate of the population multivariate correlation parameter:

$$\rho_{y,1,2,\dots,k}.$$

The sample multivariate correlation coefficient is calculated as:

$$R_{y,1,2,\dots,k} = \sqrt{C^T R^{-1} C},$$

$$\text{where } C = \begin{bmatrix} r_{1y} \\ r_{2y} \\ \cdot \\ \cdot \\ r_{ky} \end{bmatrix},$$

r_{ij} is the sample bivariate correlation coefficient
between y and x_i ,

$$R = \begin{bmatrix} 1 & r_{12} & \cdot & \cdot & r_{1k} \\ r_{12} & 1 & \cdot & \cdot & r_{2k} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{1k} & r_{2k} & \cdot & \cdot & 1 \end{bmatrix},$$

r_{ij} is the sample bivariate correlation coefficient
between x_i and x_j .

The sample multivariate correlation coefficient will vary between 0
(no correlation) and 1 (perfect correlation). For the Heslerton
study operands let:

$$y = AP,$$

$$x_1 = GEO,$$

$$x_2 = POT.$$

The matrices C and R equal:

$$C = \begin{bmatrix} -0.1036 \\ 0.2183 \end{bmatrix},$$

$$R = \begin{bmatrix} 1 & 0.0090 \\ 0.0090 & 1 \end{bmatrix}.$$

The calculated sample multivariate correlation coefficient is:

$$R_{y,1,2} = 0.2425.$$

To test to see if this correlation is significant the null and alternative hypotheses:

$$H_0: \rho_{y,1,2..k}=0,$$

$$H_1: \rho_{y,1,2..k} \neq 0,$$

can be formulated. The test statistic:

$$F = \frac{(n-k-1) R^2_{y,1,2..k}}{k(1-R^2_{y,1,2..k})},$$

has an F-distribution with k , $n-k-1$ degrees of freedom. For the Heslerton study F equals 12.4023, which is not significant at the 5% significance level. This is not surprising as the GEO operand does not seem to be correlated with either the POT or AP operands as indicated in section 9.2.1.

9.2.3 BIVARIATE REGRESSION

Given an independent variable x and a dependent variable y , a least-squares linear regression of y on x can be formulated as:

$$y = a + bx,$$

$$\text{where } b = \frac{S_{XY}}{S_{XX}},$$

$$a = \bar{y} - b\bar{x}.$$

The b coefficient in the regression equation represents the gradient of the fitted line, and the a coefficient represents the intercept. The calculated regression coefficients for the Heslerton study operands are shown in table 9.4 and are illustrated in figure 9.2. A standard test in regression analysis is to determine if y is

	AP	GEO	POT
AP	- -	0.4151 -0.0019	0.0738 0.0021
GEO	153.5135 -5.5655	- -	152.7395 0.0047
POT	19.4408 22.6031	19.5227 0.0173	- -

Table 9.4 Regression analysis; top a coefficient, bottom b coefficient, columns independent variables, rows dependent variables.

	AP	GEO	POT
AP	-	-2.0478	4.4447
GEO	-2.0777	-	0.1807
POT	4.4634	0.1791	-

Table 9.5 Regression test statistic t-values with 398 df; columns independent variables, rows dependent variables.

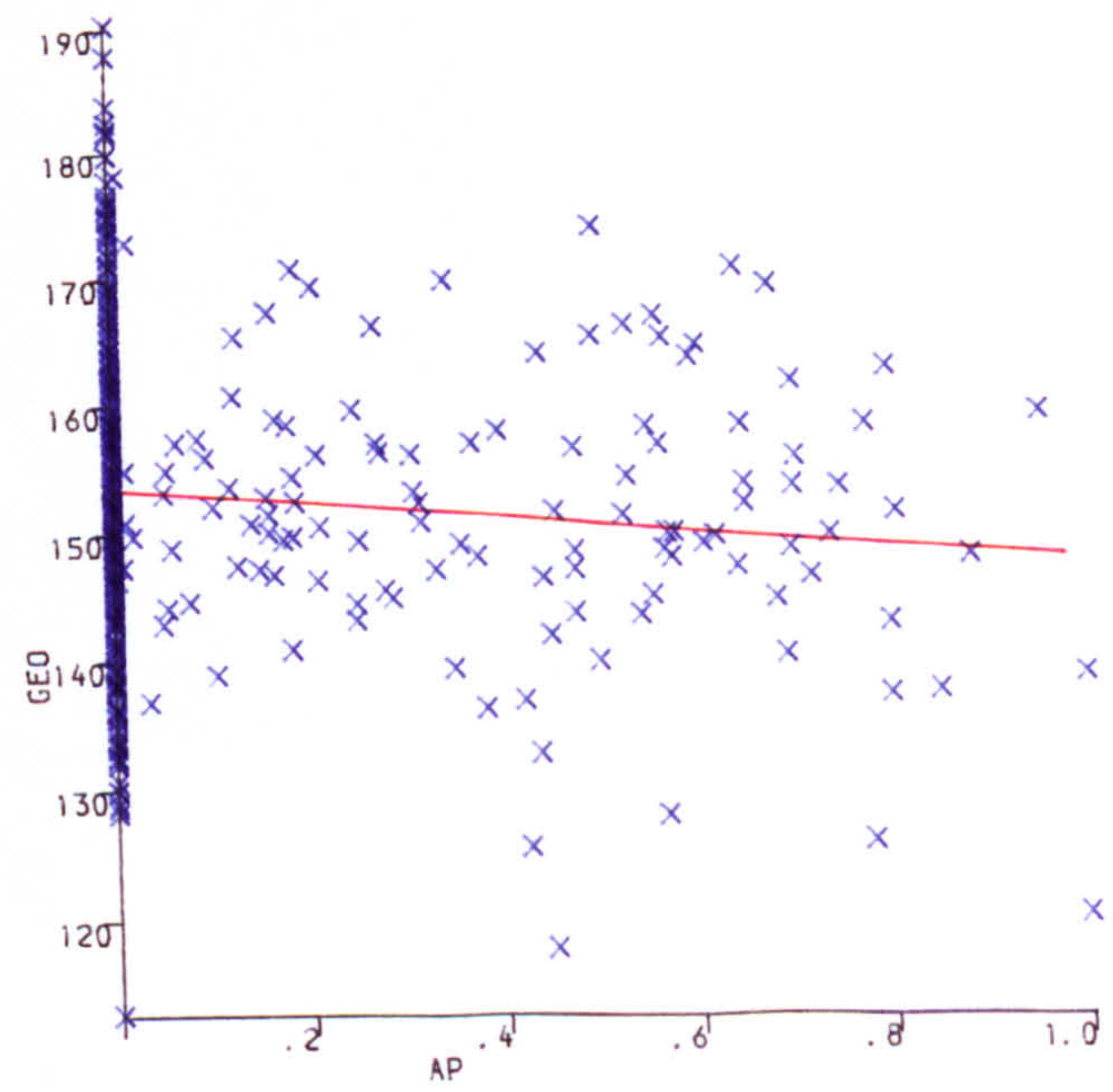
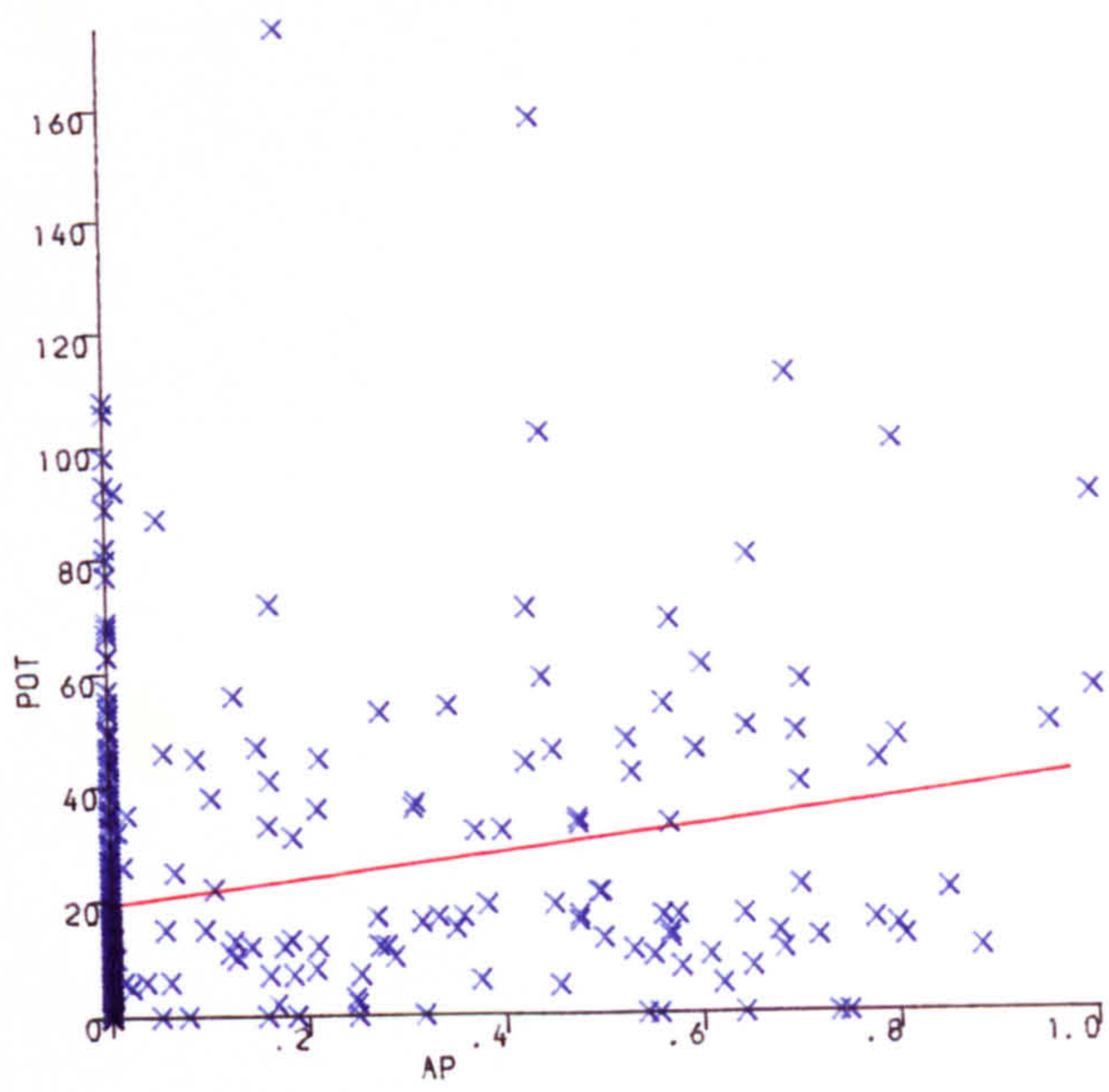
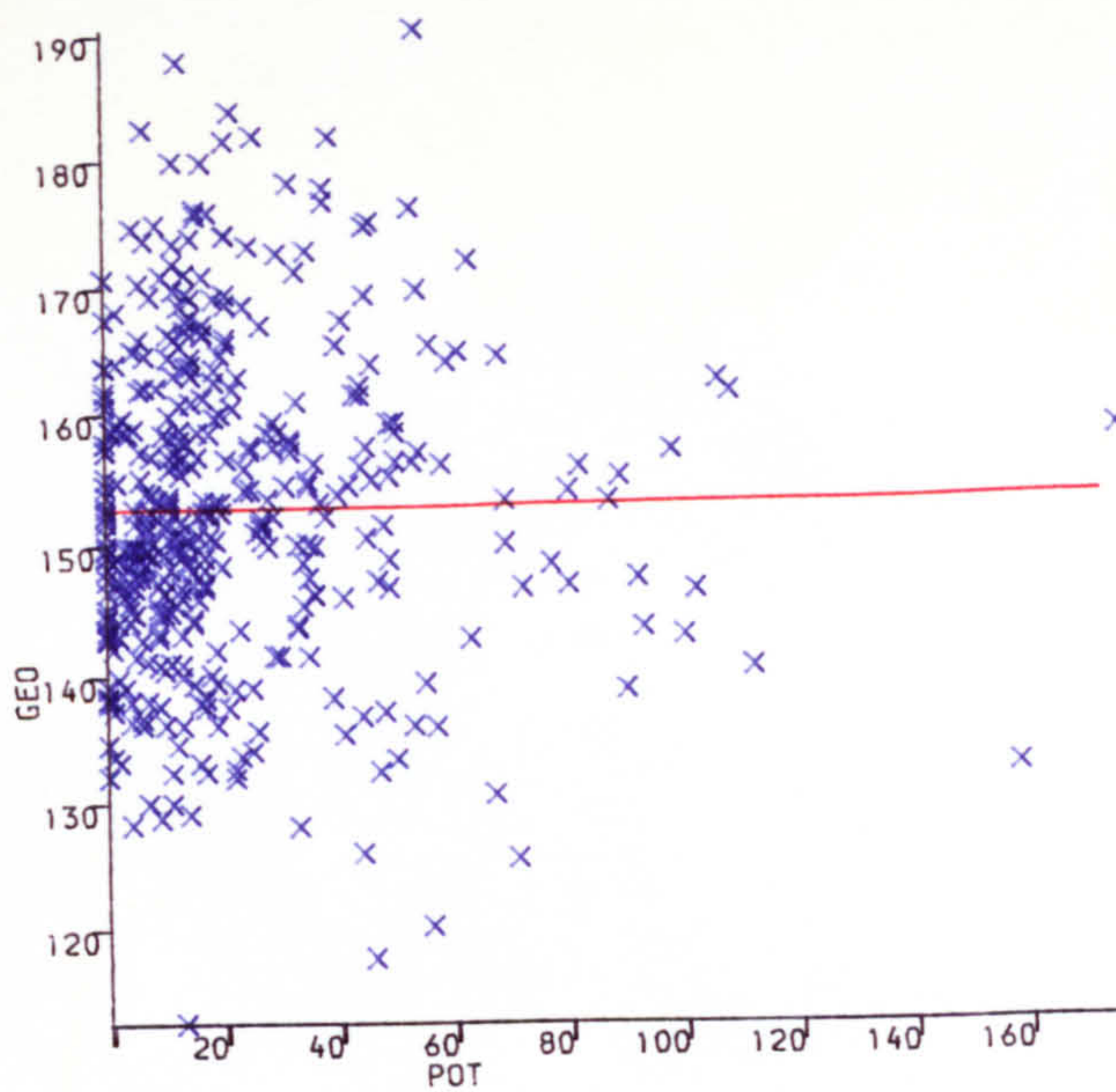


Figure 9.2 Bivariate regression.

independent of x . This is statistically determined by seeing how significantly different from 0 the calculated b coefficient is. The sample a and b coefficients are assumed to be estimates of the population regression parameters α and β . The null and alternative hypotheses:

$$H_0: \beta = 0,$$

$$H_1: \beta \neq 0,$$

can be formulated. The test statistic:

$$t = \frac{b - \beta}{S_b},$$

$$\text{where } S_b = \sqrt{\frac{SSE}{(n-2) SSX}},$$

$$SSE = SSY - \frac{(SXY)^2}{SSX},$$

has a t distribution with $n-2$ degrees of freedom. The test statistics for the Heslerton study operands are shown in table 9.5. These results show that the GEO and POT operands are not significantly dependent at the 5% significance level. The GEO and AP operands are significantly dependent at the 5% significance level but not at the 1% significance level. The AP and POT operands are significantly dependent at the 1% significance level. These regression results support the correlation analyses of sections 9.2.1 and 9.2.2.

9.2.4 MULTIVARIATE REGRESSION

Given k independent variables, x_1 to x_k , and a dependent variable

y, the least-squares linear regression analysis can be extended to the multivariate case. A multiple linear regression analysis can be formulated as:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k.$$

The b coefficients in the regression equation can be calculated by:

$$B = (RD)^{-1}C ,$$

$$b_0 = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 - \dots - b_k\bar{x}_k,$$

where R and C are defined as in section 9.2.2,

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_k \end{bmatrix} ,$$

$$D = \begin{bmatrix} d_1 & 0 & \cdot & \cdot & 0 \\ 0 & d_2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & d_k \end{bmatrix} ,$$

$$d_i = \frac{\text{standard deviation of } x_i}{\text{standard deviation of } y} .$$

For the Heslerton study operands let:

$$y = AP,$$

$$x_1 = GEO,$$

$$x_2 = POT.$$

The matrices R, D and C equal:

$$R = \begin{bmatrix} 1 & 0.0090 \\ 0.0090 & 1 \end{bmatrix} ,$$

$$D = \begin{bmatrix} 7.3300 & 0 \\ 0 & 10.1747 \end{bmatrix} ,$$

$$C = \begin{bmatrix} -0.1036 \\ 0.2183 \end{bmatrix} .$$

The calculated B matrix and b0 coefficient are:

$$B = \begin{bmatrix} -0.0144 \\ 0.2183 \end{bmatrix} ,$$

$$b_0 = 1.8471 .$$

This gives a multiple linear regression equation of:

$$y = 1.8471 - 0.0144 x_1 + 0.0214 x_2 ,$$

which is illustrated as a least squares regression plane in figure 9.3.

9.3 LOCAL OPERATORS

A local association operator measures the spatial association between operands within defined sub-regions of an analysis area. Where these sub-regions can be firmly defined by archaeological boundaries, the techniques of section 9.2 can be employed. For example an excavation may reveal several hut circles, these can act as archaeological boundaries which allow the association between operands within each hut to be measured and compared to each other. However for the Heslerton study area, and for most archaeological

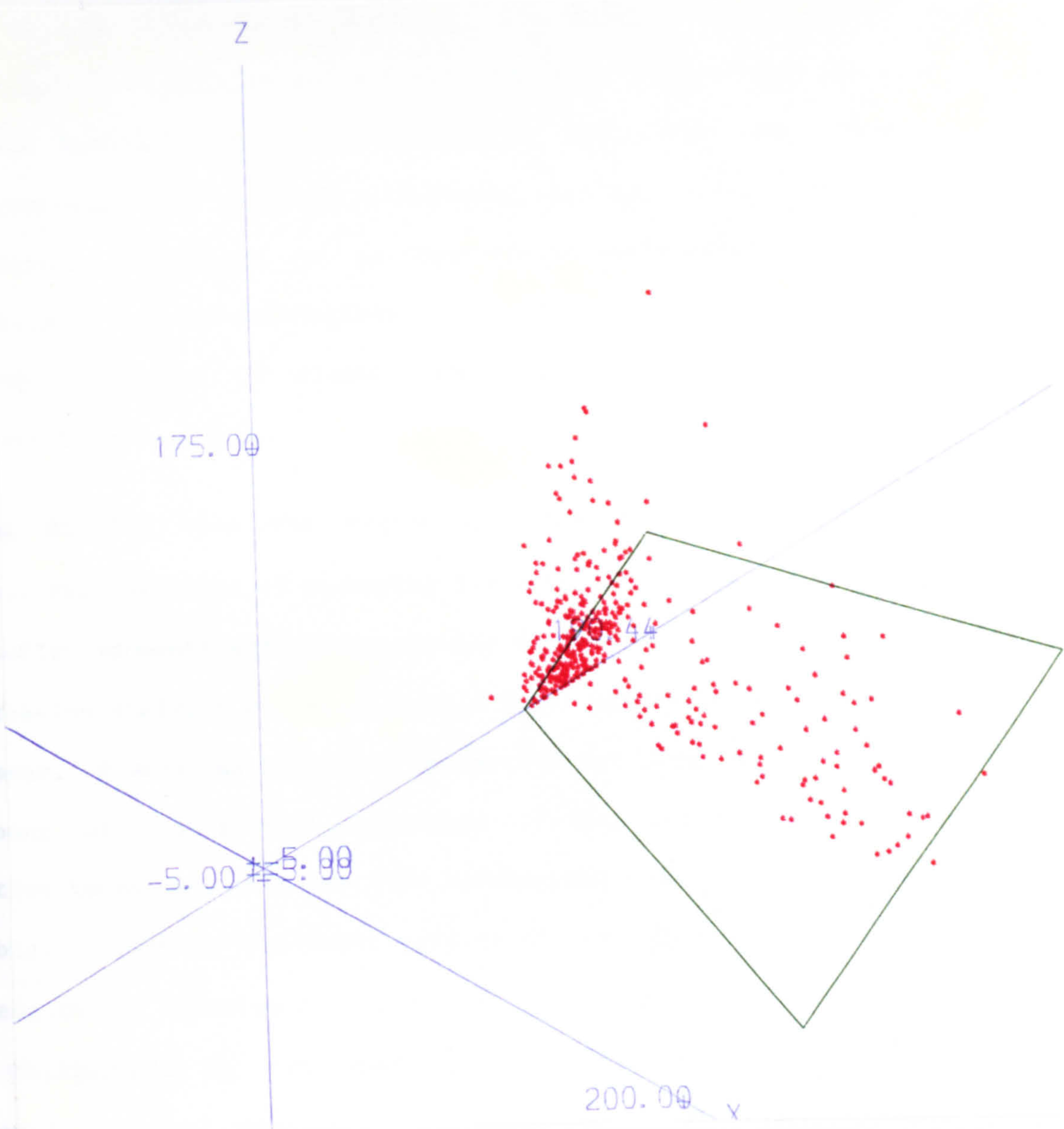


Figure 9.3 Multivariate regression.
 X = AP X 200, Y = GEO, Z = POT.

field surveys, no definable archaeological boundaries exist. It is therefore necessary to use an arbitrary sub-region definition and position the sub-regions in a systematic pattern over the analysis area. This creates problems because as Johnson (1977) has pointed out, the sub-region definition, especially its size, will effect the results of any association analysis. In fact the problems of quadrat sampling of section 6.2 are encountered. Some authors (e.g. Kintigh and Ammerman, 1982; Simek and Larick; 1983) have suggested a cluster analysis based approach to defining natural local spatial concentrations. Although it is feasible to apply association to these naturally defined sub-regions, it would not be beneficial as the basic technique of cluster analysis is of very dubious objectivity (Orton, 1980).

Having decided upon the degree of localisation for local analysis, the problems of measuring the association and displaying it are left. As mentioned previously the techniques of correlation and regression analysis can be used at a local level for association measurement. However as indicated in section 9.2 these methods take no account of the spatial locations of the operand readings. Association techniques that use this information would therefore be preferable. Given a systematic pattern of association results the data presentation techniques of 4.0 seem suitable for displaying them. Furthermore as indicated in section 4.2, the best display technique is probably contouring as implemented in BAGS (4.4).

9.3.1 SIMILARITY MAPS

The easiest and most localised method of local association

analysis is to compare two operands on the basis of their individual readings. This is done by taking the difference between the two operands and mapping it. This technique is used in geology (Davis, 1973) and is known as; similarity maps, difference maps or isopach maps.

The technique is very simple to use if the operands are of the same type and are measured in the same units. If they are not then either one operand must be converted to the units of the other, or both operands must be converted to a standardised form. The first alternative is often preferred by geologists as the final similarity map will be in units that are familiar. However problems exist in the conversion process. If operand X is to be converted to the form of operand Y, the regression of Y on X needs to be initially computed. The regression equation is then used to convert X values to \hat{Y} values. The similarity map $\hat{Y} - Y$ can then finally be drawn. As pointed out in section 9.2 a regression analysis may not reflect the overall association of an analysis area and can thus potentially produce erroneous \hat{Y} values. It is also a problem in deciding which operand to use as the estimator in the regression equation.

For these reasons it is therefore probably more useful to convert the operands to their standardised forms:

$$z = \frac{x - \bar{x}}{s},$$

where z = standardised reading,

x = operand reading,

\bar{x} = mean of operand readings,

s = standard deviation of operand readings.

The similarity maps constructed from standardised operands need however careful interpretation. It is the magnitude and not the sign of the difference that reflects the similarity between operands. Two similar operands will give small absolute difference values, while two dis-similar operands will give large absolute difference values. The use of standardised operands also allows the simple computation of a global measure of association, by multiplying the standardised operands together. If both operands deviate from their means in the same direction then the product will be positive, while if they deviate in opposite directions then the product will be negative.

The similarity maps for the standardised operands AP - GEO, AP - POT and GEO - POT are shown in figures 9.4, 9.5 and 9.6. Figure 9.4 shows several distinct areas of disassociation between the AP and GEO operands. This is not surprising given the dis-similar structures of these operands (see figures 8.3 and 8.5). Figure 9.6 shows a similar situation for the GEO and POT operands, although the extent of the disassociation here is some what more restricted. Figure 9.5 presents a more interesting picture. There are distinct areas of both disassociation and association between the AP and POT operands visible. The most notable disassociation occurs in the north west corner of the study area. Here there is a linear structure of high density pottery and low density aerial photography features. This could imply an occupation area with aerial photography features defining the boundaries of physical structures and high pottery densities defining the internal activity areas.

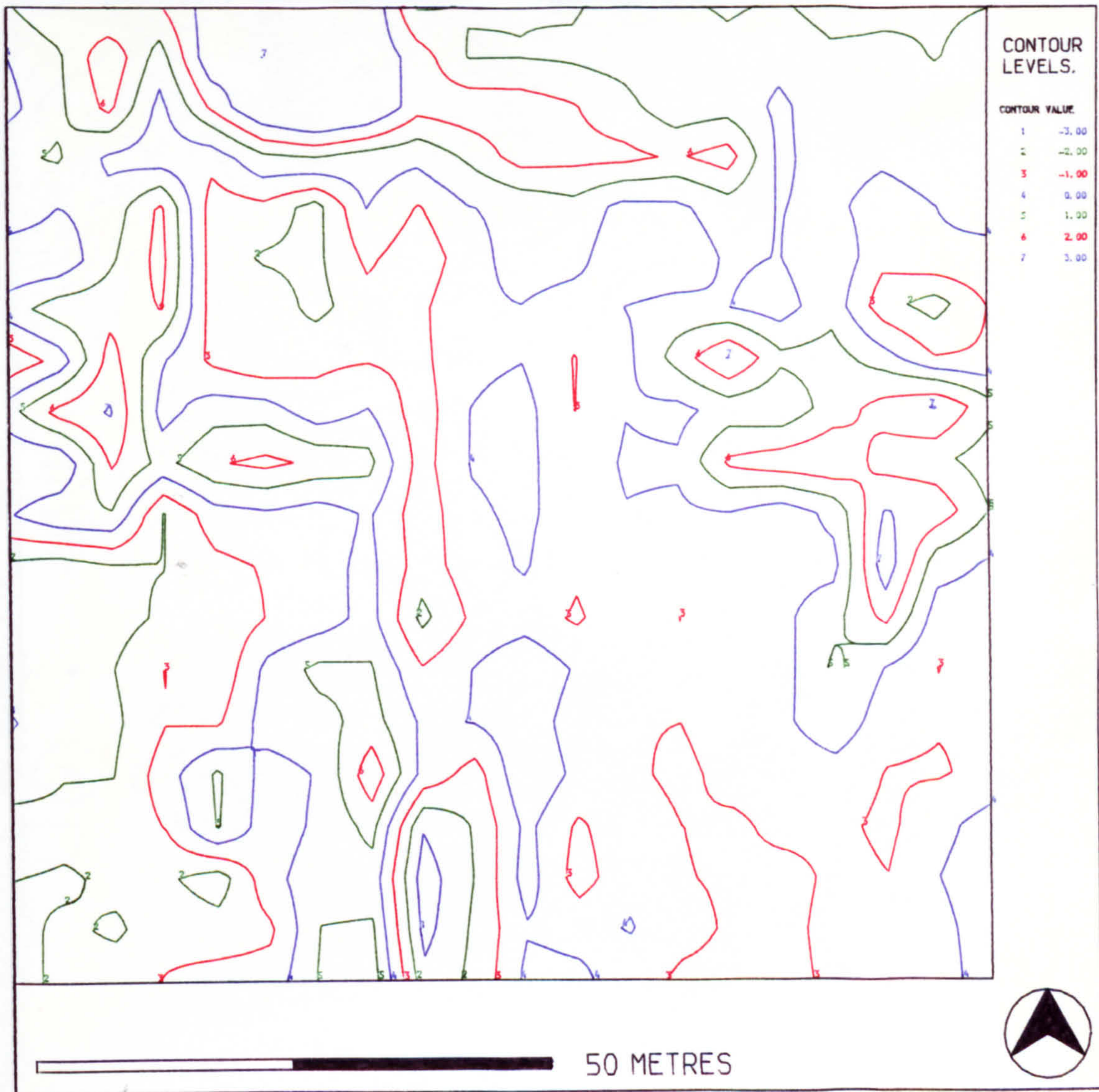


Figure 9.4 Standardised AP - GEO similarity map.

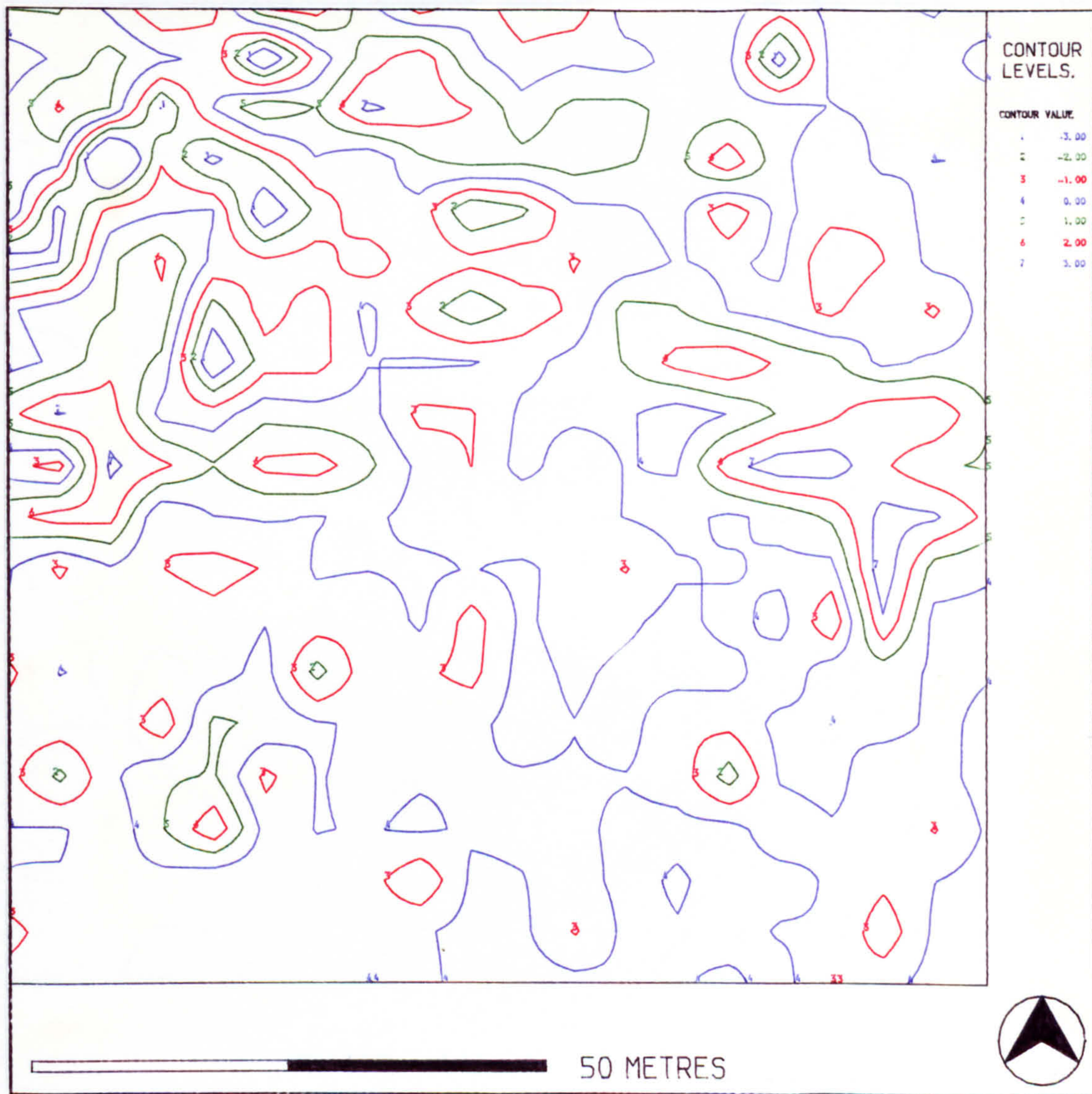


Figure 9.5 Standardised AP - POT similarity map.

Standardised GEO - POT similarity map

Firstly
either
of as
context
product
Second
signifi
probabi
therefo
previo
are pot
The
Merriam
global
trend
As dis
surface

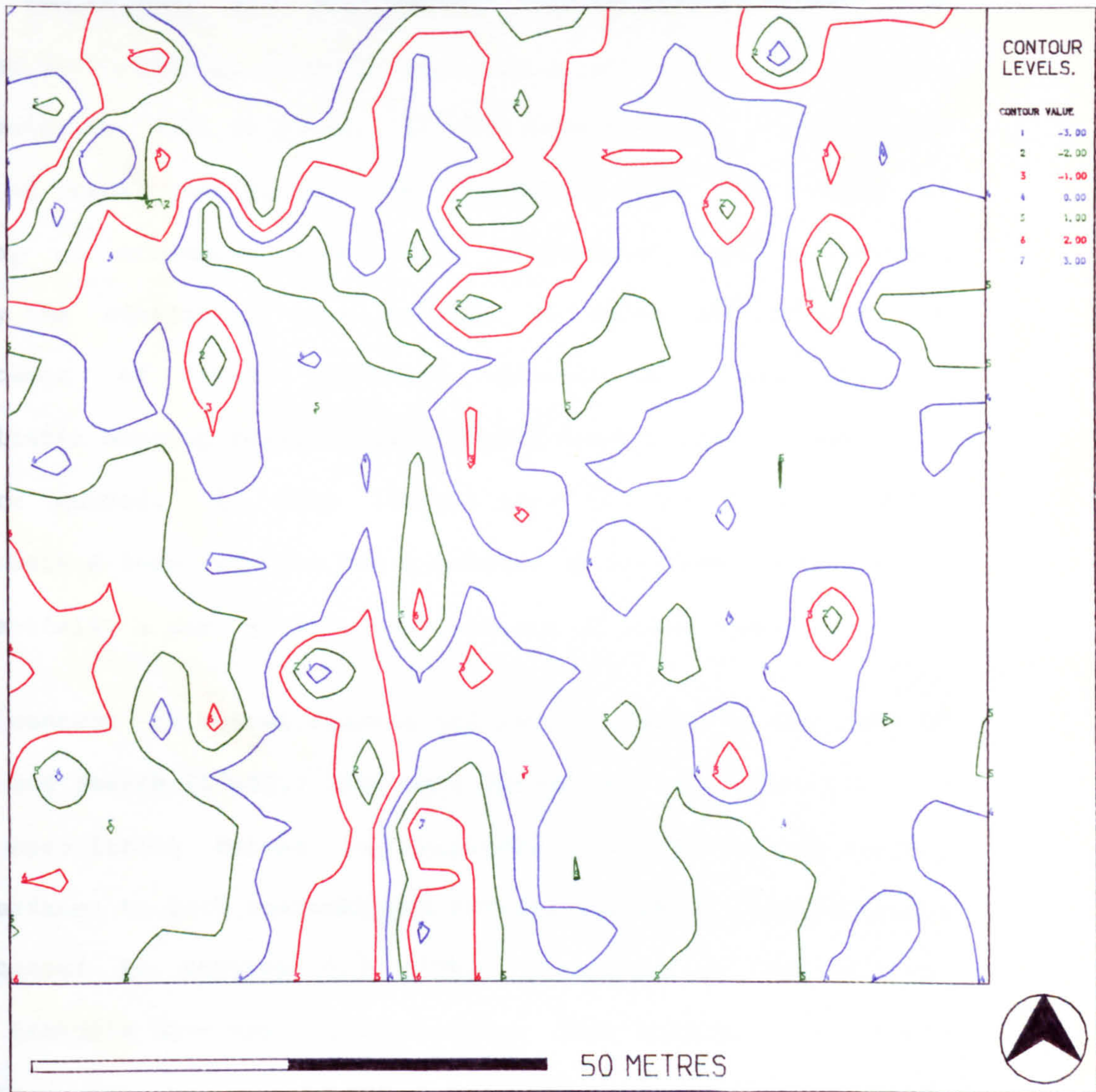


Figure 9.6 Standardised GEO - POT similarity map.

contouring
all the data
point where
than data more distant.

9.3.2 LOCAL ASSOCIATION MAPS

Although similarity maps provide a simple and obvious means of measuring local association they are not entirely satisfactory. Firstly comparisons are only made between individual readings, if either or both readings carry substantial noise then a false measure of association will be given. If both readings were placed in the context of their contiguous local readings this possibility of producing an erroneous association measurement could be reduced. Secondly the similarity maps provide no means of judging the significance of their measured associations. Some form of probabilistic mapping relative to a random unassociated situation is therefore needed. In this context local association maps, which provide both a local context and a measure of relative significance, are potentially a more satisfactory measure of local association.

The concept of local association maps is based on the ideas of Merriam and Sneath (1966). They were interested in measuring the global association between two surfaces. They did this by fitting trend surfaces to both operands and correlating their coefficients. As discussed in section 4.3 global functions as used in trend surface analysis have many disadvantages. More suitable for local association is the use of locally fitted functions. However the form of the fitting presents problems. Splines as were used for contouring (4.3) could be used, but they provide equal weighting to all the data used. For a set of data points those data close to the point where a function is to be fitted should have greater influence than data more distant. A weighted fit is therefore desirable.

A good general purpose local function is the quadratic of two variables:

$$k(x,y) = a + bx + cy + dx^2 + exy + fy^2.$$

Given a gridded data set, such a quadratic can be fitted to the data at each grid point by a standard least-squares fitting procedure, using a weighting function $W(D)$ where D is the distance from the grid point to the data concerned. As indicated, $W(D)$ should tend to zero for large values of D , so that the influence of distant data points are negligible. Simple forms of $W(D)$ include the inverse-linear and inverse-square, which assume that the grid point will take the same value as a data point coincident with it. The more general form of $W(D)$:

$$W(D) = (D + \alpha)^{-n},$$

allows an additive constant α to be included so that smoothing of grid values can be achieved if required. It is also advisable to specify a maximum cut off distance beyond which data points will not contribute to the quadratic fit. This helps reduce computations and also provides a hard edge boundary which is useful in the subsequent interpretation of results.

Although the fitting of a local quadratic can be applied directly to the operands, it is probably more sensible to use the standardised operands for the reasons outlined in section 9.3.1. The procedure defined so far therefore involves the standardisation of both operands, the definition of a weighting function and cut off distance, and the fitting of a local quadratic function at each grid point for both operands. For each quadratic fit the function

coefficients need to be correlated. As Merriam and Sneath (1966) have pointed out the initial coefficient (a) in any fitted function should not contribute to this correlation as it will unduly influence the result. A simple association measure that can be used is based on taxonomic distance (TD):

$$TD = \sqrt{\frac{S}{n}},$$

where $S = \sum(\text{coefficient } i - \text{coefficient } i')^2$,

i refers to first operand,

i' refers to second operand,

n = number of coefficients used.

For two functions that have similar coefficient values TD will be small, while for two functions that have dis-similar coefficient values TD will be large.

Although the taxonomic distance gives a measure of association between two fitted functions, it gives no criteria with which to judge significance. However using Monte Carlo simulation techniques the distribution of the taxonomic distances can in theory be approximated. This however involves modelling the processes that formed the operands, a task that is very complicated. A simpler relative measure of significance can be calculated if the coefficients of the fitted quadratics for both operands are recorded. By correlating random pairs a distribution of taxonomic distances for a randomized function match of the two recorded operands can be achieved. Using the mean and standard deviation of this simulated distribution, an observed taxonomic distance can be

standardised:

$$z = \frac{TD - \bar{x}}{s},$$

where z = standard Normal deviate,

\bar{x} = mean of simulated distribution,

s = standard deviation of simulated distribution.

The standardised distance can then be compared to standard Normal tables to judge the relative significance of the observed taxonomic distance. Or alternatively the standardised distances can be drawn as a relative probability map with various levels of significance displayed.

Both local association and relative probability maps have been implemented in a program called `local_a_map`. The source program is coded in Fortran 77 and its structure is outlined in appendix A.6. The weighting function that was used:

$$W(D) = (1 + D)^{-2},$$

gives an inverse square weighting that provides a degree of smoothing. The choice of weighting function and also cut off distance obviously presents a problem in that they will influence the function fitting and hence the subsequent association measure. From the mathematical view point there are no right or wrong choices. The definition of the extent and exact form of local association analysis must be based on both practical and archaeological considerations, which hopefully should not be in conflict.

Both local association and relative probability maps as computed

by local_a_map are shown in figures 9.7 to 9.12 for the standardised AP - GEO, AP - POT and GEO - POT operands. Two cut off distances of 12.5 and 17.5 metres have been used as they offer reasonable local analysis areas of respectively; approximately 5% and 10% of the Heselton study area. Figure 9.7 shows no significant areas of association between the AP and GEO operands with a 12.5 metre cut off. There are however two significant areas of disassociation, one in the north of the analysis area and the other on the east edge. This same pattern is shown with a 17.5 metre cut off in figure 9.8, although it is more smoothed as would be expected. Figure 9.9 shows no significant areas of association between the AP and POT operands with a 12.5 metre cut off. There are significant areas of disassociation along the north edge of the study area and towards the south east. These same areas are seen with a 17.5 metre cut off in figure 9.10, where they seem to indicate a possible linear band of disassociation running from the south west to the north. Figure 9.11 shows a significant area of disassociation between the GEO and POT operands with a 12.5 metre cut off along the north edge of the study area. With a 17.5 metre cut off a further area of significant disassociation appears along the east edge in figure 9.12.

9.4 DISCUSSION

The results of the global and local association analyses of sections 9.2 and 9.3 provide the basis for a set of recommendations for the extension of field survey into further areas of the Heselton Parish Project (see 1.0). The geophysical data at both the global and local scale shows no strong relationships with any of the other site operands. This indicates that the marginal survey

Figure 9.7 Local association map of standardised AP - GEO with cut off at 12.5 metres. A) Association map. B) Relative probability map.

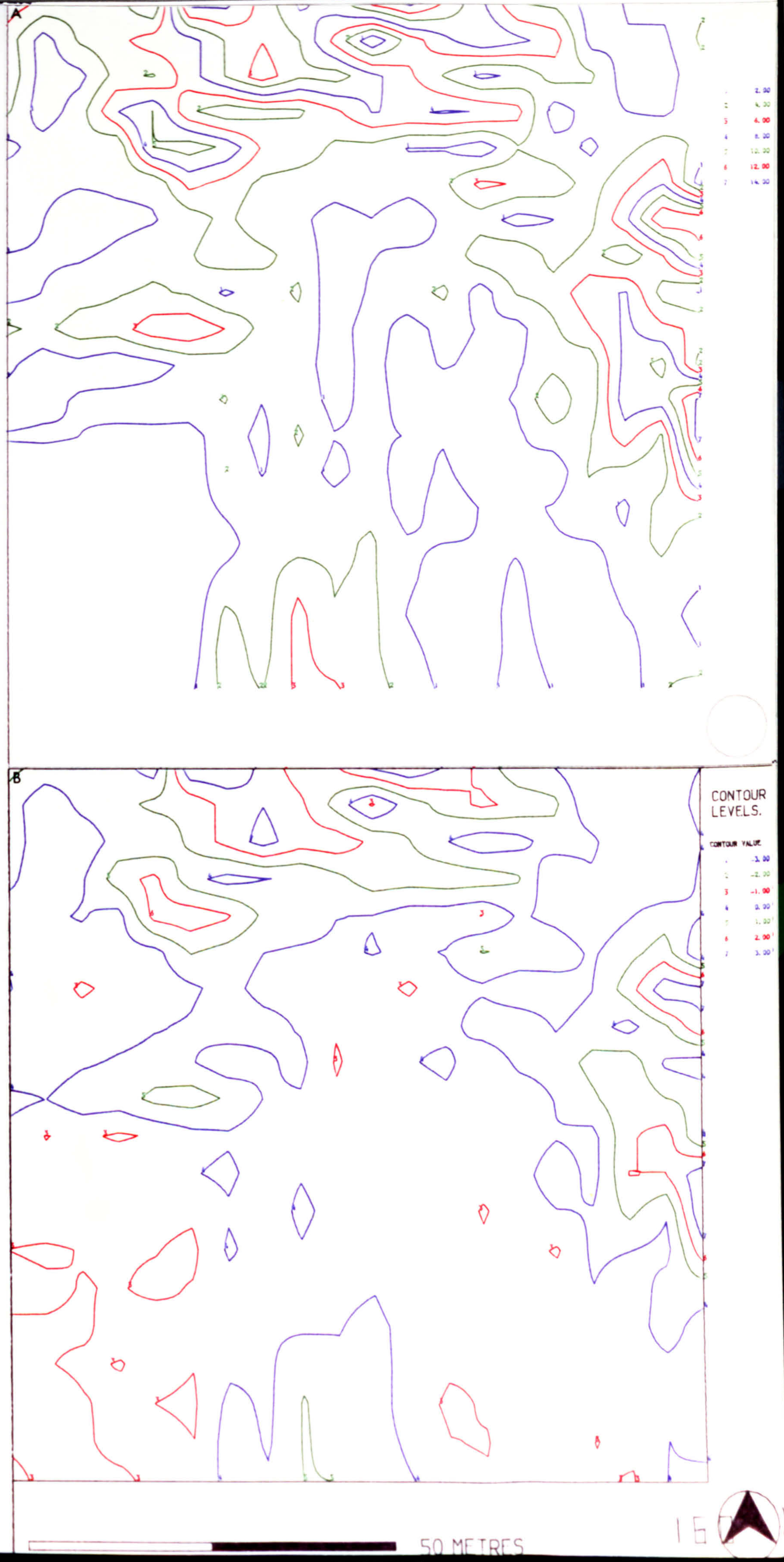


Figure 9.8 Local association map of standardised AP ~ GEO with cut off at 17.5 metres. A) Association map. B) Relative probability map.

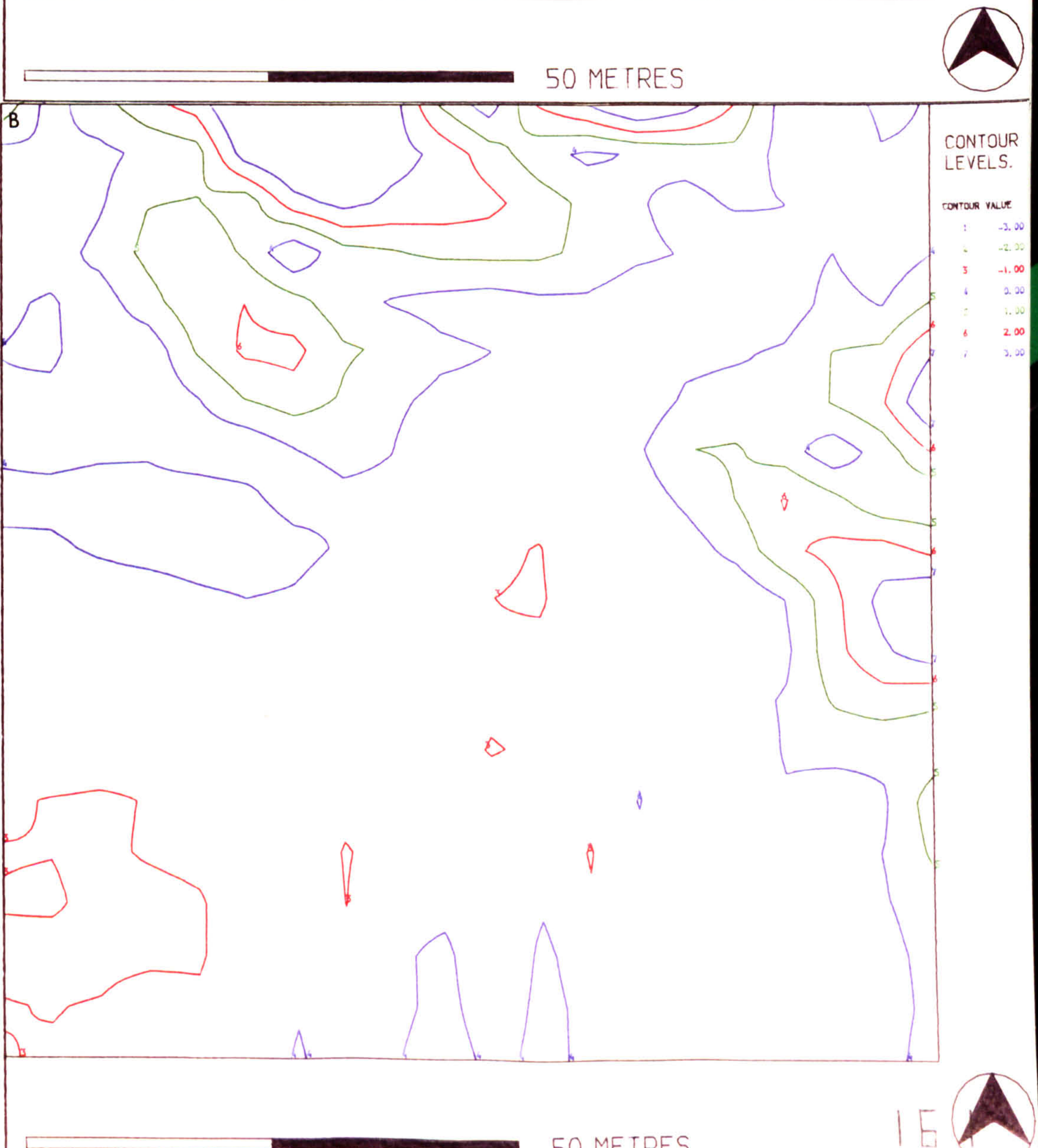
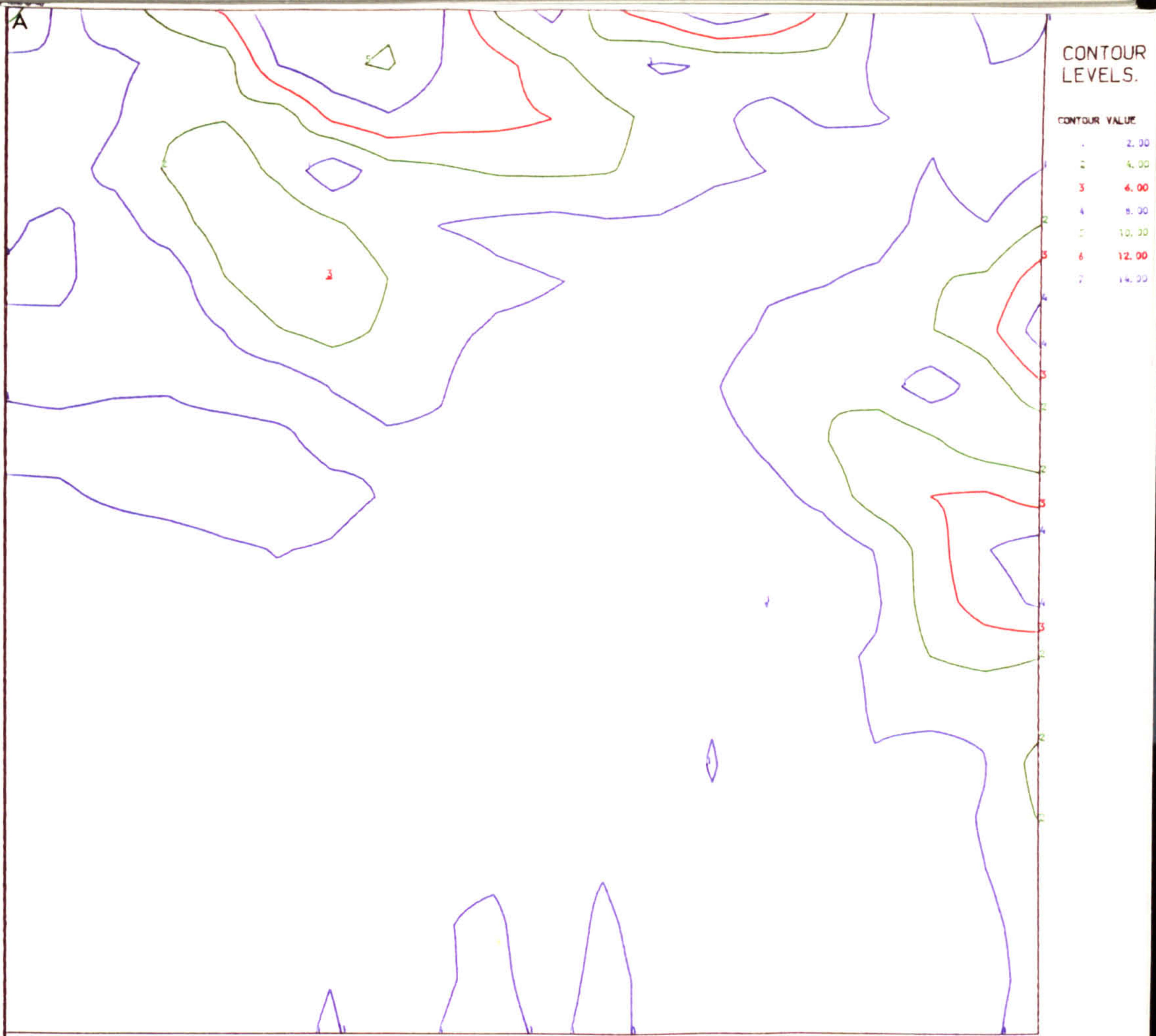


Figure 9.9 Local association map of standardised AP - POT with cut off at 12.5 metres. A) Association map. B) Relative probability map.

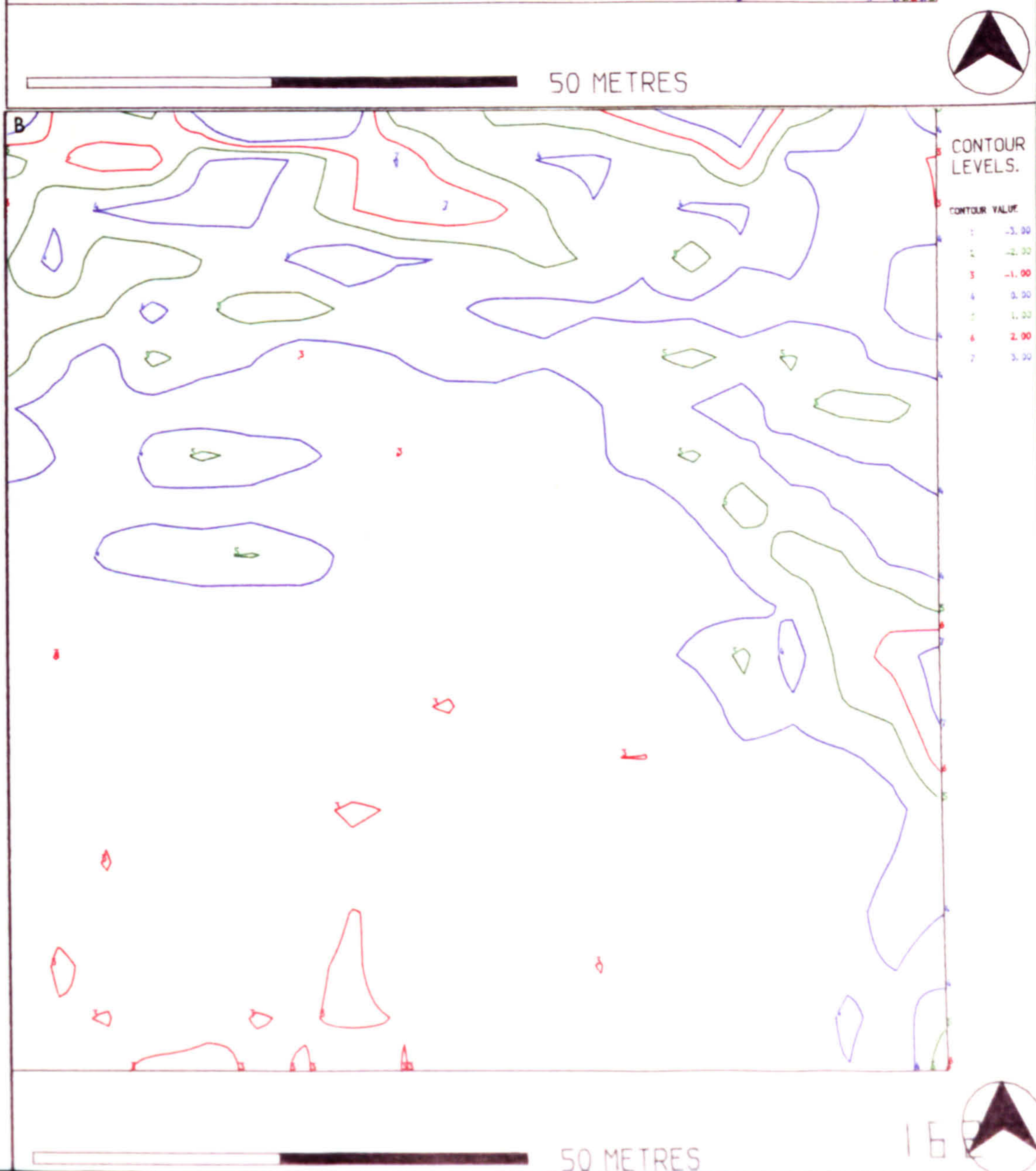
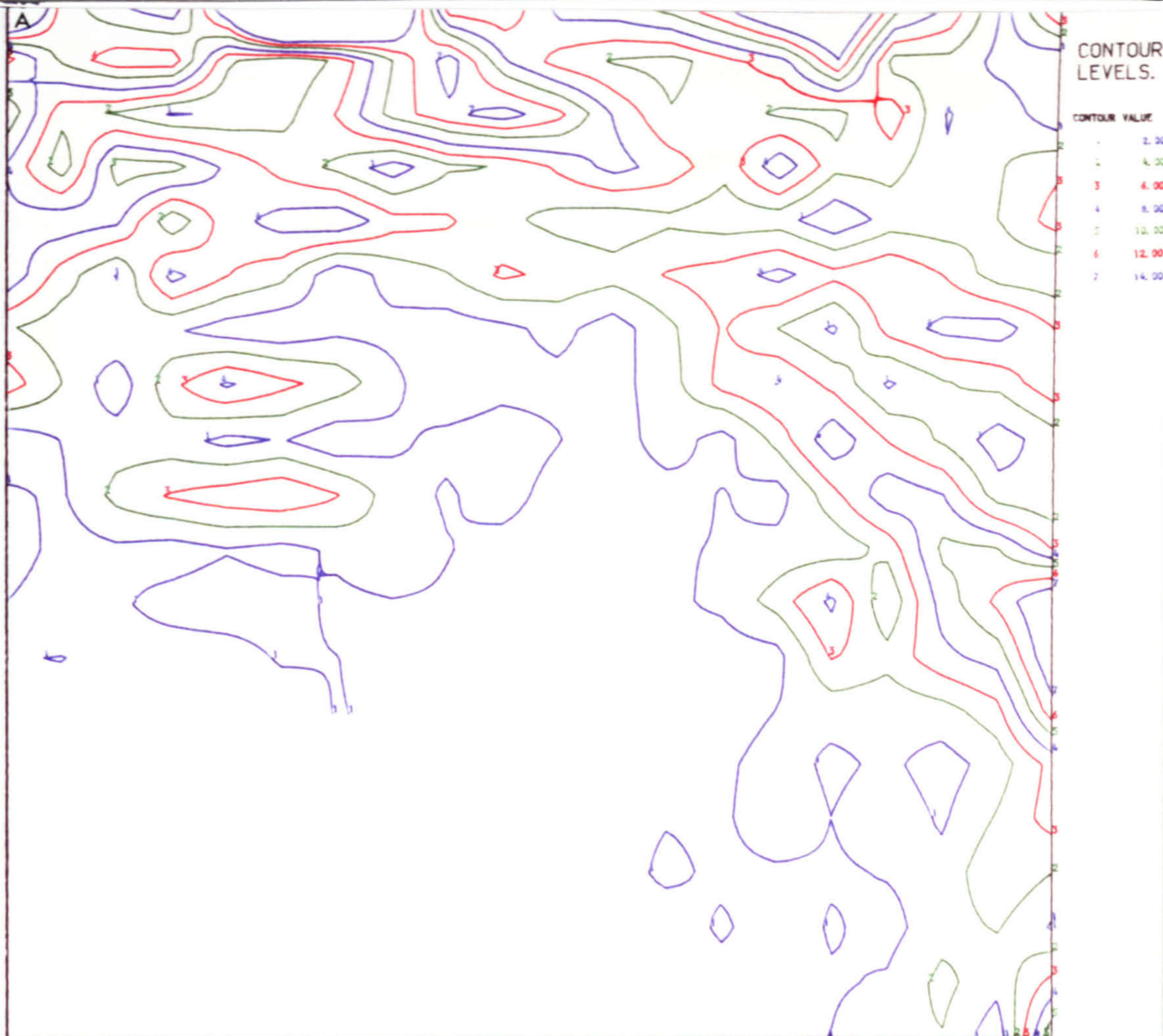


Figure 9.10 Local association map of standardised AP - POT with cut off at 17.5 metres. A) Association map. B) Relative probability map.

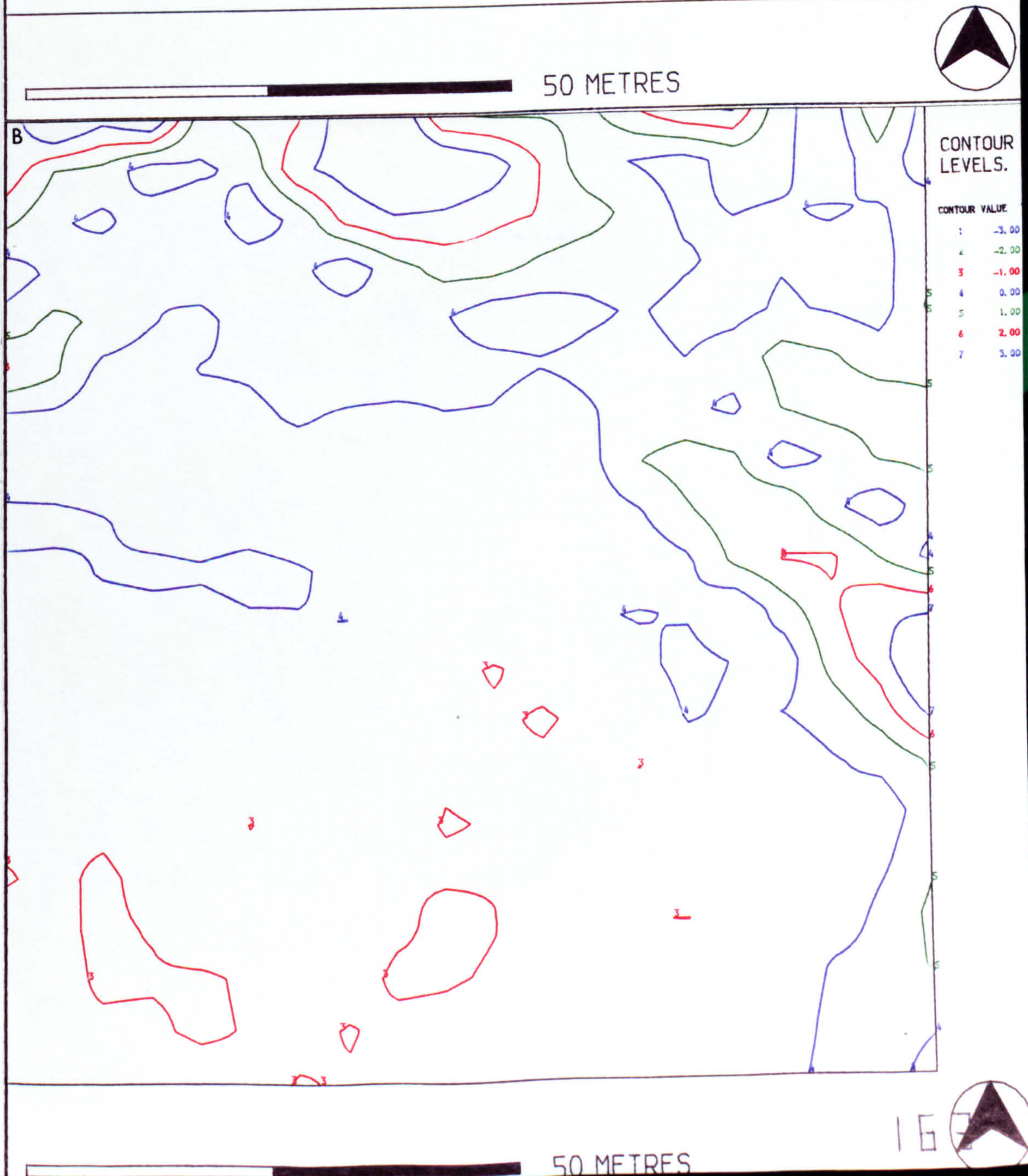
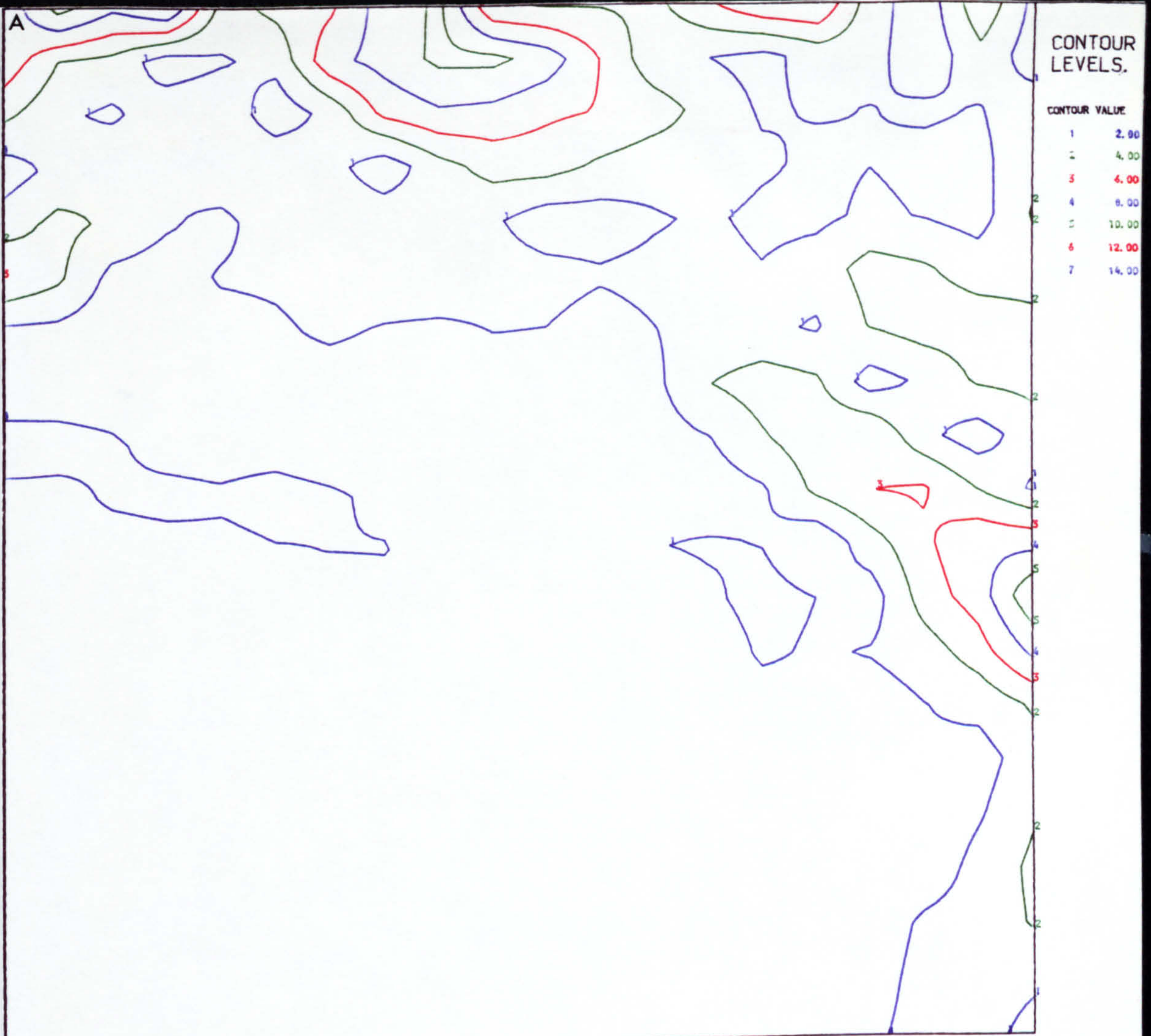


Figure 9.11 Local association map of standardised GEO - POT with cut off at 12.5 metres. A) Association map. B) Relative probability map.

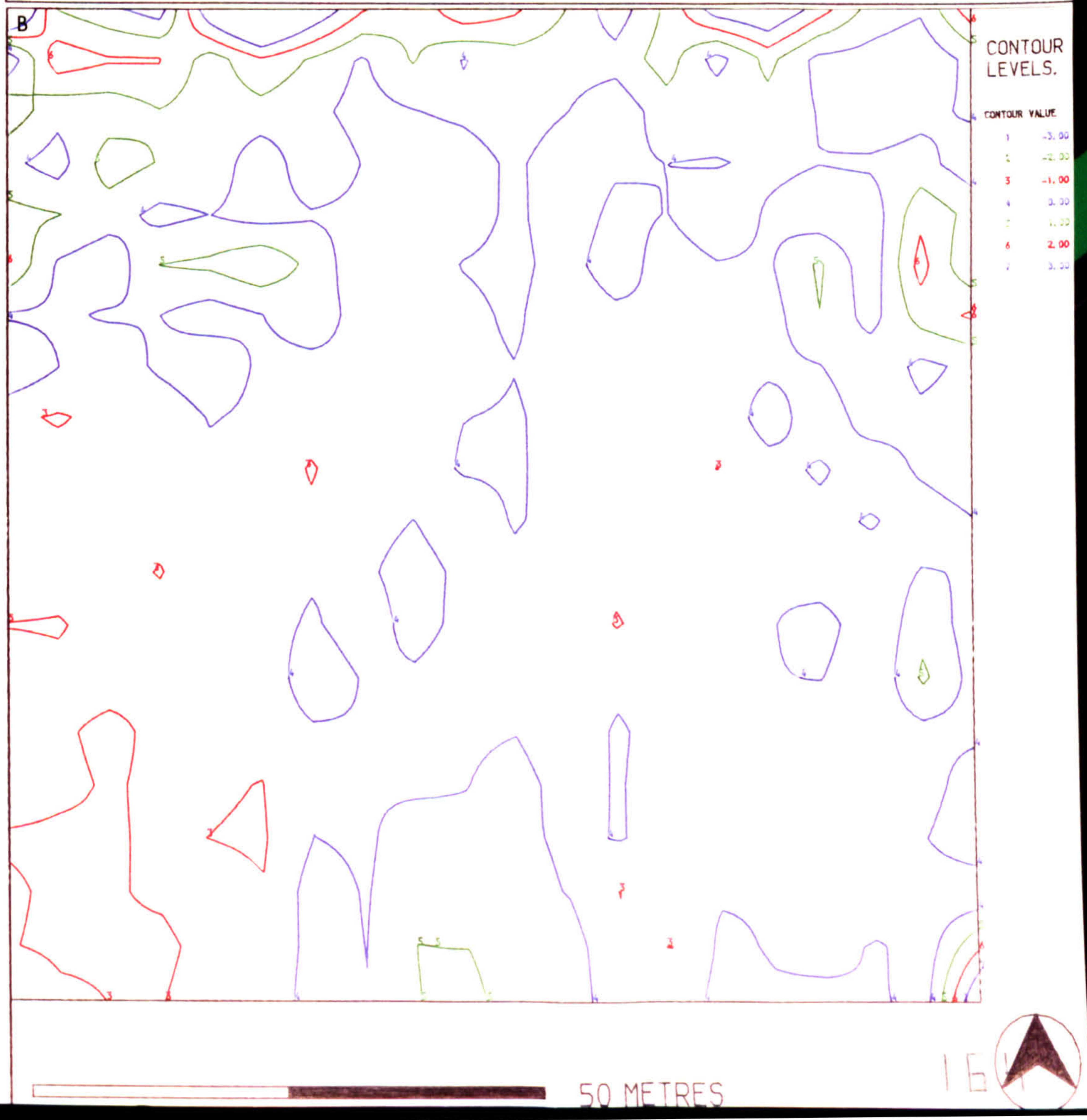
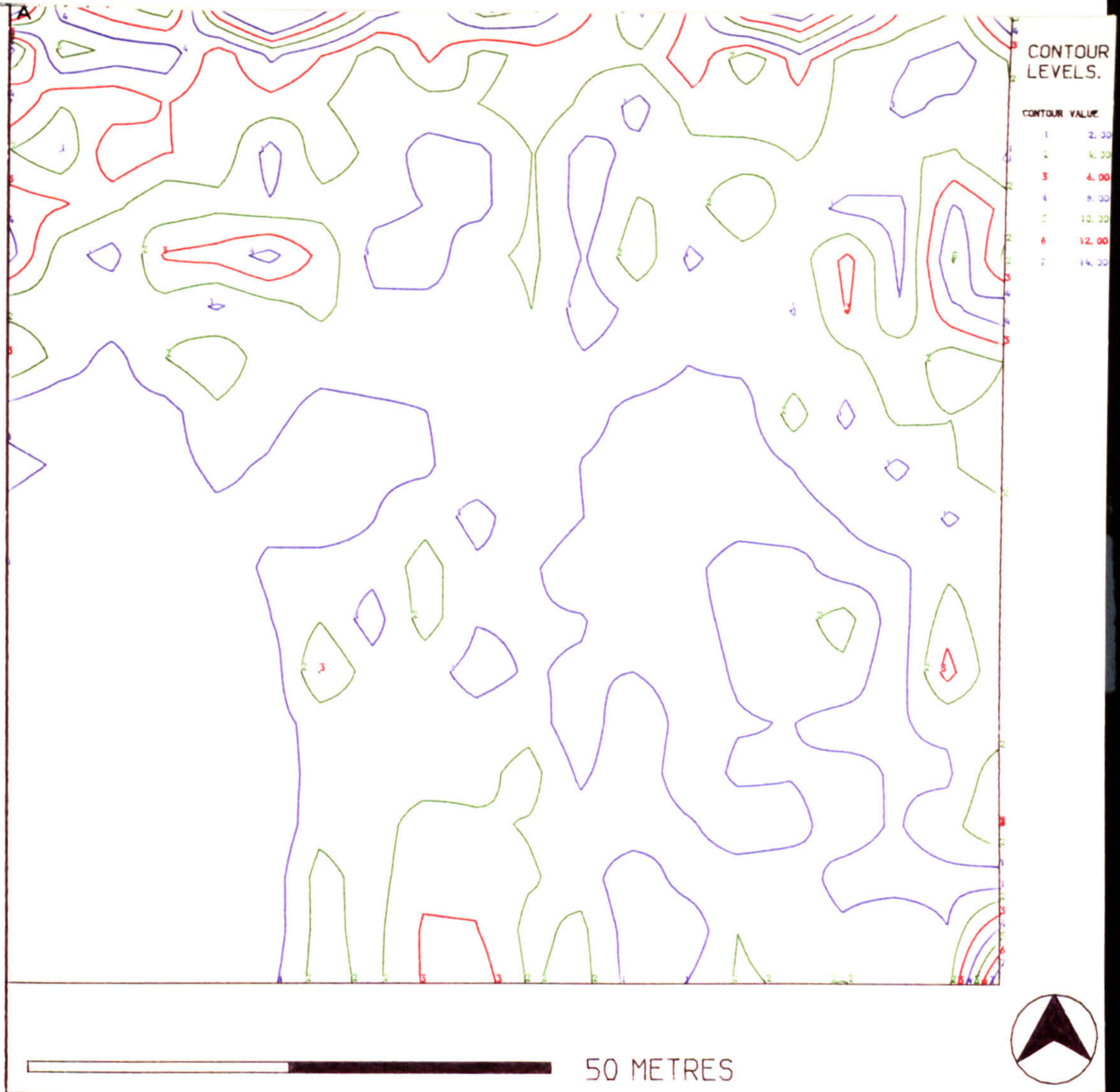
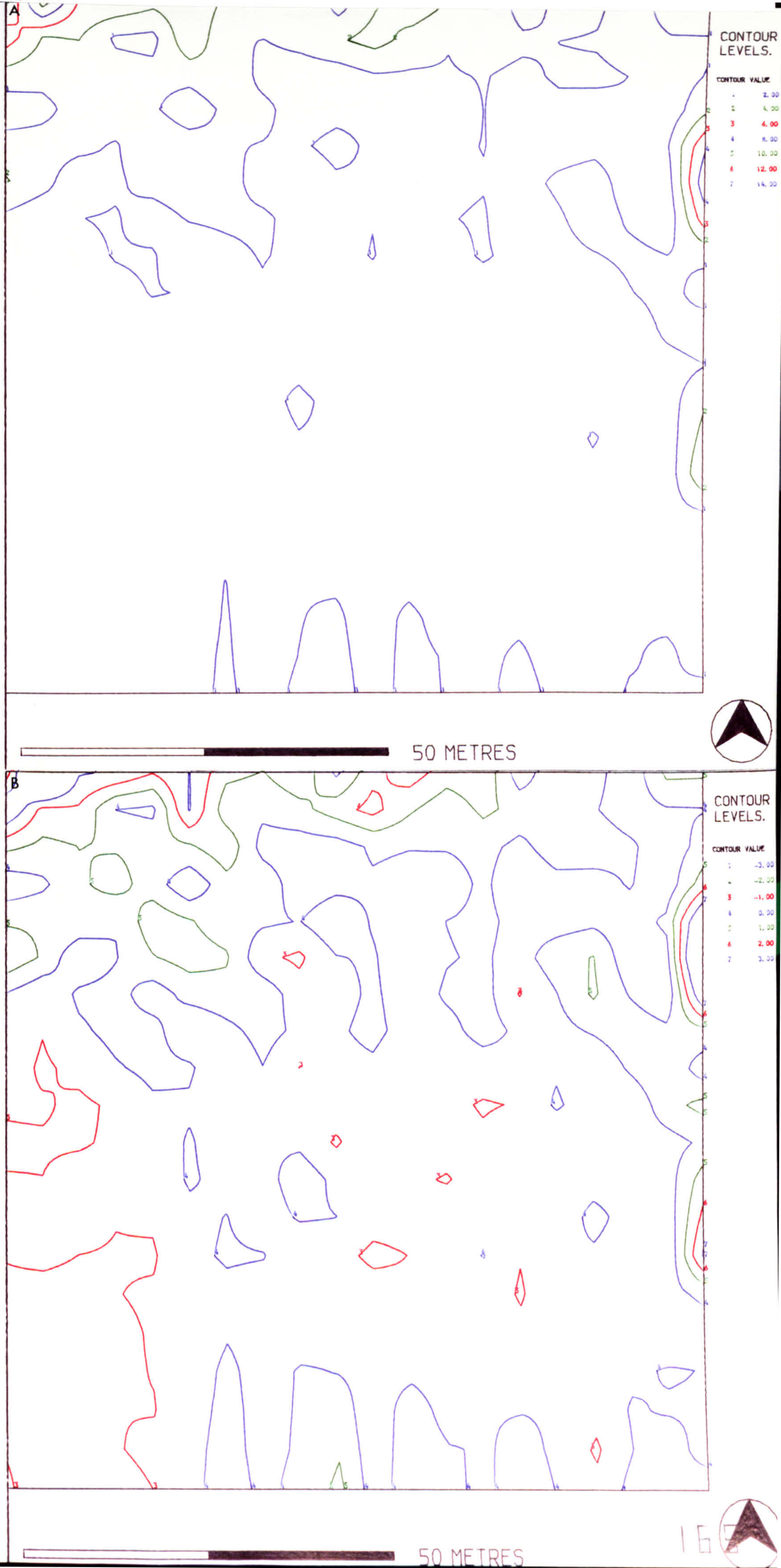


Figure 9.12 Local association map of standardised GEO - POT with cut off at 17.5 metres. A) Association map. B) Relative probability map.



conditions are in fact too marginal for useful data collection. As the earth resistance method used was the most appropriate of the commonly available methods of geophysical survey, it is recommended that no general geophysical methods are employed for further field survey. However the aerial photography and surface pottery distribution operands do show structural patterning and significant association and disassociation relationships at the local level. In particular some areas of high density pottery seem to define internal occupation areas of aerial photography features. It is therefore recommended that intensive surface collection is extended to further field survey, which in conjunction with existing aerial photography coverage may provide a valuable technique for mapping archaeological activity zones within the parish project area.

The technique of local association mapping as presented in section 9.3.2 is very useful but has several problems associated with it. The first is in the determination of relative significance by using Monte Carlo simulation techniques. This provides only an approximate measure of relative significance, an analytical measure of exact significance would be more accurate and a lot easier to apply. Such a measure would have to be based on the underlying spatial distributions of the operands. Although such an analytical measure is theoretically feasible, it would be complex, therefore the simulation approach provides an acceptable alternative for the present.

Secondly, as defined local association mapping can only be applied to two operands. Given $2n$ operands this will mean that a local analysis will produce n individual local association maps that need to be compared. However there are no restrictions on combining

a standardised local association map with a standardised operand or even another standardised map. Thus in theory local association mapping can be applied in a iterative manner to produce an association map for an arbitrary number of operands.

10.0 REGIONAL ASSOCIATION FOR LATTICE AND POINT PROCESSES

10.1 LATTICE PROCESSES

The previous two chapters have looked at spatial association in terms of site operands and operators. These operands can be regarded as the results of surface processes, i.e. processes that generate values of variables at all points in the plane. However, as Orton (1982b) has pointed out, lattice and point processes are also responsible for producing important classes of archaeological spatial data, for which measures of association are often desirable.

A lattice process generates values of variables at specified loci in the plane. These loci can be either points or regions. However most data at a regional scale, even if defined as areas, can be effectively represented as point loci. For example, measured attributes of archaeological sites, such as size, are best represented by a lattice process defined at point loci when the data are to be used for regional analysis.

Few techniques are available for the spatial analysis of lattice processes. The best known method is that of spatial autocorrelation (Cliff and Ord, 1973) which is a global analysis technique that tests for the existence of spatial autocorrelation (i.e. clustering of values of variables). Local analysis techniques, as described in section 9.3, are difficult to define for lattice processes because the values of variables other than at loci have no meaning. Most regional archaeological data is however often intuitively treated as the result of a surface process, even if actually produced by a

lattice process. For example the spatial distribution of percentages of pottery types at sites is often treated as a surface process when analysing marketing distribution systems (Fulford and Hodder, 1974). If a regional lattice process can be assumed to be defined at point loci, then a representation as a surface process can be achieved by regarding the loci as scattered sampling points of the surface. As described in section 4.2, contouring is the best method for illustrating data produced by surface processes, therefore methods of contouring scattered data need to be considered.

10.1.1 METHODS OF CONTOURING SCATTERED DATA

There are two approaches to the contouring of scattered data (Sabin, 1980), either contours are drawn directly from the data points or an intermediate grid of values is first calculated and then contoured as described in section 4.3. Directly drawn contours can be based on both global and local functions. Schagen (1982) uses a global interpolating function which is defined as a correlation between the current position in the contouring plane and all the data points. This correlation function takes the given value at each data point, but tends to an average value in regions away from the data, producing a series of undesirable peaks and troughs surrounding each point. Green and Sibson (1978) in contrast use local functions. They tessellate the plane as a set of Dirichlet triangles, with the data points at their vertices. Local linear functions can then be defined within each triangle, allowing the contours to be drawn. This method as implemented in the SIMPLEPLOT graphics package (Butland, 1984) has the advantage of

restricting contours to the region covered by the data, but produces inelegant lengthy straight-lined contours when the data are sparse.

Methods involving the definition of an intermediate grid, as for directly drawn contours, can be based on both global and local functions. However, the disadvantages of using a global function to fit all the data and define the grid values have already been emphasised in section 4.3 and are discussed in Hietala and Larson (1979). It is difficult, if not impossible, to define explicit global functions which are sufficiently complex to give a good fit to the data and which have reasonable behaviour at large distances. Global functions are therefore of little practical use. Local functions may be fitted to the data points in the neighbourhood of a grid point to determine its value. As discussed in section 9.3.2 this fit is usually weighted in terms of the distance from the grid point to the data point, but it may also involve allowance for accuracy, when the errors on the data are known. The GHOST graphics package (Noland, 1979) provides calculation of grid points by means of weighted averages, effectively giving a linear fit to neighbouring data. McLain (1974) uses a quadratic function to obtain grid values. This allows the fitted function to reach its extremes at points between the data rather than at themselves, as is the case with the linear fit. It is a matter for consideration whether or not the fitted local functions should match the data values. When the data are known to be prone to significant error, it is better to allow an overall fit, without matching each data value exactly. Either option can be achieved by a suitable choice of weighting function. If exact matching is required, then a data point should be given infinite weighting at a coincident grid point;

when some degree of smoothing is required, infinite weightings should be avoided.

The main disadvantage in the use of an intermediate grid is that grid values may be calculated and contoured for regions not logically associated with the data. A typical example is the contouring of data such as riverbed temperatures measured at random points in a river. The constructed intermediate grid may well generate contours that cross the riverbanks. Although untidy these spurious contours are not detrimental and can be eliminated either by using a mask or by marking a boundary outline.

Of the contouring methods discussed the direct methods are most problematic. They are potentially powerful but are complex both in theory and implementation. The intermediate grid methods although less natural are more satisfactory in terms of efficient implementation and tractability. They also offer considerable advantages in the development of local association measures as outlined in the next section. For these reasons the contouring of scattered data using an intermediate grid calculated by locally fitted functions is recommended.

As an illustration of this approach a least-squares fitting of a local quadratic function with an inverse-square weighting, as defined in section 9.3.2, was used on a typical regional lattice process. The lattice process chosen was the size of Bronze age barrow cemeteries as recorded in the North Yorkshire Sites and Monuments Record Database (NYSMRDB) (see section 7.3). The attribute of size was measured by counting the number of barrows

within a cemetery, the criteria for a cemetery being:

'at least three barrows occur in close proximity to each other as extant structures, or are recorded as having existed as such', (Griffiths, 1982).

From the NYSMRDB seventy seven possible Bronze age cemeteries were identified and their sizes measured. As the physical sizes of the cemeteries are very small compared to the regional analysis area, they can effectively be regarded as point loci. The resultant scattered data contour plot is shown in figure 10.1. The intermediate grid size was based on cells of 5 km X 5 km. The red boundary indicates the present extent of the NYSMRDB within the county, therefore contours outside this boundary should be ignored.

10.1.2 ASSOCIATION OF LATTICE PROCESSES

The generalisation of the spatial autocorrelation method to the multivariate case to allow the measurement of global association is theoretically possible but, as Orton (1982b) comments, likely to be very complicated. Local association measures for lattice processes are however more feasible. By assuming for regional analysis that lattice processes are defined at point loci and treating them as surface processes, contouring methods for scattered data, as described in the previous section, can be used. With the intermediate grid calculation contouring method, data structures equivalent to the site operands of 8.0 are created. This means that the local association methods of section 9.3 (i.e. similarity maps and local association maps) can be applied.

10.2 P
attempt

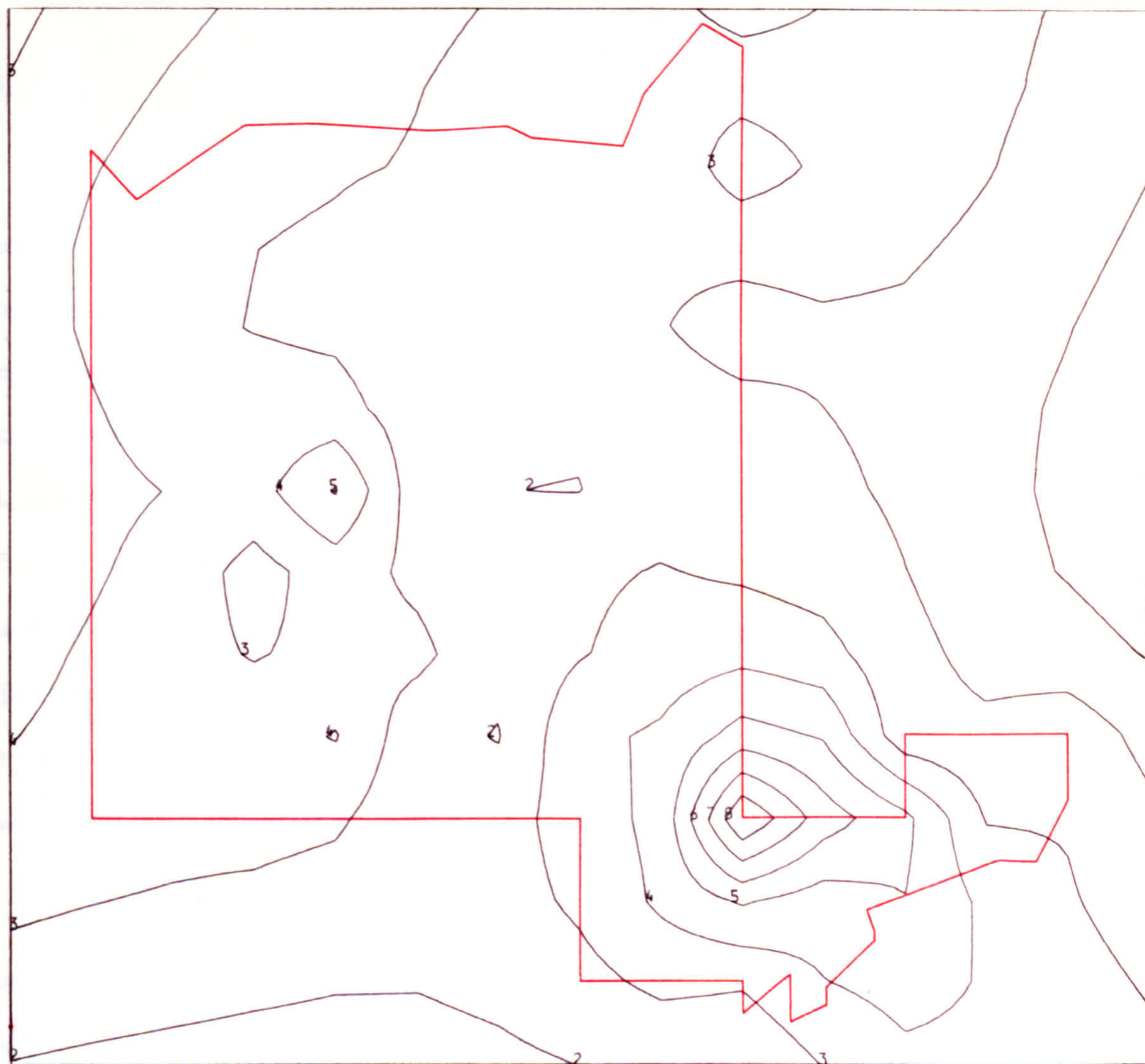


Figure 10.1 Bronze age barrow cemeteries. Bottom left hand corner at 435 km, 455 km absolute. Scale 1 cm = 5 km. Contours 1-2, 2-4, 3-6, 4-8, 5-10, 6-12, 7-14, 8-16.

In practise the results of these local measures of association must be interpreted carefully as the lattice processes are being treated as surface processes. It is therefore probably best to look for general association trends within a regional analysis and not attempt to use apparently significant details.

10.2 POINT PROCESSES

The spatial analysis of point processes has received considerable attention within quantitative archaeology. A point process generates events at points in the plane. These events, such as artifacts within a site or sites within a region, form distribution maps which are a traditional interpretative tool in archaeology. Much effort has been expended in developing quantitative analysis techniques (Hodder and Orton, 1976) to try and overcome the problems of subjective interpretation. The main developments have been in distance methods, i.e. methods that use event-event distances or point-event distances. The principal technique used is that of nearest neighbour analysis.

However there has been growing dissatisfaction with these approaches to the spatial analysis of point processes (e.g. Kintigh and Ammerman, 1982) on the basis that they do not answer relevant archaeological questions. This dissatisfaction is strengthened by the inadequate extension of these methods to association analysis. No really effective local or global association methods for archaeological point processes exist at the present time.

10.3 CLASSICAL K'TH NEAREST NEIGHBOUR ANALYSIS

Classical K'th nearest neighbour analysis was first proposed by Clark and Evans (1954) for first order neighbours. Thompson (1956) extended the analysis to K neighbours and Dacey (1963) extended it to n dimensions. For two dimensions and K neighbours the classical analysis can be described in the following manner. Given an infinite Euclidean plane in which there is a random distribution of point particles with density λ ; the probability density function for the random variable x, the distance from a random point to its K'th nearest neighbour particle, is:

$$f(x; \lambda, k) = \frac{2(\lambda\pi)^k}{(k-1)!} x^{2k-1} e^{-\lambda\pi x^2} I_{(0, \infty)}(x),$$

where $\lambda > 0$ and $k=1, 2, 3, \dots$

The mean and variance of this probability density function are:

$$E[x] = \frac{C(k)}{\sqrt{\lambda}},$$

$$\text{Var}[x] = \frac{1}{\lambda} \left[\frac{k}{\pi} - C(k)^2 \right],$$

$$\text{where } C(k) = \frac{1.3.5 \dots (2k-1)}{2^k (k-1)!}.$$

In the normal use of the analysis a random sample of n particles is taken. The distance from each particle in the sample to its K'th nearest neighbour particle (x_k) is then measured and the mean K'th

nearest neighbour distance ($\overline{x_k}$) calculated:

$$\overline{x_k} = \frac{\sum x_k}{n} .$$

The expected mean K'th nearest neighbour distance for a sample from a random spatial distribution of particles with the same density is used to test the significance of the sample statistic:

$$z = \frac{\overline{x_k} - E[x]}{\sqrt{\frac{\text{Var}[x]}{n}}} ,$$

where z is the standard Normal deviate.

For significantly negative z , the particle distribution is judged to be significantly clustered in comparison to a random distribution. For significantly positive z , the particle distribution is judged to be significantly systematic in comparison to a random distribution. For non-significant z , the particle distribution is judged not to be significantly different from a random distribution (see figure 10.2).

It should be noted that as the distances are measured from random particle to particle rather than from random point to particle, the correct density is calculated as:

$$\frac{n-1}{\text{area}} ,$$

instead of:

$$\frac{n}{\text{area}} .$$

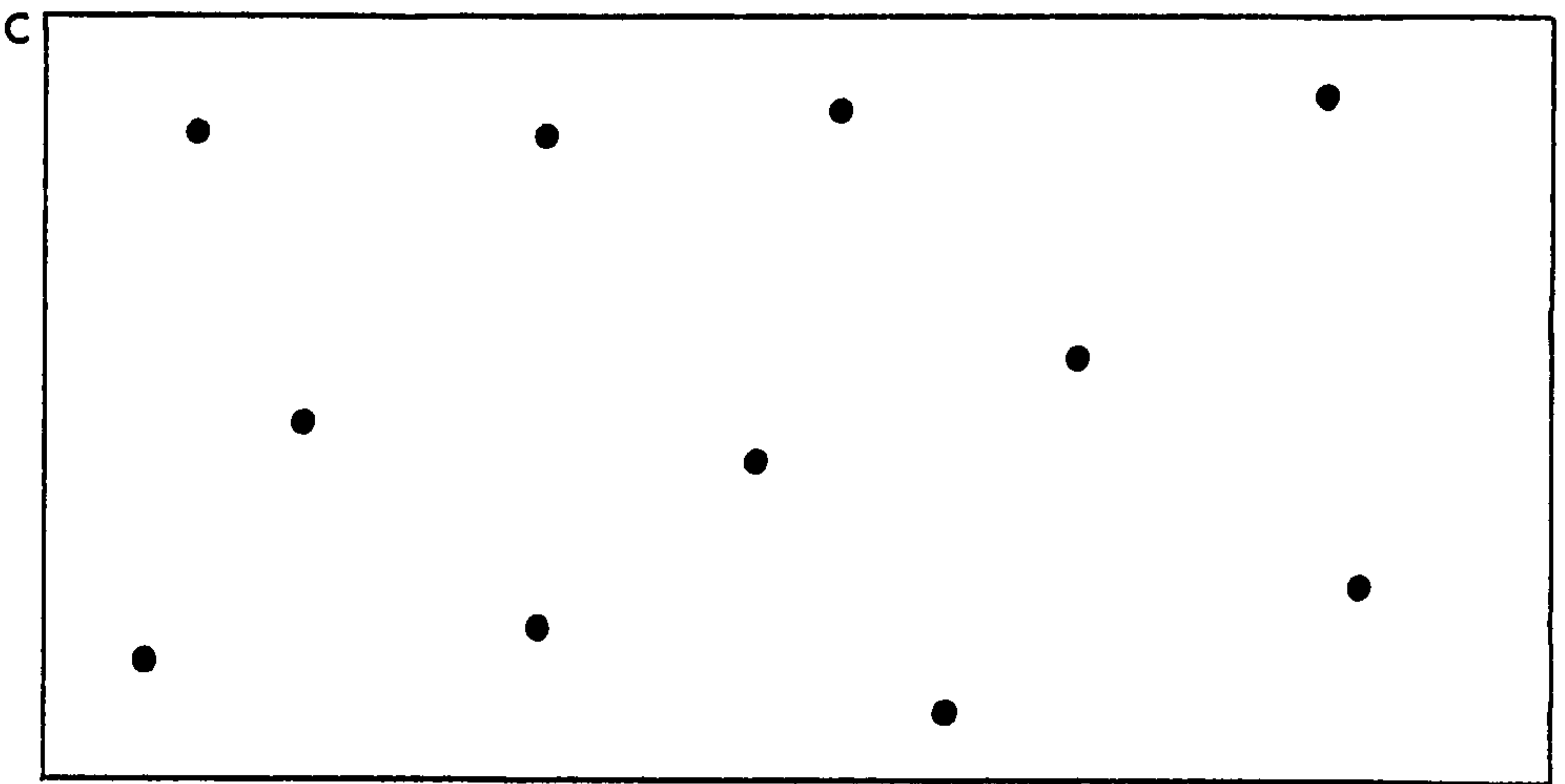
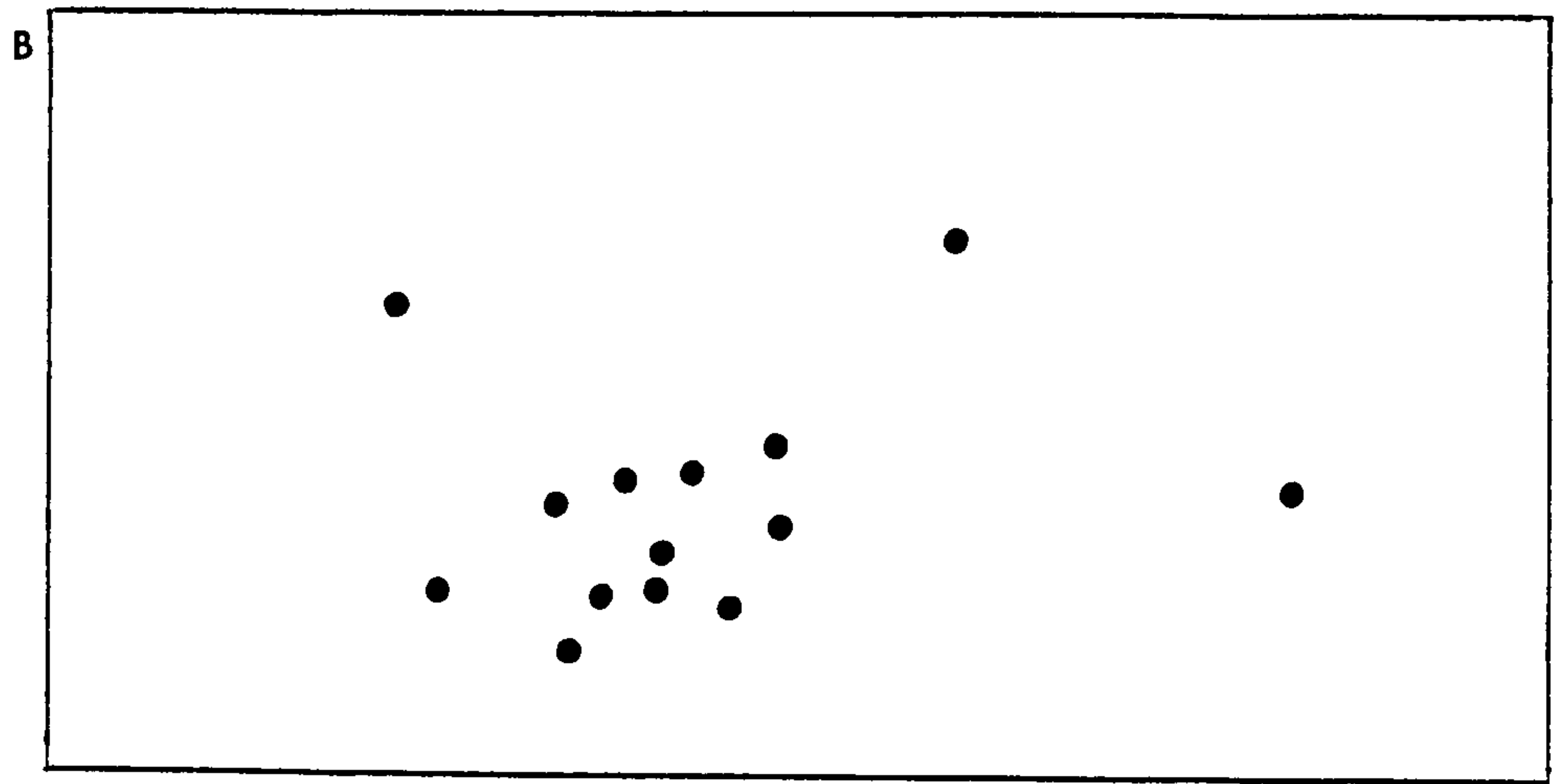
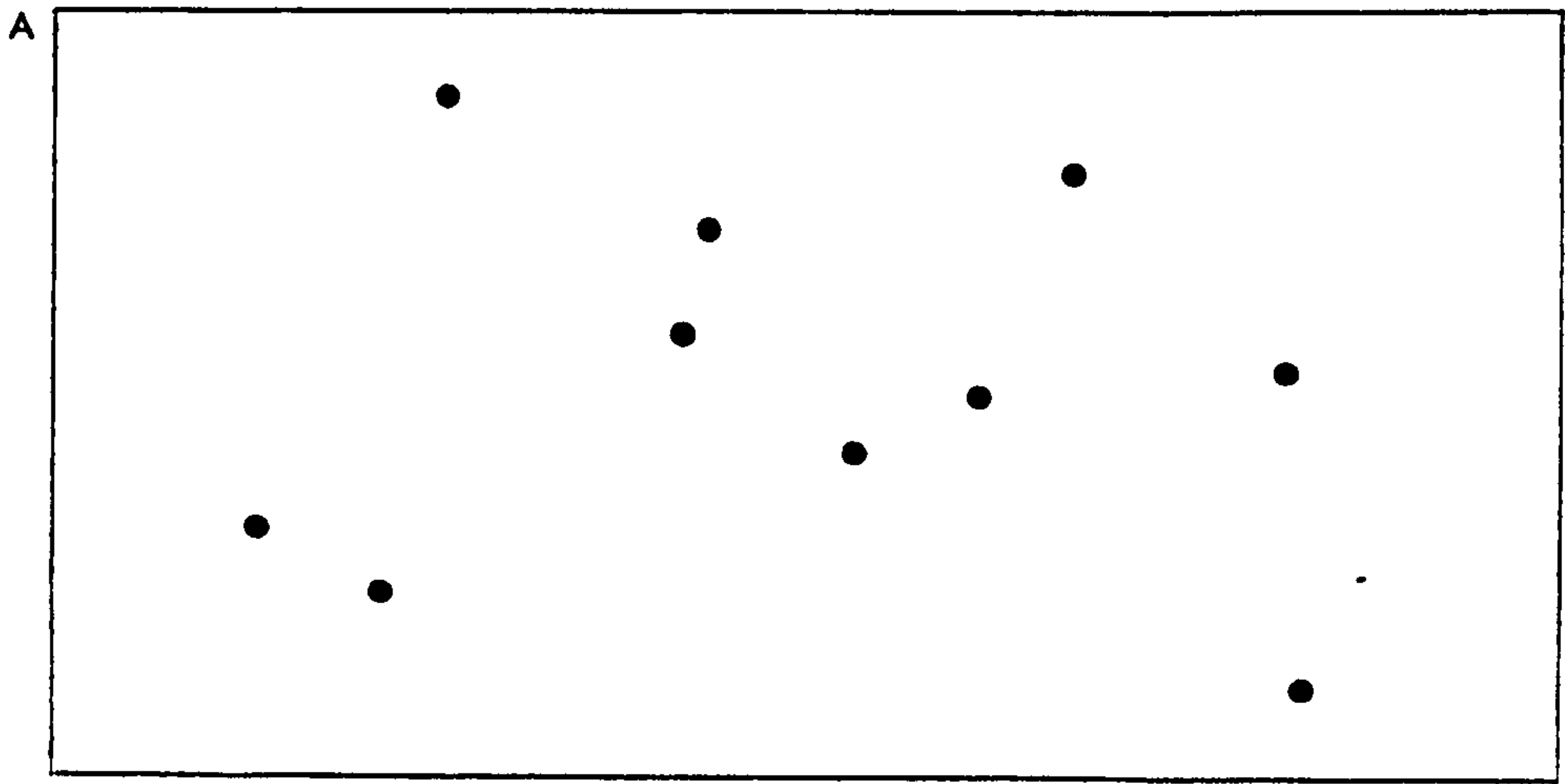


Figure 10.2 Point distributions. A) Random. B) Clustered. C) Regular.

For large n there is little difference, but small values of n will give significant variations. It is therefore recommended that the exact density calculation is always used.

10.4 PROBLEMS WITH THE CLASSICAL ANALYSIS

A number of problems exist in the application of classical K' th nearest neighbour analysis to archaeology. The technique was originally developed by Clark and Evans (1954) to measure spatial distribution patterns in plant ecology. In a typical application a population of plants would be analysed by defining a sampling area well within the spatial extent of the population. A random sample of plants within this area would then be drawn to allow the K' th nearest neighbour distances to be measured. As De Vos (1973) has pointed out, archaeological distributions are often too sparse and too valuable to be sampled in this manner. In most archaeological applications we wish to analyse a population rather than a sample. This means that the classical variance measure is no longer valid as we now have dependent distance measurements. Moore (1954) has suggested generating a random distribution of points from which to measure K' th nearest neighbour distances to particles, this will give a random sample of distances which will allow the classical variance to be used. Although this is a feasible solution it still does not guarantee a full utilisation of the collected data even for very high sampling densities.

In the previous plant ecology example the situation could arise where a sample plant's K' th nearest neighbour lies outside the sampling area. This presents no problems as the distance can still

be measured. However most archaeological distributions are defined within a rigid boundary, such as an excavation area. Points lying close to the boundary cannot see potential K'th nearest neighbours beyond it. Thus the mean K'th nearest neighbour distance for a random distribution within a rigid boundary will be larger than predicted by the classical analysis. This is the problem of edge effects.

Various solutions to overcome edge effects have been proposed. Hodder and Orton (1976) have suggested constructing a buffer zone within the analysis area which excludes points close to the boundary from the analysis. This has the obvious disadvantage of reducing the population size being analysed. For large K the buffer zone will be so large that the population will be effectively negated. Pinder, Shimada and Gregory (1979) have proposed an empirical correction for the first order nearest neighbour statistic for a square area based on the simulation data of Ebdon (1976). Their corrected formulae are:

$$E[x] = C\sqrt{A/n} ,$$

$$\sqrt{\frac{\text{Var}[x]}{n}} = V.E[x] ,$$

where A = area ,

$$C = 0.497 + 0.127\sqrt{A/n} ,$$

$$V = 0.0098 + 0.4701 (C\sqrt{A/n}) .$$

The disadvantage of this correction is that it does not work. The corrected formulae are only applicable to the simulation conditions used, a square of area four units, and when applied under different conditions, as shown later, give completely spurious results.

McNutt (1981) has suggested a semi-analytical correction for first order nearest neighbours based on the observation that edge effects effectively reduce the density. He has put forward a series of geometrical density corrections for a limited range of shapes. For example for a square of side S the density correction is:

$$\lambda = \frac{(n-1-n_0)}{A},$$

$$\text{where } n_0 = \frac{2S\sqrt{n-1}}{3\sqrt{A}} - \frac{1}{12}.$$

The classical analysis can then be used with this corrected density. Although McNutt has used a range of shapes, his correction is only valid for first order nearest neighbours, a limitation which has important archaeological significance as discussed later. Donnelly (1978) has taken the simulation approach of Ebdon (1976) to produce empirical corrections for the first total nearest neighbour distance statistic. He has used however a large variety of shapes, an acceptable range of densities and very large numbers of simulation experiments to produce corrections for edge effects and non-independence of measurements:

$$E[T] = 0.5\sqrt{(nA)} + \left[0.051368 + \frac{0.041}{\sqrt{n}} \right] L,$$

$$\text{Var}[T] = 0.0703 A + 0.037 L\sqrt{(A/n)},$$

where T = total nearest neighbour distance,

L = boundary length.

However Donnelly's corrections, as for the previous ones, are limited to first order neighbours.

A careful examination of the classical analysis will show that the statistical significance of any observed mean K' th nearest neighbour distance is dependent on density. For plant ecology this presents no problems as the sampling area is chosen to lie well within the plant population being analysed. However the definition of an area in archaeological applications is often problematic. For example in an excavation it may seem obvious to use the area defined by the excavation limits. But if the excavation only contains spatial distributions localised to a small occupation area this will produce a clustered result. Although this is a valid answer most archaeologists would be more interested in the distribution patterns within the actual occupation area. Solutions to this area definition problem are not obvious. A statistical approach would involve defining the area before any analysis or even any data collection, but this may produce archaeologically irrelevant results. An intuitive approach would involve fitting an area to the data to formulate specific archaeological questions, but this may produce subjective results. The best solution is probably to use a compromise, clearly state the area definition and the criteria for its choice. Analysis results can then be openly judged in this context.

The final problem in the application of the classical analysis to archaeology is in the limitation of its use. The classical analysis has been applied to a wide range of archaeological problems; e.g. Hodder and Hassall (1971), Zubrow (1971), Adams and Nissen (1972), Hodder (1972), Hammond (1974), Plog (1974), Washburn (1974), Whallon (1974), Earle (1976), and Clark, Effland and Johnstone (1977). In the vast majority of applications the analysis has been limited to

first order nearest neighbours. As several authors have pointed out (Greig-Smith, 1964; Hodder, 1972; Hodder and Hassall, 1971; Hodder and Orton, 1976; King, 1969; Pielou, 1961, 1969) limiting the K'th nearest neighbour analysis to first order neighbours, gives only a very limited measure of spatial patterning. Of much greater use is an analysis over a range of K's, however for increasing K the problems of non-independence of measurements and edge effects substantially increase. Because no effective corrections for higher order neighbours exist it has not been possible objectively to use higher order neighbours in archaeological applications.

10.5 A NEW CORRECTION FOR K'TH ORDER NEIGHBOURS

To produce a corrected K'th nearest neighbour analysis there are two basic approaches. The first involves the use of Monte Carlo simulation techniques for each individual analysis. Given an analysis area and a point density, a large number of simulations of random spatial distributions are carried out. These are then used to construct sampling distributions for the mean K'th nearest neighbour distances. The significance of the observed mean K'th nearest neighbour distances can then be determined. The second approach is to produce corrected formulae for the expected mean K'th nearest neighbour distances and their sampling variances for random distributions, using analytical methods.

Both approaches have their advantages and disadvantages. The simulation technique is applicable to all possible analyses. The degree of accuracy of an analysis can also be controlled by the number of simulations used. Simulation techniques are however very

much a brute force approach, they give no real insight into the actual analysis and often involve considerable processing, usually on a computer. In contrast analytical methods produce results that are very simple and quick to apply to analyses. They can also give considerable insight into how an analysis performs. But analytical corrections are very difficult to formulate for every conceivable analysis, as problems such as edge effects are dependent on the geometry of the analysis area.

To produce a corrected K'th nearest neighbour analysis it was decided to use a semi-analytical approach. By formulating the general analytical corrections (after McNutt (1981)), their actual values could then be estimated by simulation (after Donnelly (1978)). This would hopefully produce corrected formulae that could be applied to most analyses.

10.5.1 SIMULATION EXPERIMENTS

For the simulation experiments a standard area of 100 units was taken. Six shapes with this area were then devised (figure 10.3), which ranged from simple shapes such as the circle to complex shapes such as the Z and T which were expected to present severe boundary effects. For each shape random distributions of points were simulated using a proprietary random number generator (NAG, 1982). The number of points used for each shape were:

5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80,
85, 90, 95, 100.

For each shape and number of points; 500 simulations were carried

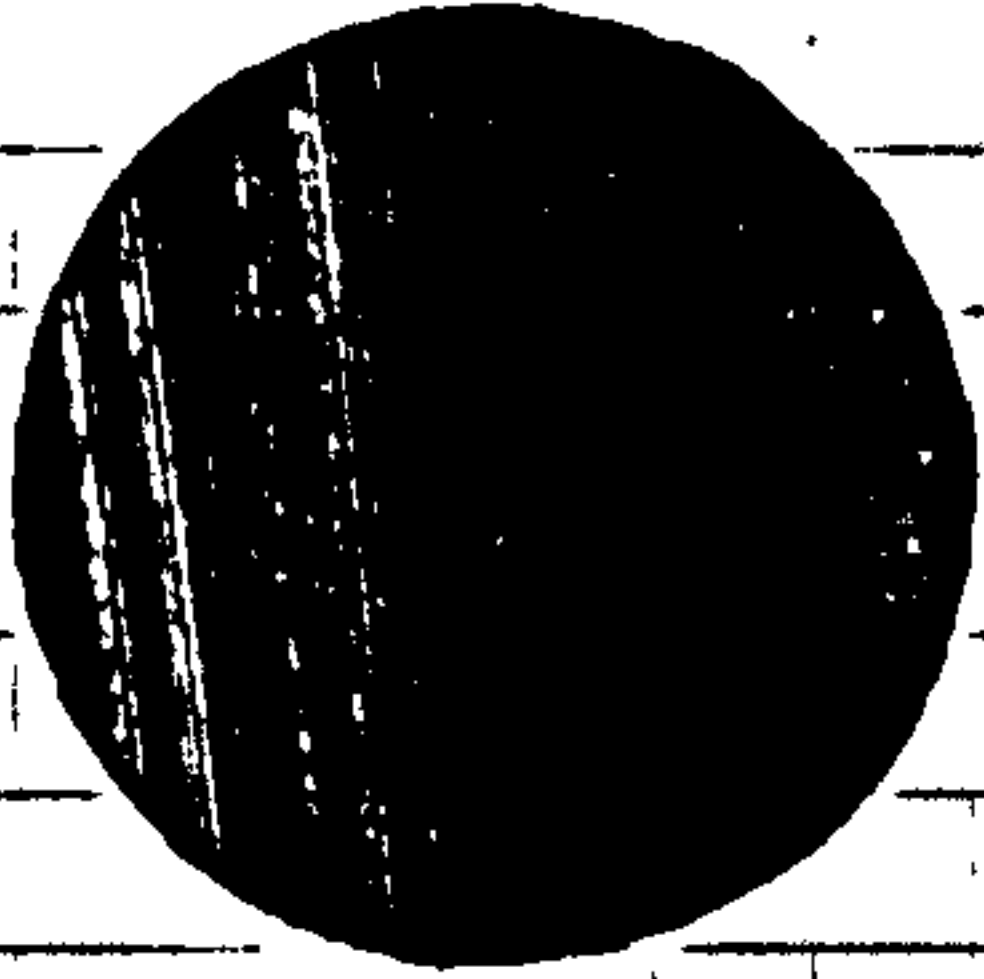
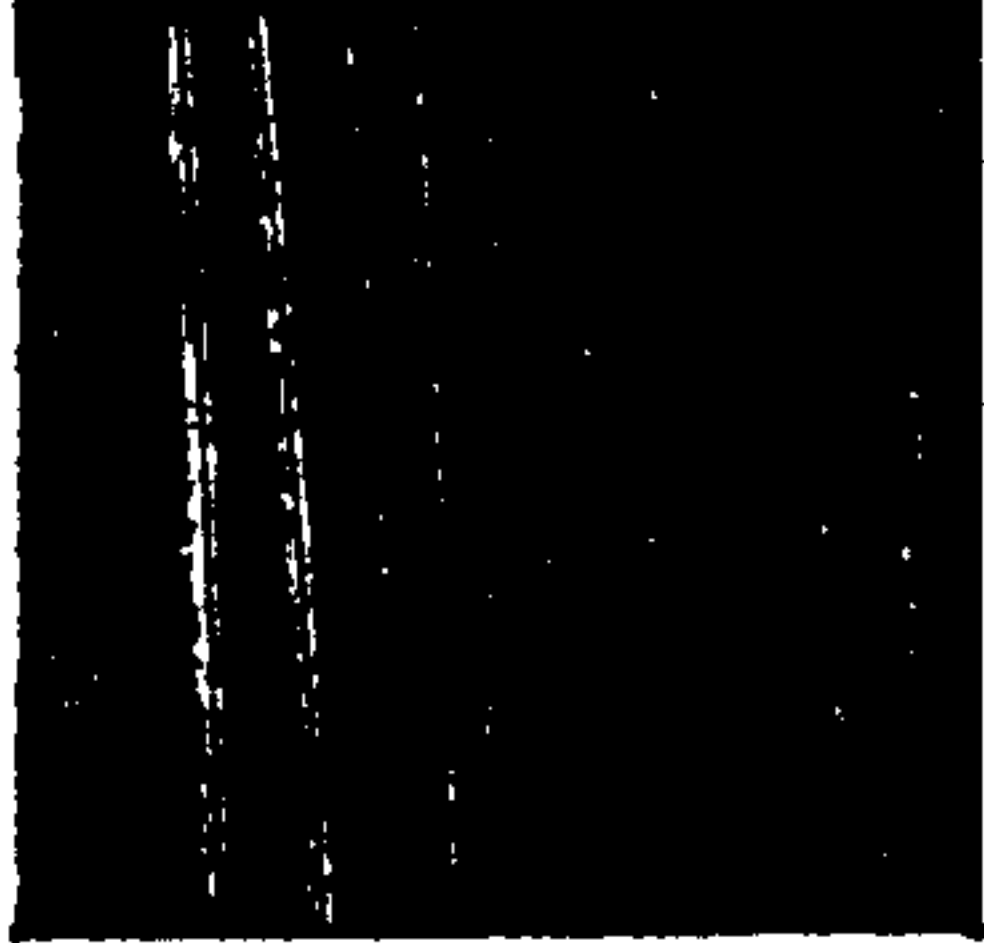

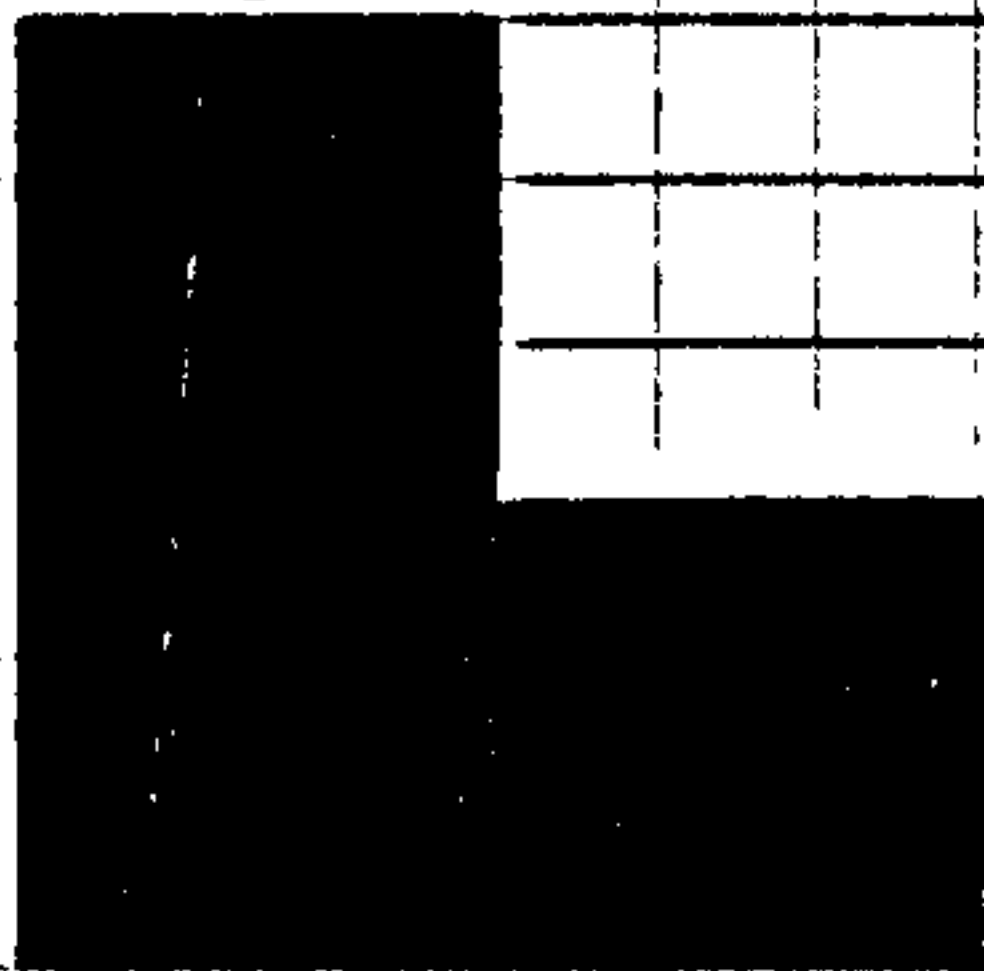
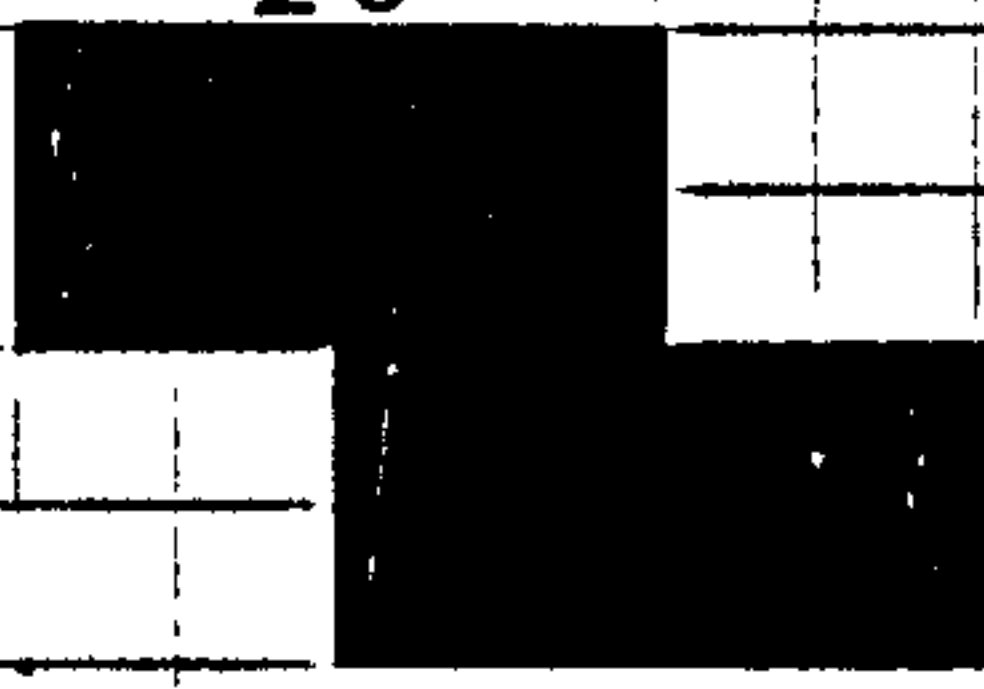
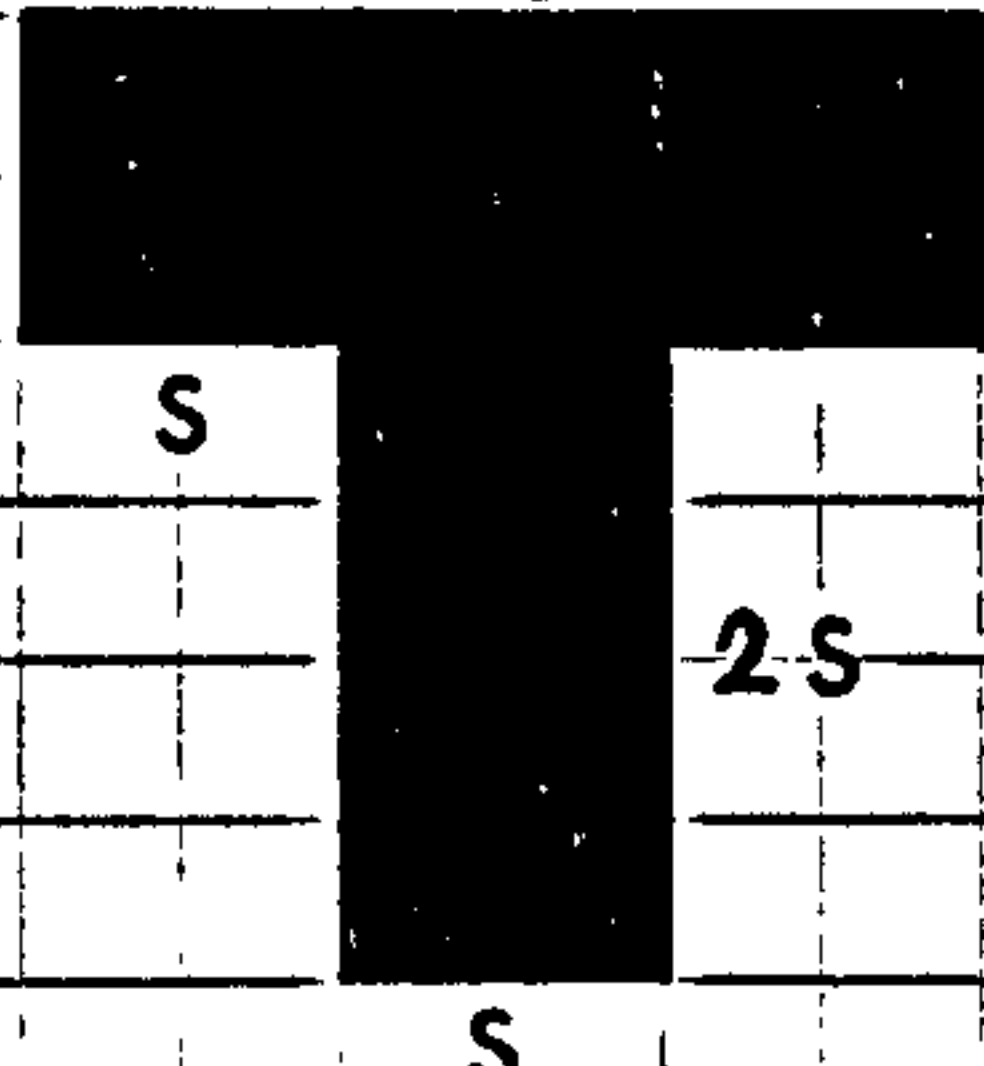
SHAPE	FORM	AREA (A)	BOUNDARY (L)
Circle		$A = \pi S^2$	$L = 2\pi S$
Square		$A = S^2$	$L = 4S$
Rectangle		$A = 2S^2$	$L = 6S$
L		$A = 3S^2$	$L = 8S$
Z		$A = 4S^2$	$L = 10S$
T		$A = 5S^2$	$L = 12S$

Figure 10.3 Simulation shapes.

out, giving a total of 60,000 simulation experiments in all. For each such simulation the mean K'th nearest neighbour distances were measured up to the twentieth neighbour where possible. The 500 simulations for each shape and number of points were used to produce sampling distributions of the mean K'th nearest neighbour distances. The mean and variance of these sampling distributions were calculated and recorded.

10.5.2 MEAN K'TH NEAREST NEIGHBOUR DISTANCE CORRECTION

As McNutt (1981) has pointed out, edge effects essentially reduce the point density towards the edge of the analysis area. McNutt corrected for this density drop by reducing the number of points in the density calculation. However it is probably more sensible to see the density drop in terms of an increased area. This increased or apparent area can be defined as:

$$A' = A + Q,$$

where A' = apparent area,

A = actual area,

Q = area correction.

The area correction will depend on the estimate of the mean K'th nearest neighbour distance ($E[x]$) as calculated by the classical analysis. In fact Q will be directly proportional to $E[x]$ because as $E[x]$ gets larger the edge effects will increase, so that Q also needs to be increased to correct for this. Q is also directly proportional to the length of the boundary (L) around the analysis area. If A remains the same and L is made larger the edge effects will increase, so that Q also needs to be increased to correct for

this. Therefore:

$$Q \propto L E[x],$$

or:

$$A' = A + b L E[x],$$

where b is a constant between 0 and 1.

Using this apparent area the corrected mean K' 'th nearest neighbour distance ($E[x]'$) can be defined as:

$$E[x]' = C(k) \sqrt{\frac{A'}{n-1}},$$

or:

$$E[x]' = C(k) \sqrt{\frac{A + b L E[x]}{n-1}}.$$

Donnelly (1978) in his simulation experiments calculated a correction constant similar to b . He plotted his correction constant against the reciprocal of the square root of n , and found a significantly straight line. Using the same approach the above formula for the corrected mean K' 'th nearest neighbour distance can be expanded using a binomial expansion. Let:

$$Y = \frac{b L E[x]}{A},$$

therefore:

$$E[x]' = C(k) \sqrt{\frac{A}{n-1}} (1 + Y)^{\frac{1}{2}}.$$

A binomial expansion to three terms gives:

$$E[x]' \doteq C(k) \sqrt{\frac{A}{n-1}} + \frac{L}{n-1} \left[\frac{C(k)^2 b}{2} - \frac{C(k)^3 b^2}{2\sqrt{n-1}} \right].$$

Let:

$$CL = \frac{C(k)^2 b}{2} - \frac{C(k)^3 b^2}{2\sqrt{n-1}},$$

therefore:

$$E[x]' \doteq C(k) \sqrt{\frac{A}{n-1}} + \frac{L}{n-1} CL,$$

therefore:

$$CL \doteq \frac{n-1}{L} \left[E[x]' - C(k) \sqrt{\frac{A}{n-1}} \right].$$

Using this last formula, where $E[x]'$ was taken from the simulation experiments, CL was plotted against the reciprocal of the square root of $n-1$ for K equals 1 to 20. All the graphs showed significantly straight lines except for small values of n where the CL values for the more complex shapes, such as Z and T, started to diverge. Ignoring this divergence for the moment, the gradients of the graphs were all effectively zero. This is in strong contrast with Donnelly's graph which has a significant gradient. This suggests that Donnelly's correction may very well have been correcting for the effect of his having used n rather than $n-1$ in his density evaluation, and have nothing to do with the actual correction for edge effects.

By rearranging the formula for the corrected mean K 'th nearest neighbour distance the following can be obtained:

$$b = \left[\frac{(n-1) E[x]'}{C(k)^2} - A \right] \frac{1}{L E[x]} .$$

Using this formula, where $E[x]'$ was taken from the simulation experiments, b was plotted against $n-1$ for K equals 1 to 20. The resulting graphs again displayed significantly straight lines except for a divergence at low values of n for the complex shapes. This complexity effect related to the shape of the analysis area needs next to be considered.

Given an analysis area, the distance from its centroid to the closest point on the boundary can be thought of as defining the radius of an enclosed circle (figure 10.4). The diameter of this circle (D) can be used to define a complexity limit beyond which any analysis could be judged unreliable. Ignoring edge effects let:

$$E[x] < \alpha D,$$

where α is some constant,

therefore:

$$C(k) \frac{\sqrt{A}}{\sqrt{n-1}} < \alpha D,$$

therefore:

$$n-1 > \frac{A C(k)^2}{\alpha^2 D^2} .$$

The value of α is arbitrary, but a value of 0.5 is a reasonable

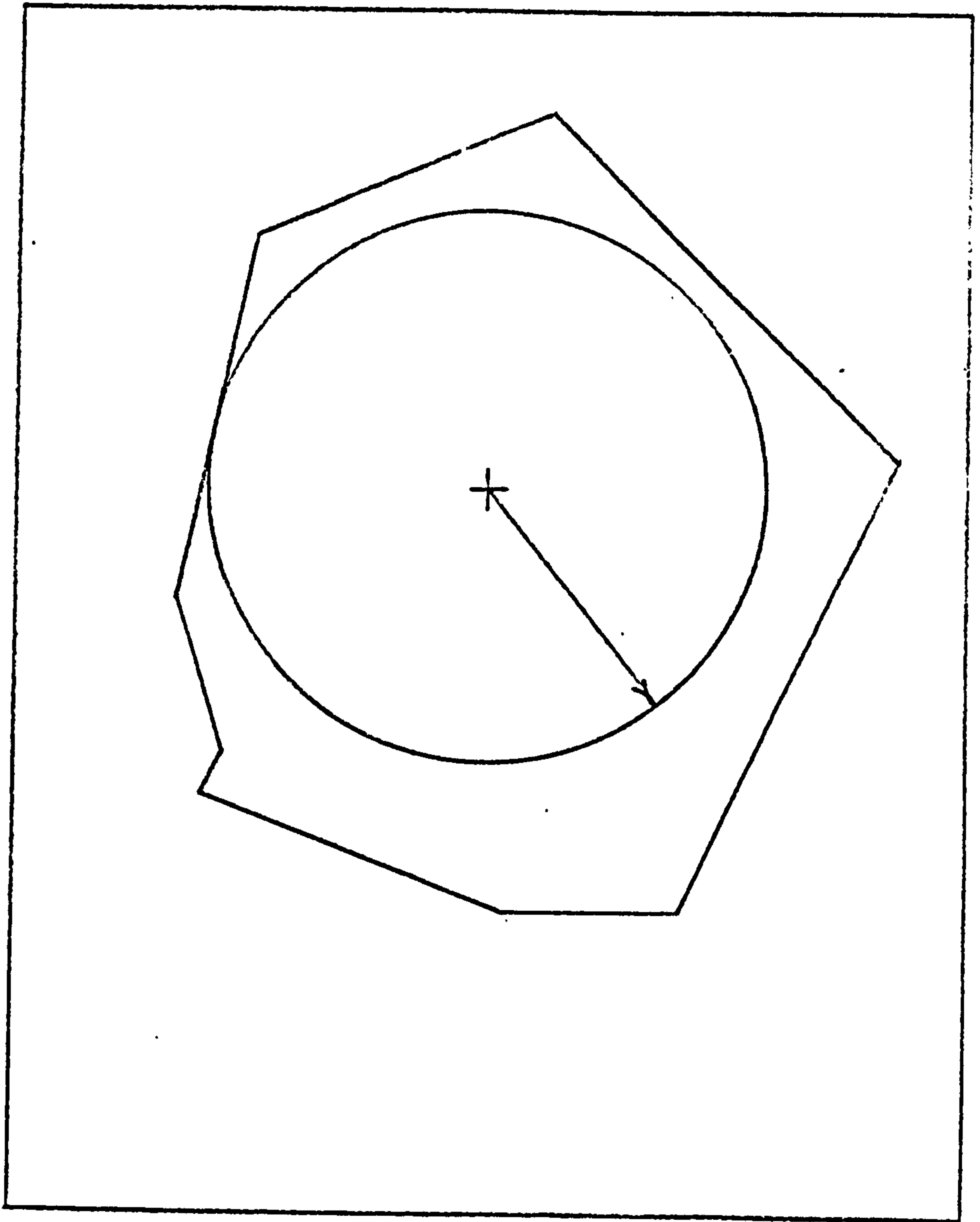


Figure 10.4 Complexity effect.

choice. This gives a complexity limit of:

$$n-1 > \frac{4 A C(k)^2}{D^2} .$$

Using this complexity limit, all the over complex data were removed from the graphs of b against n-1. The new graphs showed significantly linear trends with zero gradients and no divergences. The intercept values of these graphs are shown in table 10.1. A function of the form:

$$b = \alpha \ln k e^{\beta k} + \gamma ,$$

was fitted to this data to give:

$$b = 0.3934 - 0.0425896 \cdot \ln k \cdot e^{-0.180336 k} .$$

This function and the fitted data are shown in figure 10.5.

10.5.3 SAMPLING VARIANCE CORRECTION

To produce a corrected sampling variance ($\text{Var}[\bar{x}]'$) for the mean K'th nearest neighbour distance, the edge effects correction of the previous section can be applied to the classical variance formula:

$$\text{Var}[x] = \frac{A'}{n-1} \left[\frac{k - C(k)^2}{\pi} \right] ,$$

or:

$$\text{Var}[x] = \frac{A + b L E[x]}{n-1} \left[\frac{k - C(k)^2}{\pi} \right] .$$

This corrected formula now needs to be divided by n to give the

K	b
1	0.3934
2	0.3677
3	0.3653
4	0.3652
5	0.3657
6	0.3683
7	0.3705
8	0.3731
9	0.3755
10	0.3776
11	0.3794
12	0.3832
13	0.3830
14	0.3844
15	0.3849
16	0.3842
17	0.3850
18	0.3900
19	0.3917

Table 10.1 b coefficients calculated from simulations.

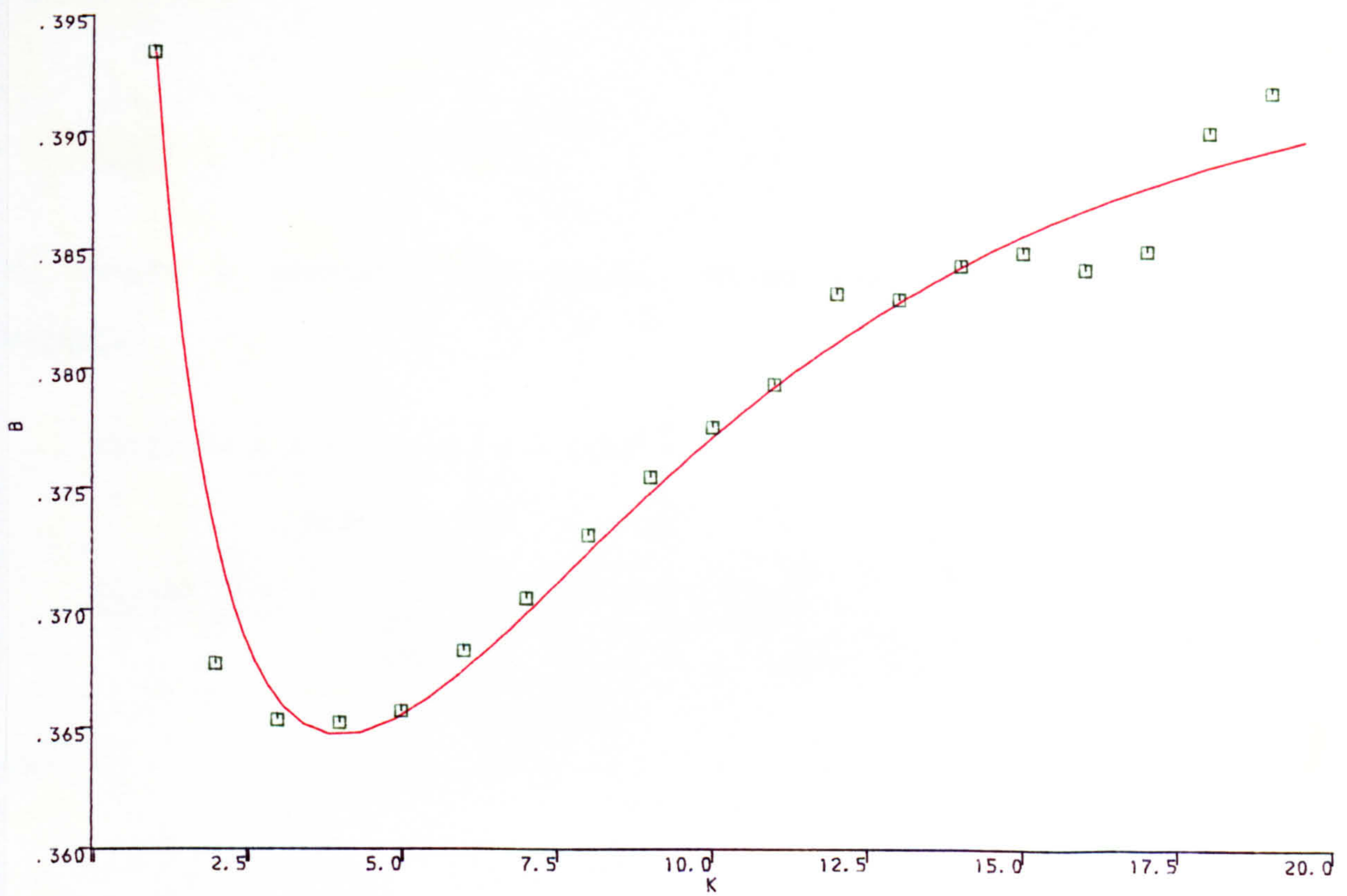


Figure 10.5 b values and fitted function.

variance of the sampling distribution of the mean K'th nearest neighbour distance:

$$\text{Var}[\bar{x}]' = \frac{A + b L E[x]}{n(n-1)} \left[\frac{k}{\bar{\pi}} - C(k)^2 \right].$$

Donnelly (1978) in his simulation experiments plotted:

$$\frac{\text{Var}[T]}{A},$$

against:

$$\frac{L}{\sqrt{nA}},$$

to obtain a straight line graph. Given the sampling variance formula:

$$\text{Var}[\bar{x}]' = \frac{A + b L E[x]}{n(n-1)} \left[\frac{k}{\bar{\pi}} - C(k)^2 \right],$$

$$\frac{\text{Var}[\bar{x}]' n(n-1)}{A} = \left[\frac{k}{\bar{\pi}} - C(k)^2 \right] \left[\frac{1 + b L E[x]}{A} \right],$$

let:

$$p = \left[\frac{k}{\bar{\pi}} - C(k)^2 \right],$$

therefore:

$$\frac{\text{Var}[\bar{x}]' n(n-1)}{A} = p + \frac{p b L E[x]}{A},$$

$$\frac{\text{Var}[\bar{x}]' n(n-1)}{A} = p + \frac{p b L C(k)}{\sqrt{A} \sqrt{n-1}},$$

let:

$$q = p C(k) b,$$

therefore:

$$\frac{\text{Var}[\bar{x}]'n(n-1)}{A} = p + \frac{q L}{\sqrt{A} \sqrt{n-1}} .$$

After removing all the over complex data, as outlined in section 10.5.2, graphs of:

$$\frac{\text{Var}[\bar{x}]'n(n-1)}{A} ,$$

against:

$$\frac{L}{\sqrt{A} \sqrt{n-1}} ,$$

were plotted using the simulated data for K equals 1 to 20. All the graphs showed significant linear trends and had equivalent intercepts at 0.0703. Linear functions were fitted to the graphs using linear regression. The gradients of these graphs are shown in table 10.2. A function of the form:

$$g = \alpha K^\beta ,$$

where g = gradient,

was fitted to this data to give:

$$g = 0.03059 k^{1.357} .$$

This function and the fitted data are shown in figure 10.6. Using the intercept and gradient the expected sampling variance can be

K	g
1	0.033
2	0.080
3	0.136
4	0.182
5	0.260
6	0.350
7	0.430
8	0.520
9	0.570
10	0.670
11	0.780
12	0.890
13	0.990
14	1.090
15	1.200
16	1.320
17	1.480
18	1.570
19	1.680
20	1.950

Table 10.2 g gradients calculated from simulations.

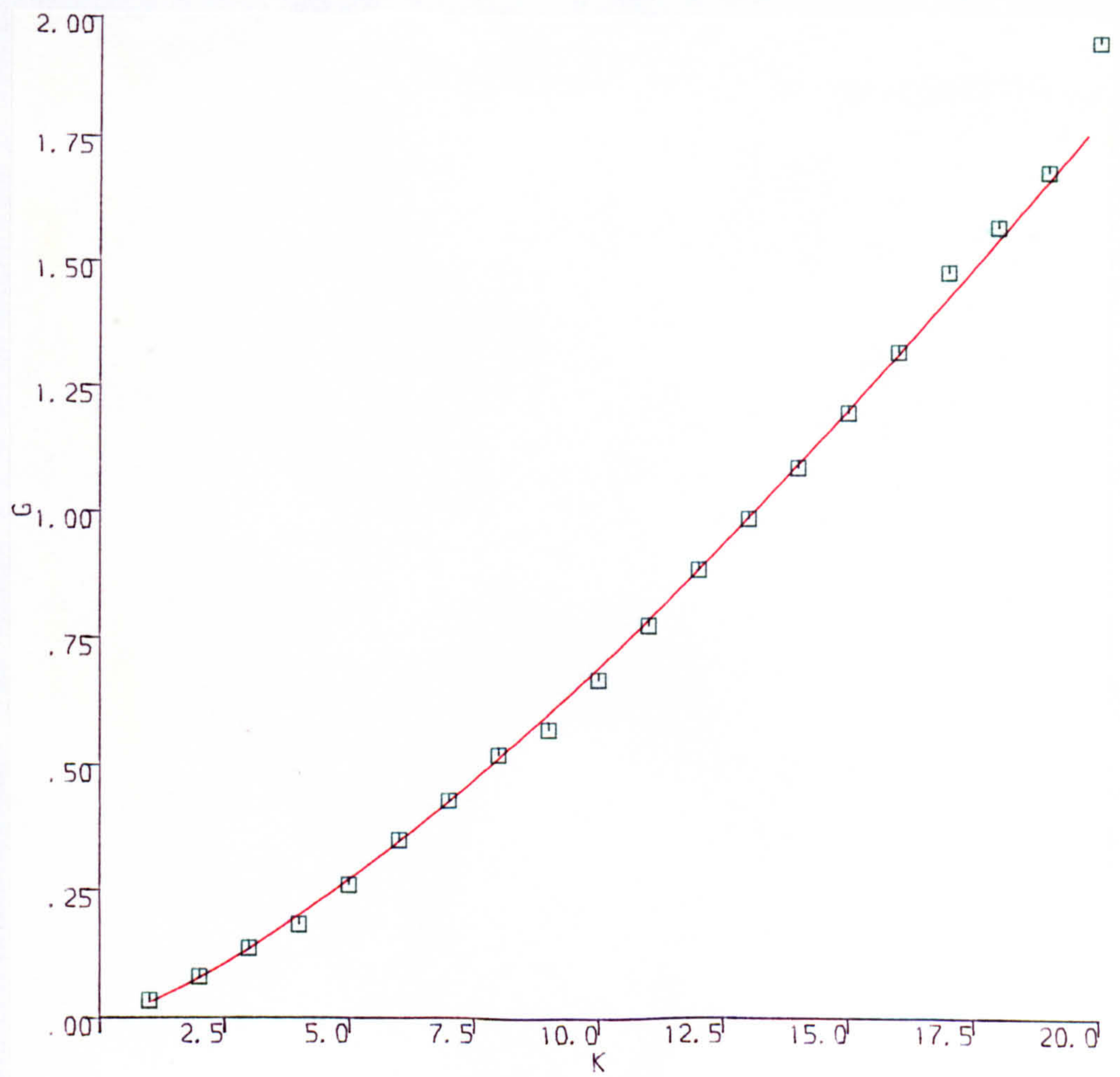


Figure 10.6 g values and fitted function.

calculated as:

$$\frac{\text{Var}[\bar{x}]'n(n-1)}{A} = 0.0703 + \frac{g L}{\sqrt{A} \sqrt{n-1}},$$

therefore:

$$\text{Var}[\bar{x}]' = \frac{A}{n(n-1)} \left[0.0703 + \frac{g L}{\sqrt{A} \sqrt{n-1}} \right].$$

10.5.4 COMPARISON

To compare the relative effectiveness of the new correction for K'th order neighbours, simulated random distributions of 10, 20, 30, 40 and 50 points were generated within a square of 100 units area. For the first nearest neighbour the actual mean nearest neighbour distance and the predicted using the classical analysis, the previous corrections presented in section 10.4 and the new correction are given in table 10.3. As expected the classical analysis gives an under estimate of the actual values. The new correction and those of McNutt (1981) and Donnelly (1978) give close approximations to the simulated data, with the new correction probably giving the best approximation. In contrast the correction of Pinder, Shimada and Gregory (1979) gives large over estimates for all the simulated data.

Table 10.4 shows the variance of the mean nearest neighbour distance for the simulated data and for the classical analysis and corrections. Again the classical analysis gives an under estimate of the actual values. The new correction and those of Donnelly (1978) and McNutt (1981) give closer approximations to the simulated

N	10	20	30	40	50
(1)	1.8905	1.2254	1.0041	0.8496	0.7536
(2)	1.6667	1.1471	0.9285	0.8006	0.7143
(3)	2.8417	1.7463	1.3307	1.1033	0.9569
(4)	1.8787	1.2431	0.9903	0.8461	0.7502
(5)	1.8385	1.2390	0.9913	0.8484	0.7528
(6)	1.8725	1.2463	0.9940	0.8496	0.7534

Table 10.3 Mean nearest neighbour distance in a square of area 100 for:

- (1) Simulated data,
- (2) Classical analysis,
- (3) Pinder, Shimada and Gregory (1979),
- (4) McNutt (1981),
- (5) Donnelly (1978),
- (6) New correction.

N	10	20	30	40	50
(1)	0.1514	0.0234	0.0095	0.0055	0.0040
(2)	0.0759	0.0180	0.0079	0.0044	0.0029
(3)	14.6232	2.1046	0.7148	0.3399	0.1934
(4)	0.0964	0.0211	0.0089	0.0049	0.0031
(5)	0.1171	0.0258	0.0108	0.0059	0.0036
(6)	0.1234	0.0259	0.0105	0.0058	0.0036

Table 10.4 Variance of mean nearest neighbour distances in a square of area 100 for:

- (1) Simulated data,
- (2) Classical analysis,
- (3) Pinder, Shimada and Gregory (1979),
- (4) McNutt (1981),
- (5) Donnelly (1978),
- (6) New correction.

data, with again the new correction giving the best overall approximation. The correction of Pinder, Shimada and Gregory (1979) gives ridiculously large estimates. This is not surprising as their correction has been formulated for a specific size and shape of analysis area as indicated in section 10.4.

Overall the new correction seems to give good results when compared to simulated data. It also seems to behave marginally better than previous corrections, however these corrections are only applicable to first order neighbours while the new correction can be applied to K'th order neighbours.

10.6 ASSOCIATION OF POINT PROCESSES

The problems involved in the association of point processes are far from simple. This is reflected in the development of measures of spatial association for point data. Pielou (1961) developed a coefficient of segregation (S) based on a contingency table analysis of the number of times a type A point has a type B point as its K'th nearest neighbour. The disadvantage of this global method is that it gives only a simple measure of segregation or non-segregation and does not use the actual distances between points. Clark, Effland and Johnstone (1977) use a related method where they define a radial cut off, construct circles around type A points and count the number of type B points in these circles. Using this information they construct an index of similarity (I) between spatial distributions (a similar method is used by Johnson (1977) in local density analysis). As for the previous method the actual distances between points are not used. The index has also not been related to any

formal statistical tests. Hodder and Okell (1978) calculate an index of association (A) based on the mean A-A, B-B and A-B distances. To test the significance of this index they use simulation techniques. This has the obvious disadvantage of requiring considerable computing time. Graham (1980) uses A-A and A-B distances to construct a test statistic in a similar manner to the index of association. In common with all the other global measures of association described, this only gives a very limited description of the relationship between spatial distributions.

Given a distribution of type A points and a distribution of type B points, there are four main relationships to be considered in any form of association analysis. The spatial structure of A-A and B-B distances can be determined for K'th order neighbours using the corrected analysis. These two relationships characterise the A and B distributions independently of each other. For the A distribution it is important to compare A-A distances to B-A distances to determine spatial association. Similarly for the B distribution the B-B distances need to be compared to the A-B distances. Thus the four main relationships are A-A, B-B, B-A and A-B distances for K'th order neighbours. The B-A and A-B distances can be analysed using the corrected K'th nearest neighbour analysis. Using these four relationships a descriptive assessment of spatial association can be produced. Archaeologically this is more useful than a simple measure of association or non-association and is more powerful than simple subjective description as it is based on a series of objective analyses.

To implement this association method a program called k_nna was written in Fortran 77. An outline of the structure of this program

is given in appendix A.7. To facilitate the easy analysis of a wide range of point data sets a format for a standard point data file (SPDF) was devised. This is described in appendix D. The program `k_nna` reads a standard point data file. It then allows you to nominate the type A and type B distributions. With these distributions it calculates corrected k' 'th nearest neighbour analyses for A-A, B-B, A-B and B-A distances. The results are then presented as graphs with the expected mean distances drawn in black, the 5% and 1% significance levels drawn as dashed red lines, and the observed mean distances drawn in green. These graphs of the four main relationships form the basis for the association analysis.

As an example of the use of this approach to the spatial association of point processes, the work of Radley (1974) can be considered. Radley was interested in the relationship between prehistoric settlement and soil type in the Vale of York. He plotted distribution maps of polished axes representing the Neolithic period and bronze implements and axe hammers representing the Bronze age. He then equated the extent of these distributions with occupation areas and decided that the distribution of these occupation areas was not random but tended to cluster on the well drained soils.

To test these assumptions the Bronze age distributions of bronze implements and axe hammers can be considered first (figure 10.7). The bronze implement distribution shows significant spatial clustering. The axe hammer distribution has a lower density and shows just significant clustering. The axe hammer to bronze implement distances show no significant association below the sixth nearest neighbour. The bronze implement to axe hammer distances

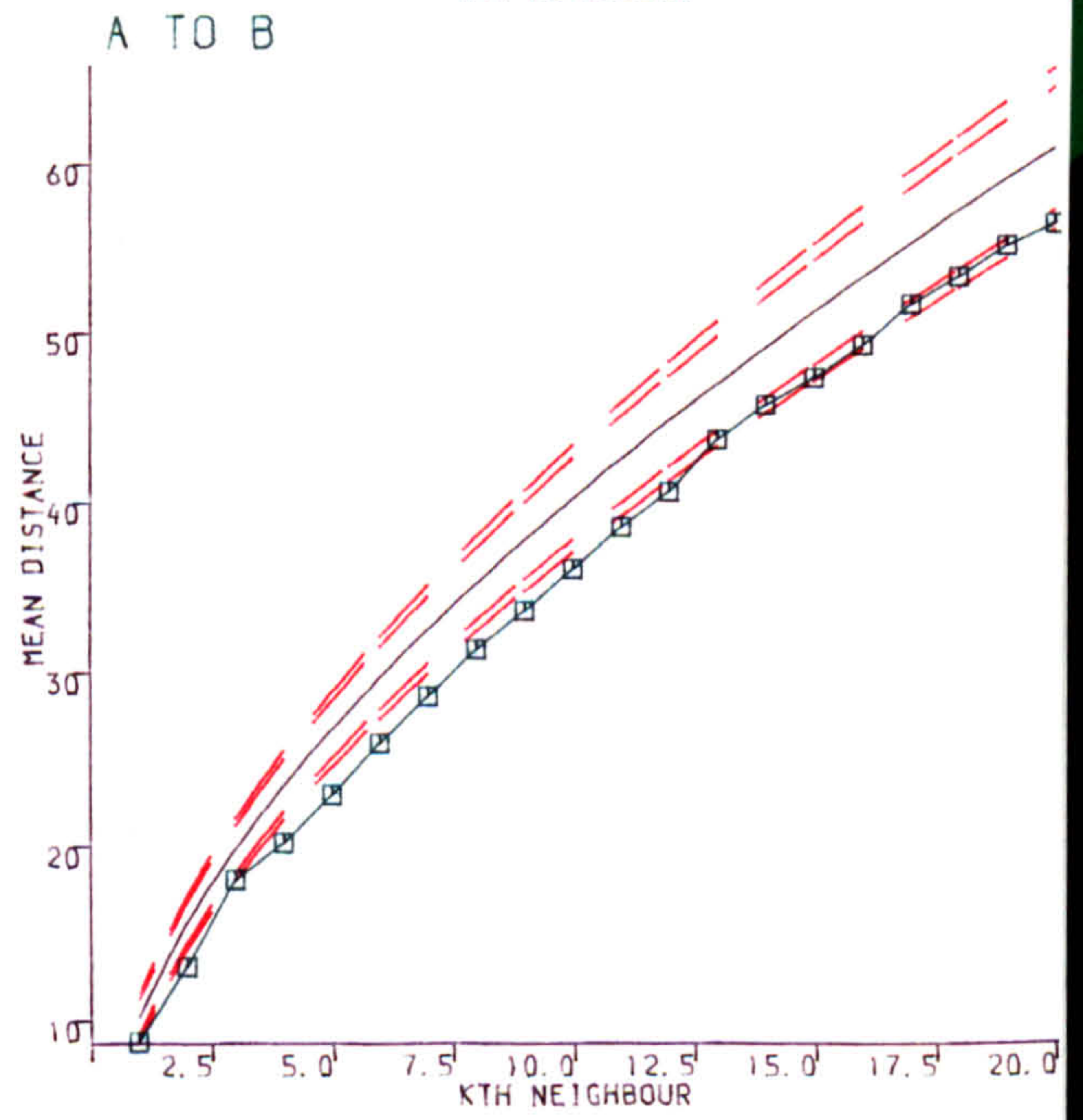
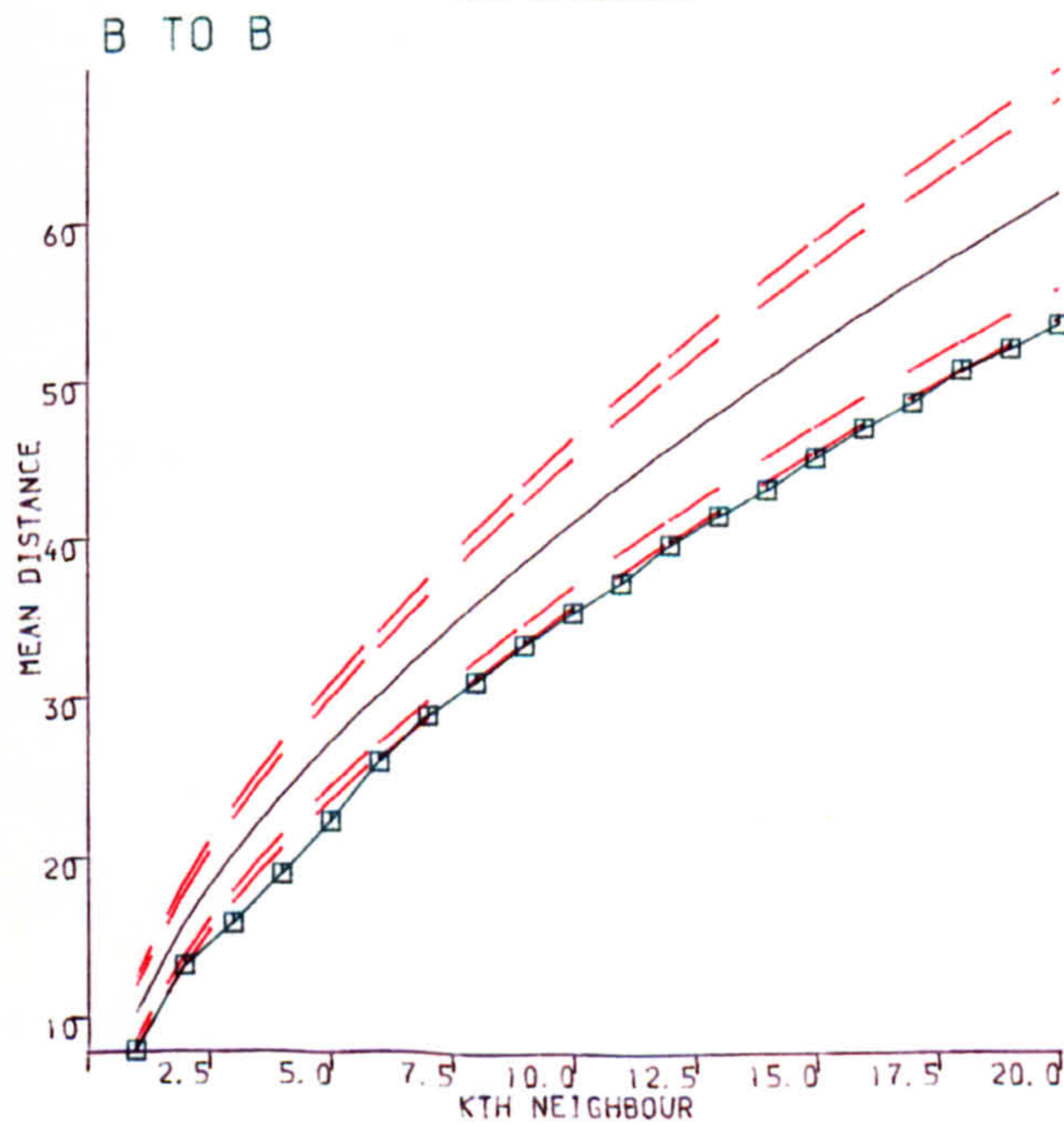
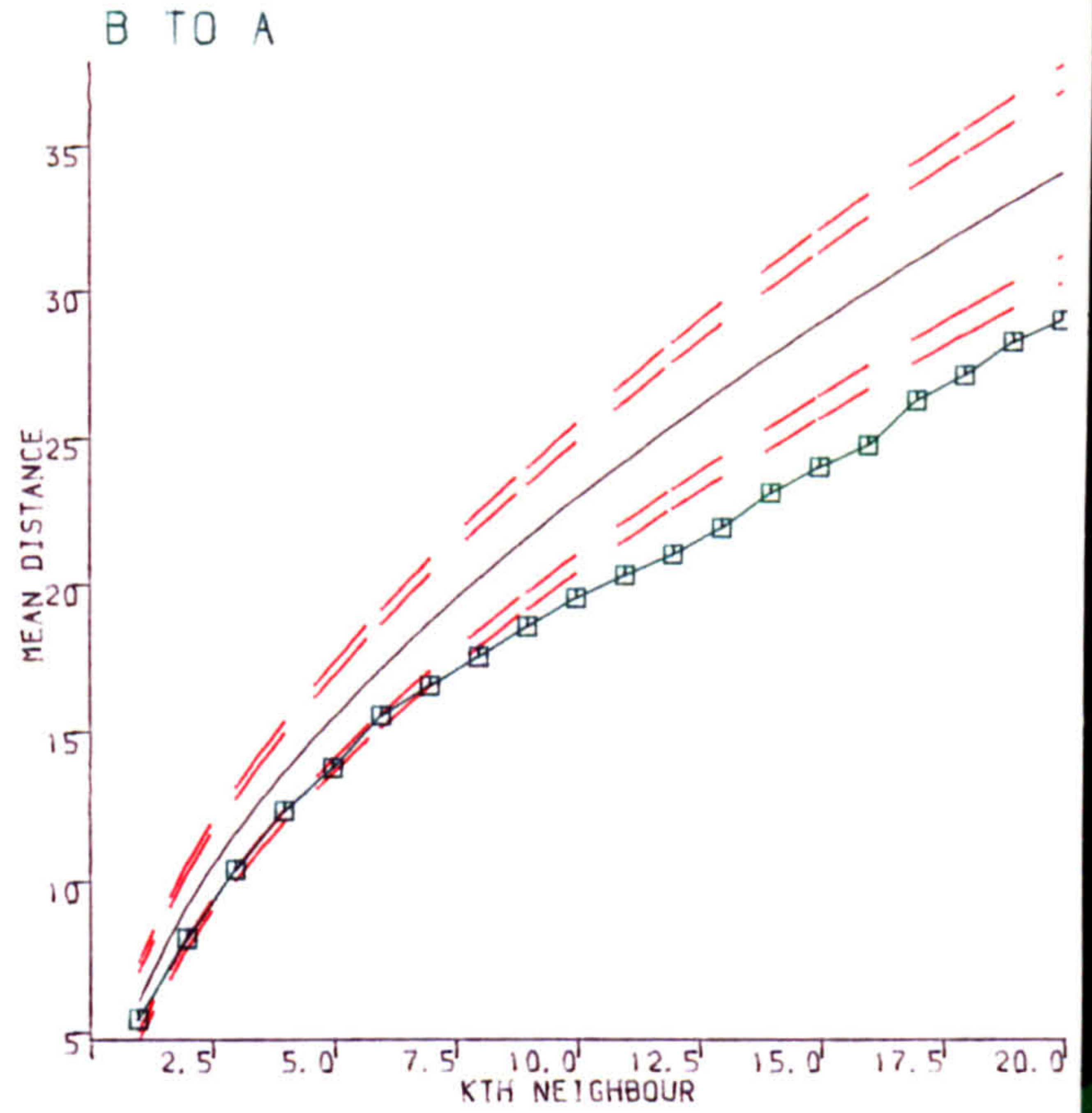
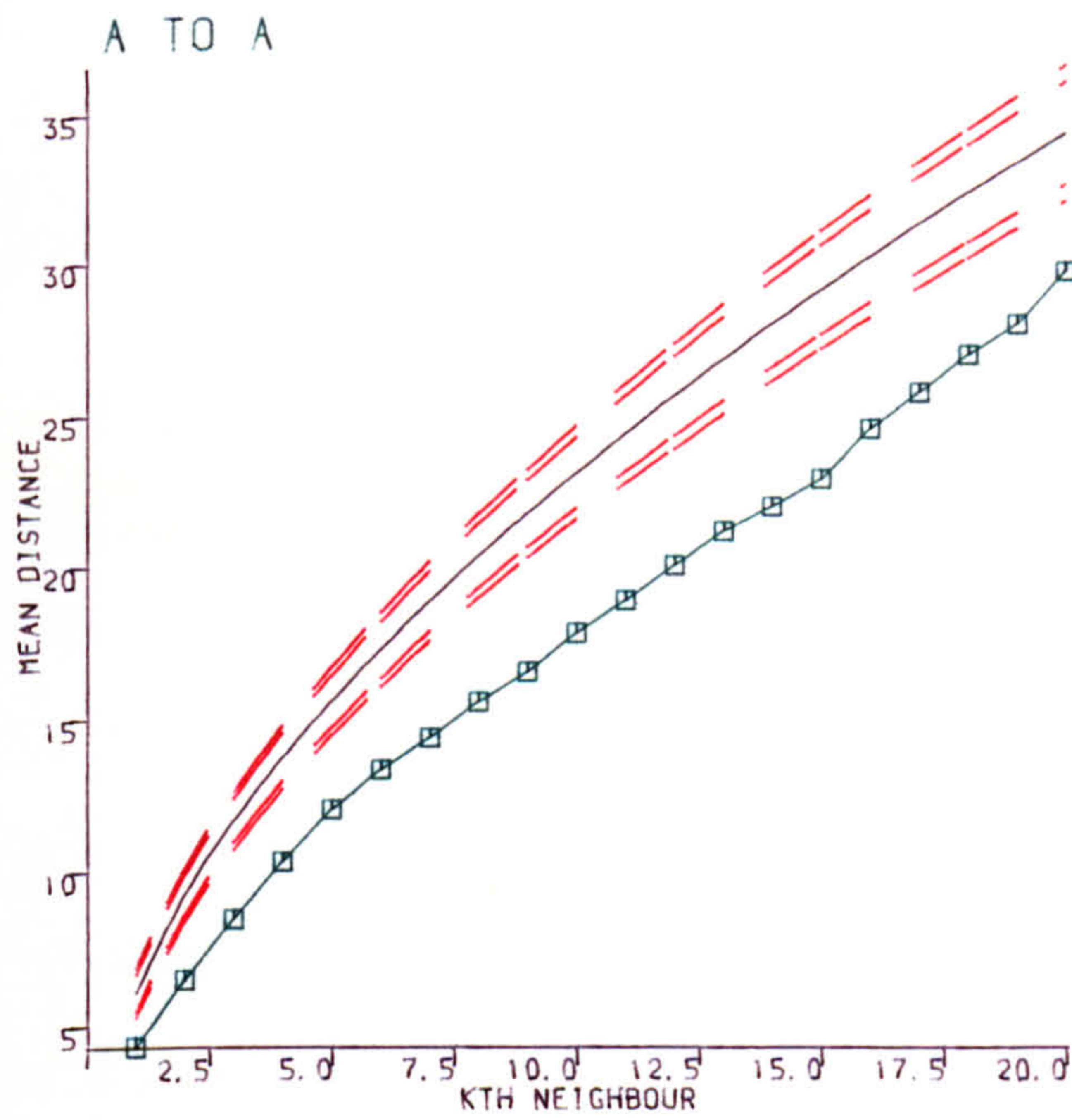
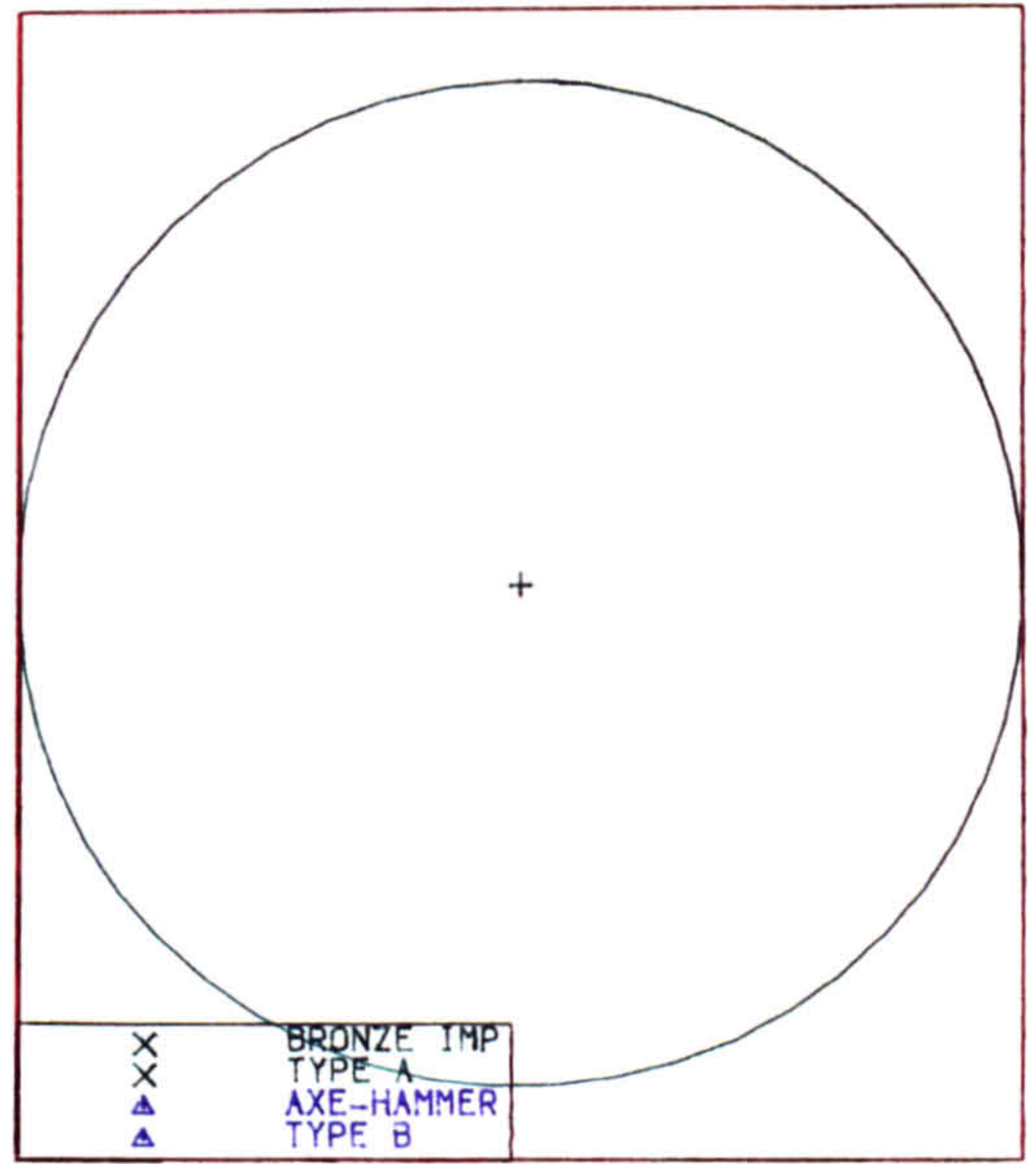
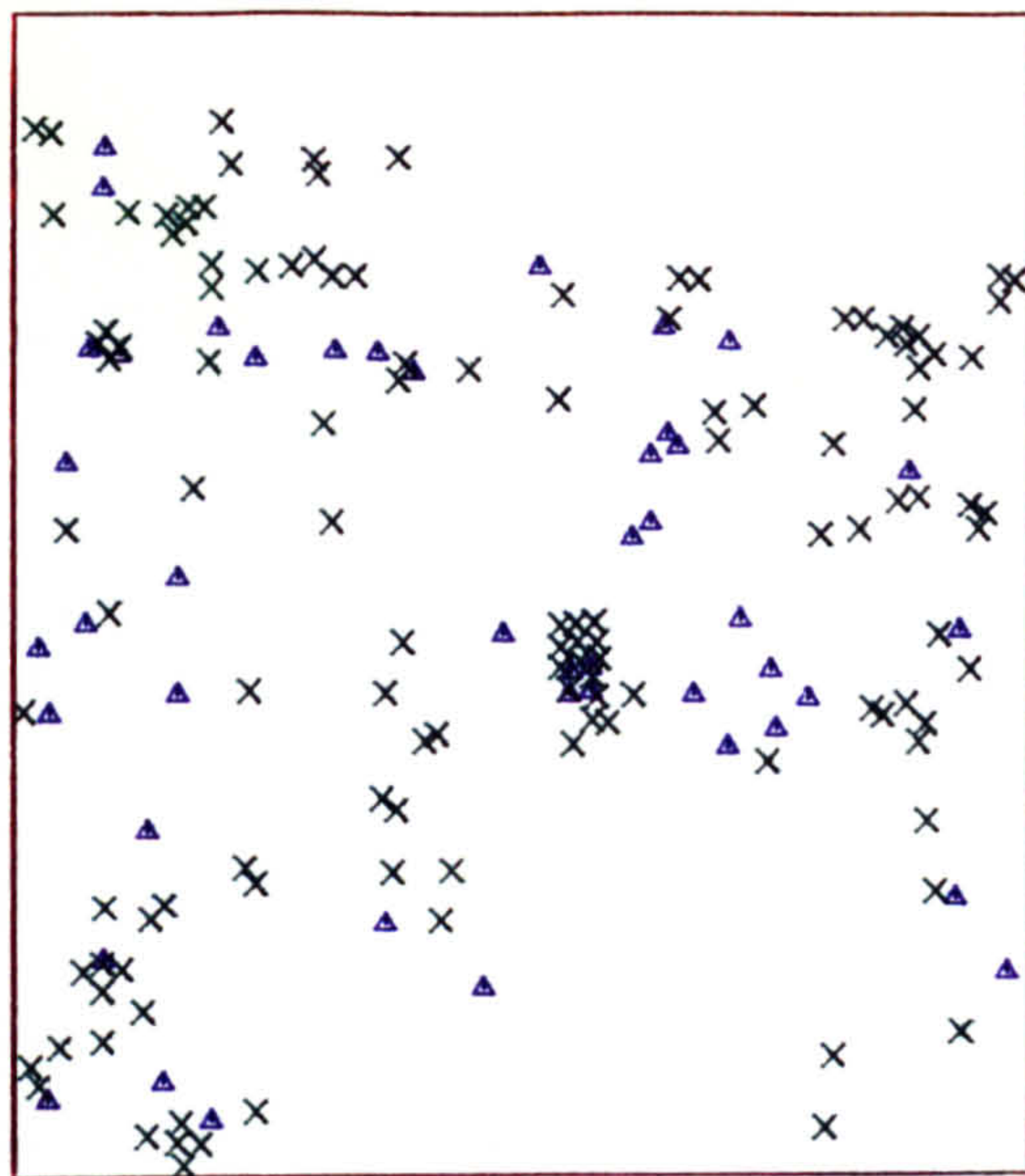


Figure 10.7 Vale of York analysis I.

just show significant association for all twenty nearest neighbours of the analysis. These results seem to indicate that both the bronze implement and axe hammer distributions are non-random and therefore likely to be associated with the well drained soils. However for the axe hammer to bronze implement relationship there seems to be no local spatial association up to the sixth nearest neighbour. This could indicate that although both axe hammers and bronze implements were deposited on the areas of well drained soils during the Bronze age, the depositional patterns within these areas were not related, i.e. the equating of both distributions to occupation areas is questionable.

The Bronze age distribution of bronze implements and the Neolithic distribution of polished axes can be considered next (figure 10.8). The bronze implement distribution as before is strongly clustered. The polished axe distribution also shows significant spatial clustering. Both the polished axe to bronze implement distances and the bronze implement to polished axe distances show significant association. These results clearly show that the spatial distributions of bronze implements and polished axes are non-random and are strongly spatially associated. This supports Radley's interpretation.

Finally the Bronze age distribution of axe hammers and the Neolithic distribution of polished axes can be considered (figure 10.9). The axe hammer distribution just shows significant clustering while the polished axe distribution is strongly clustered. The polished axe to axe hammer distances show significant association while the axe hammer to polished axe distances show a fluctuating pattern of association and non-

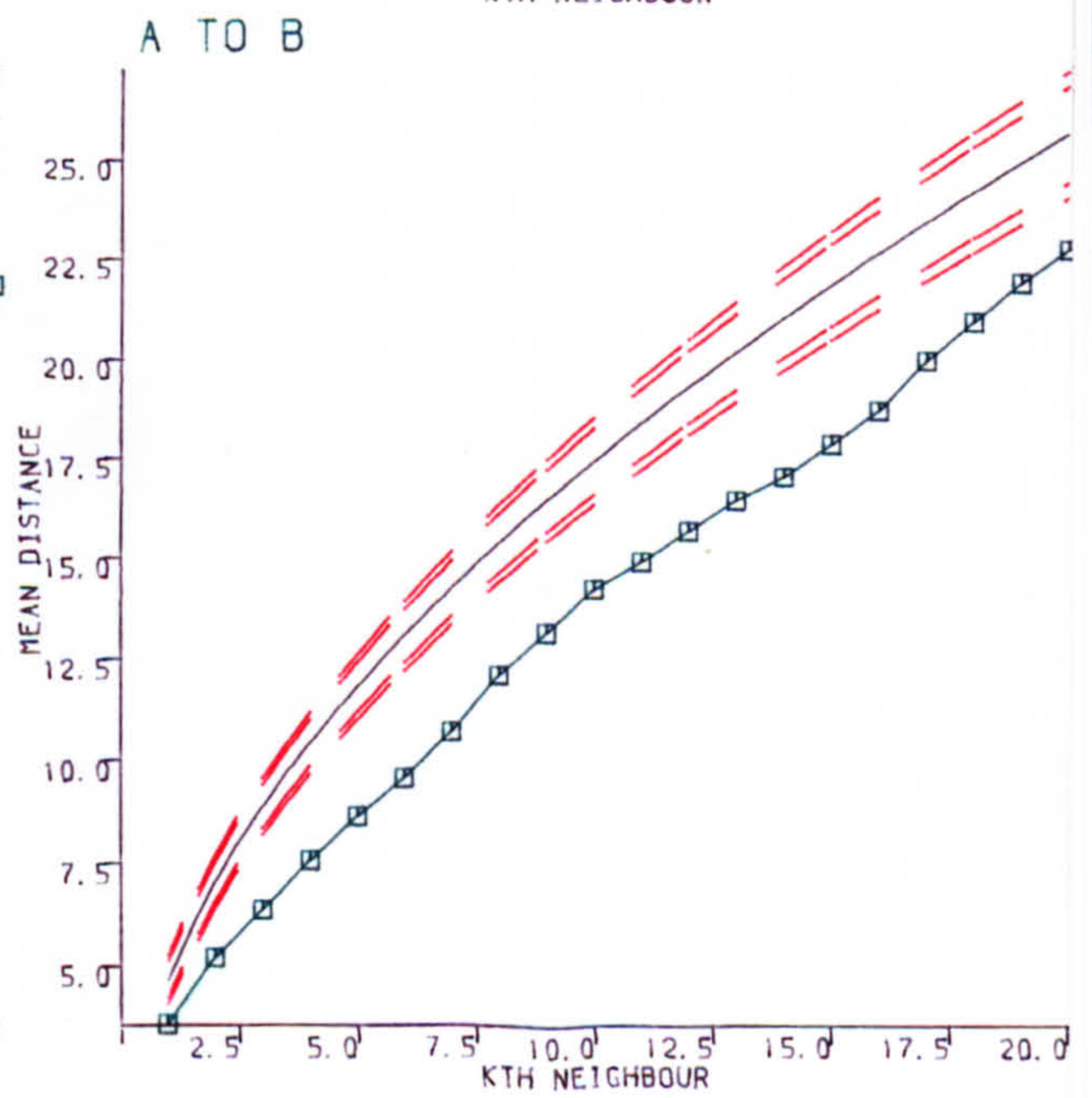
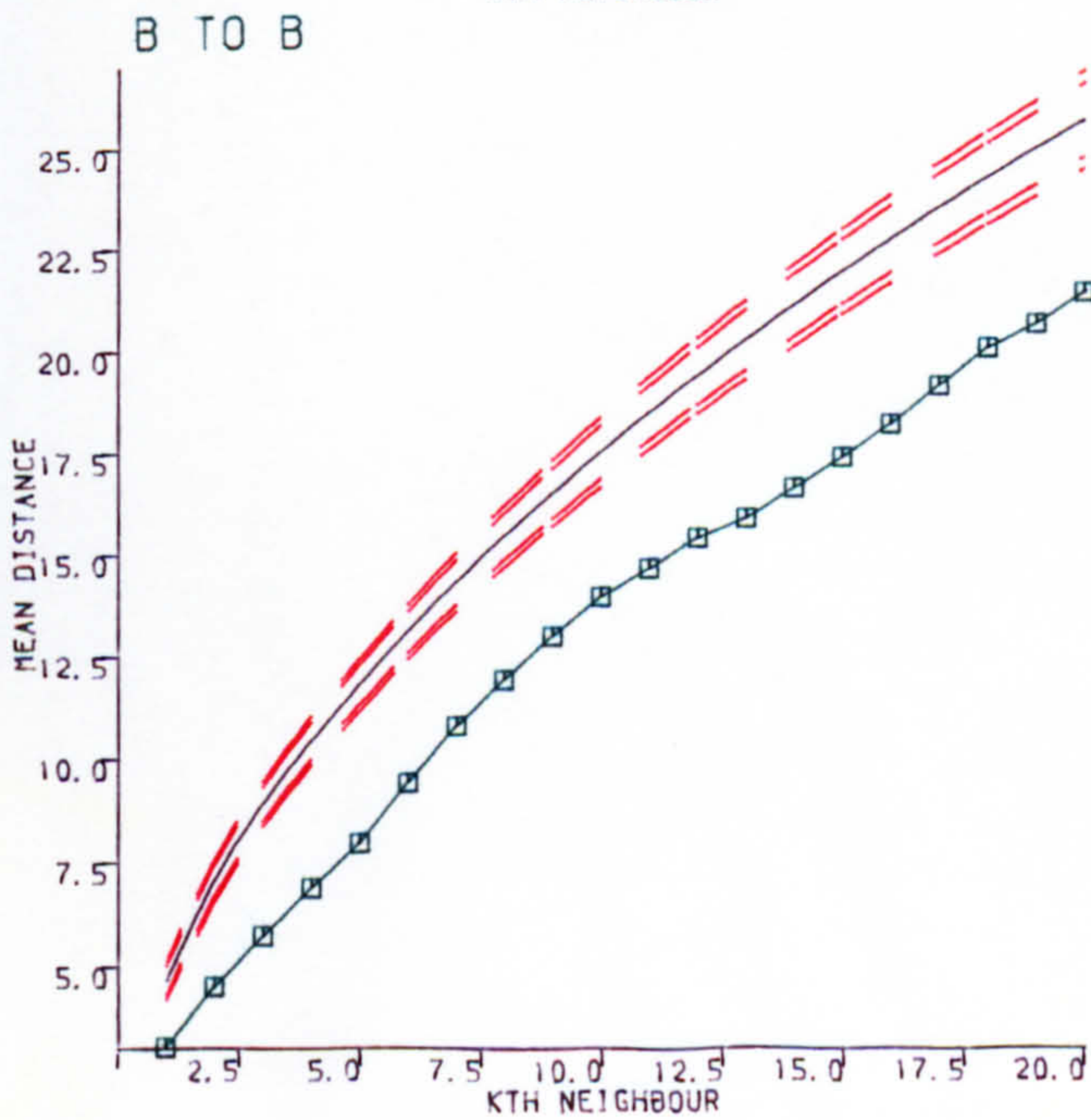
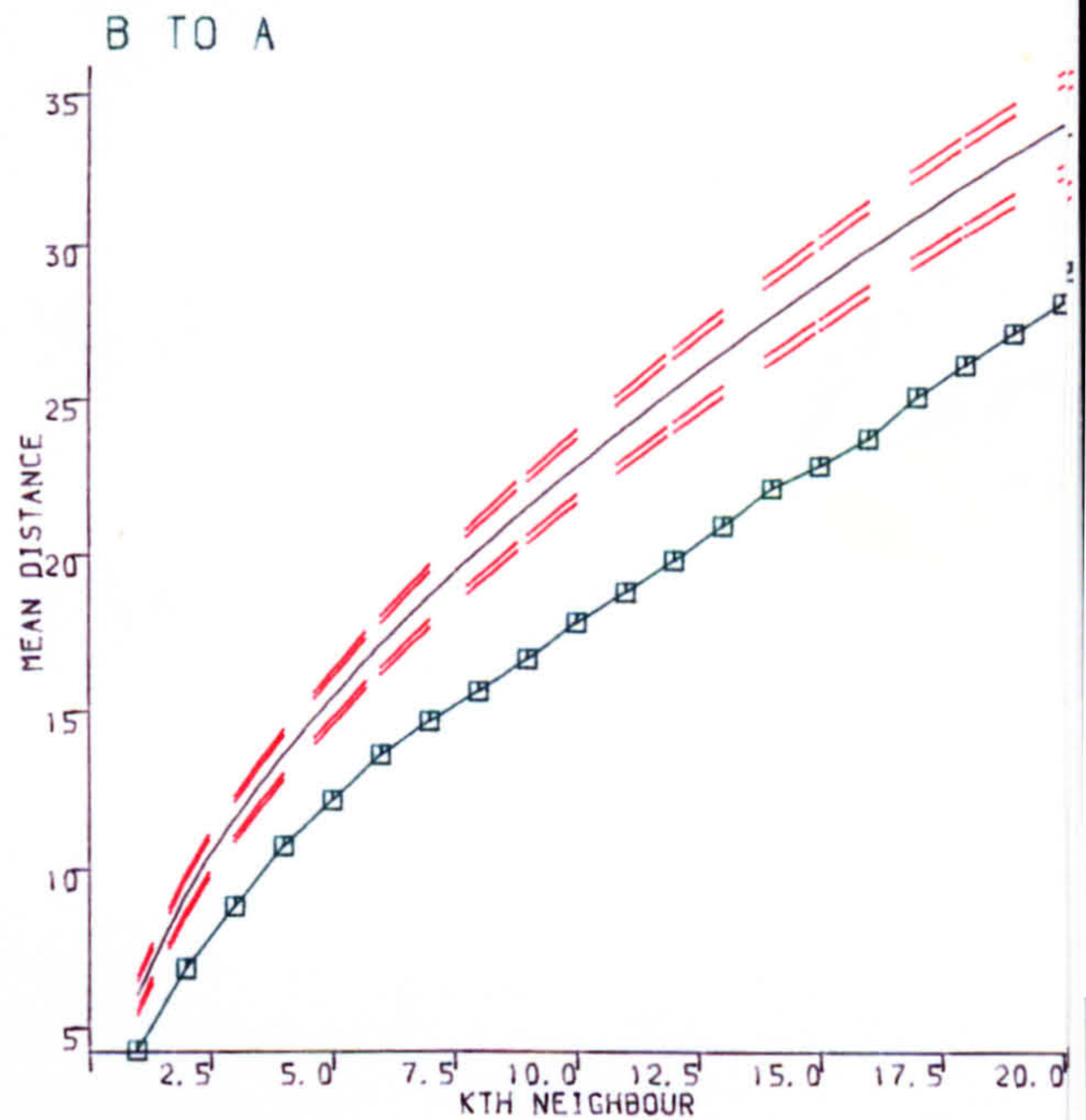
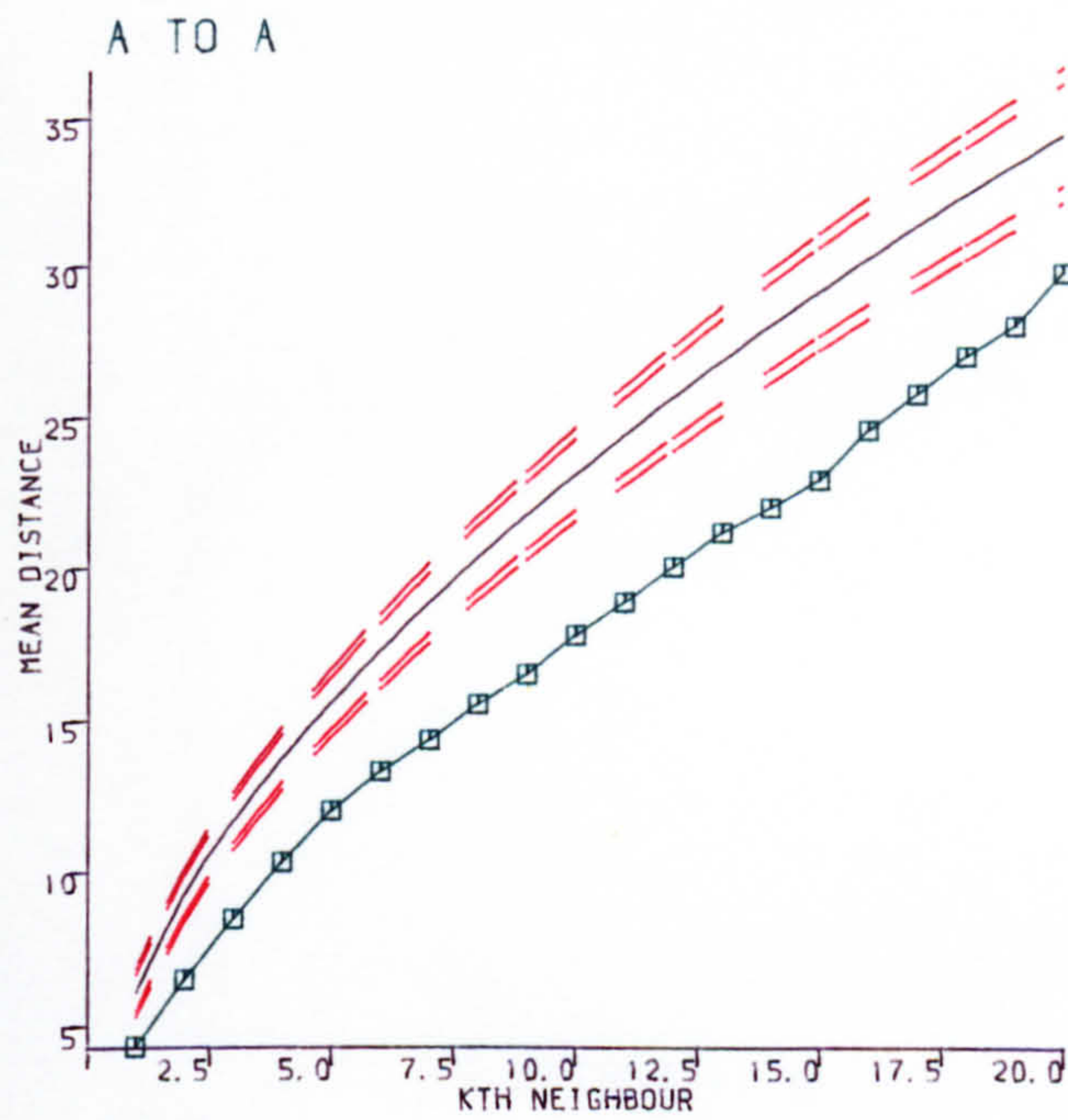
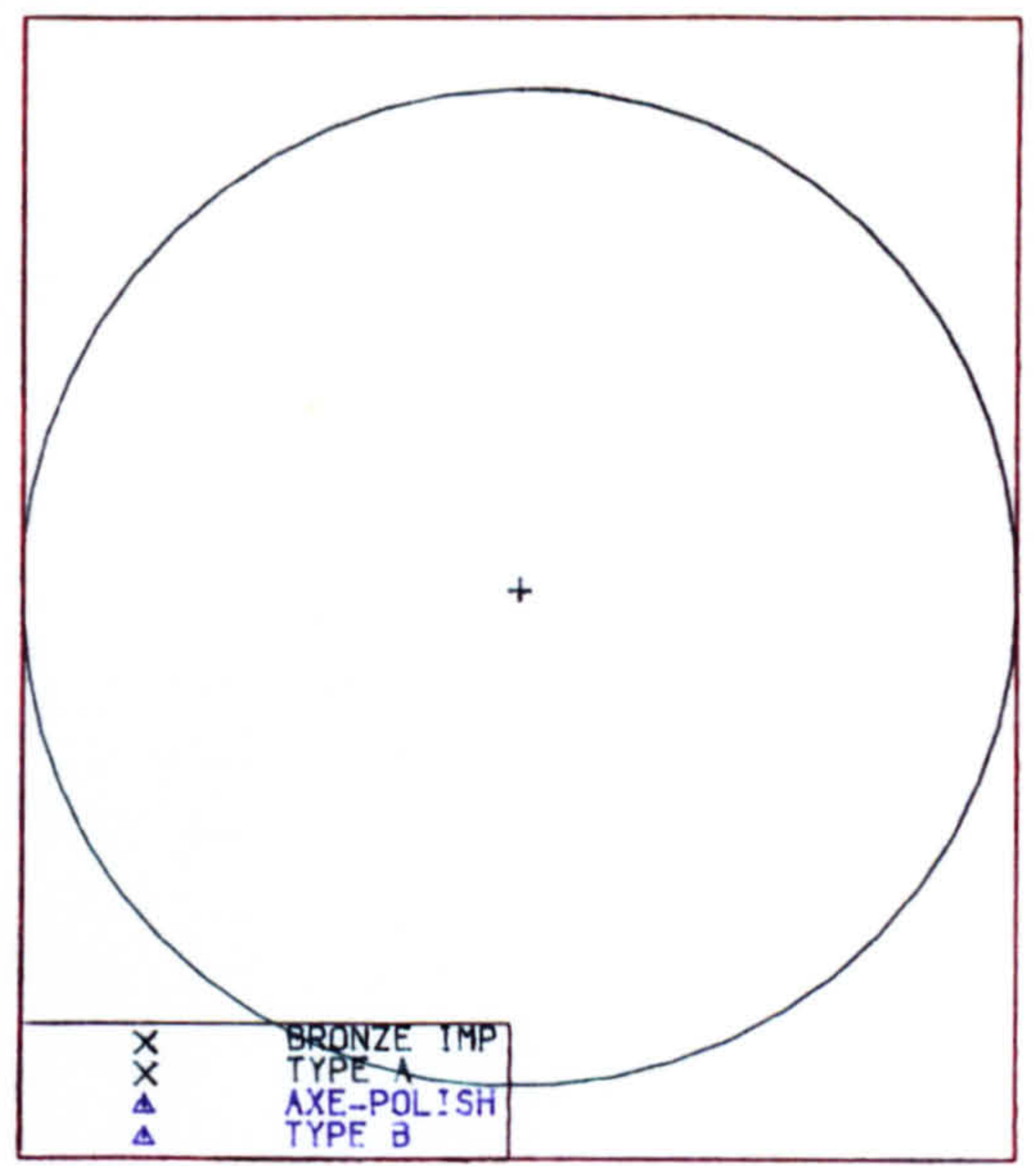
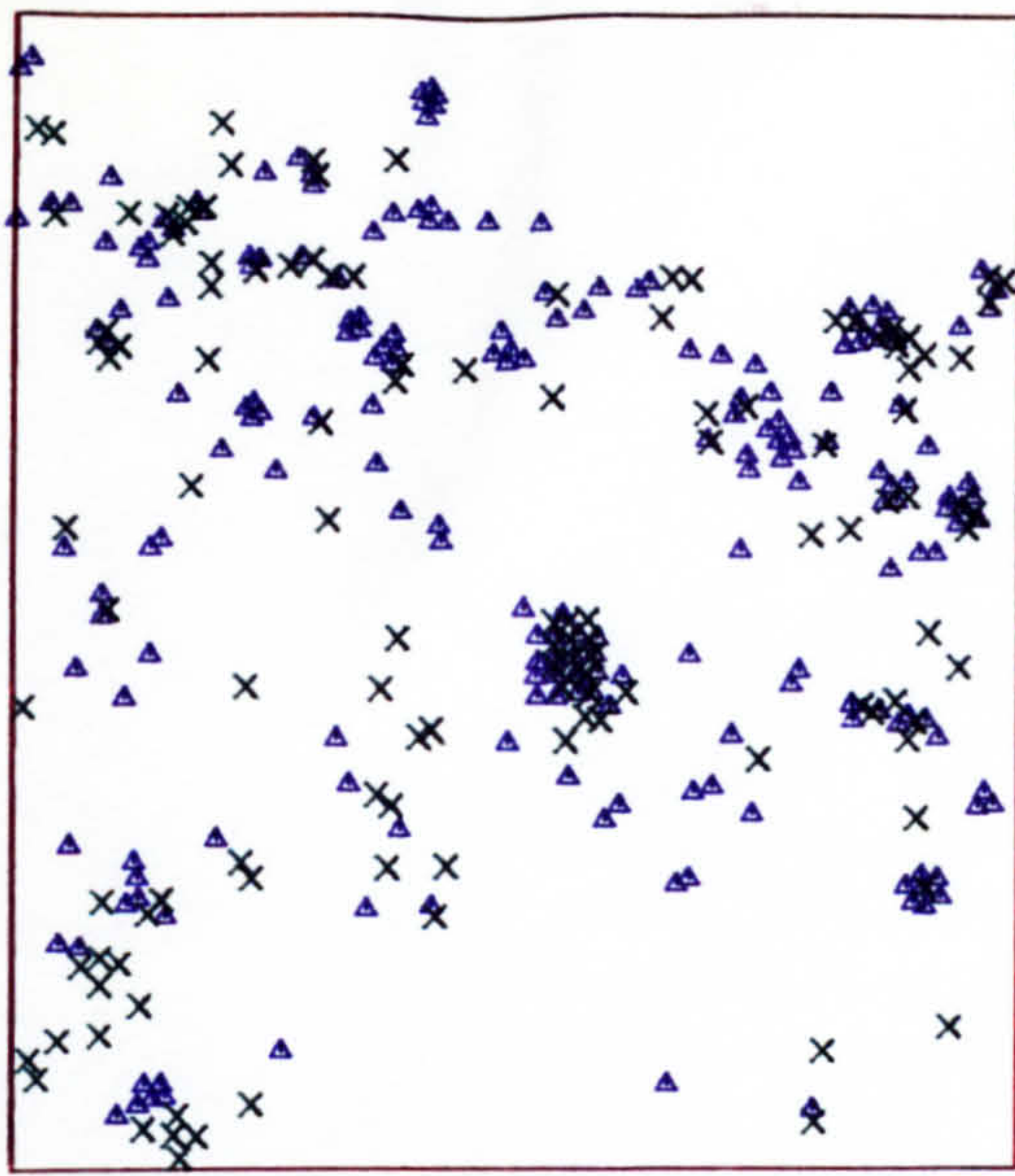


Figure 10.8 Vale of York analysis II.

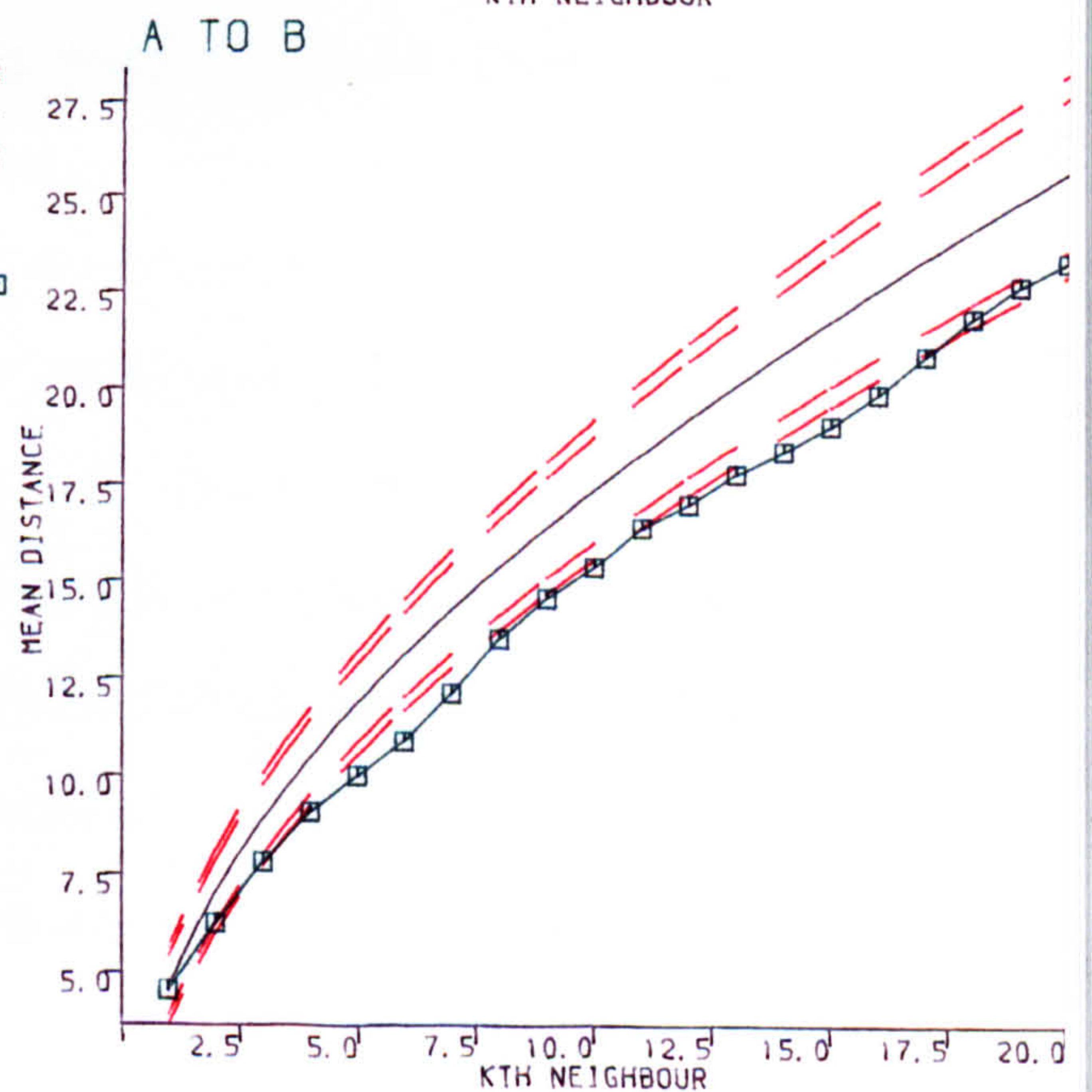
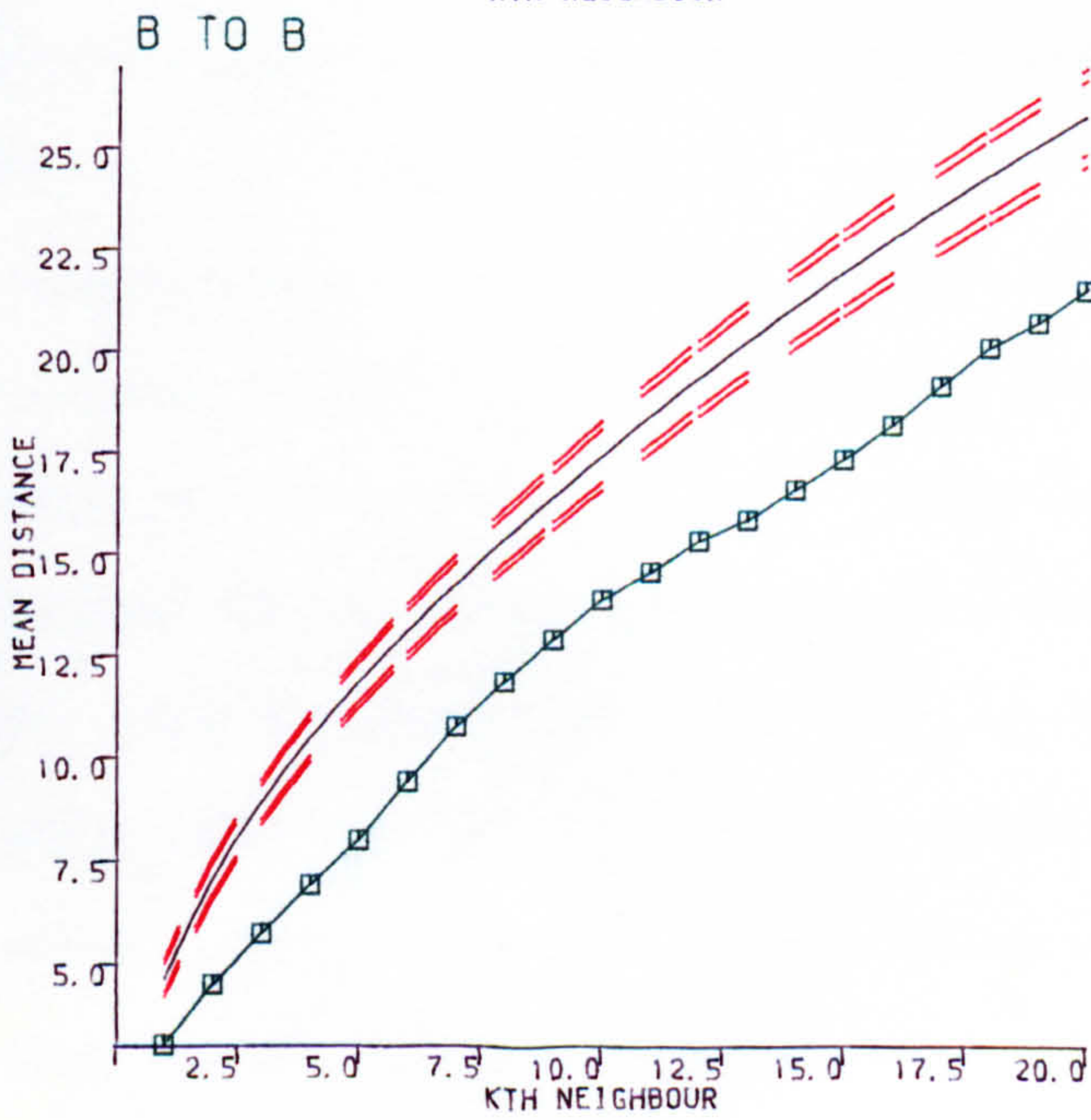
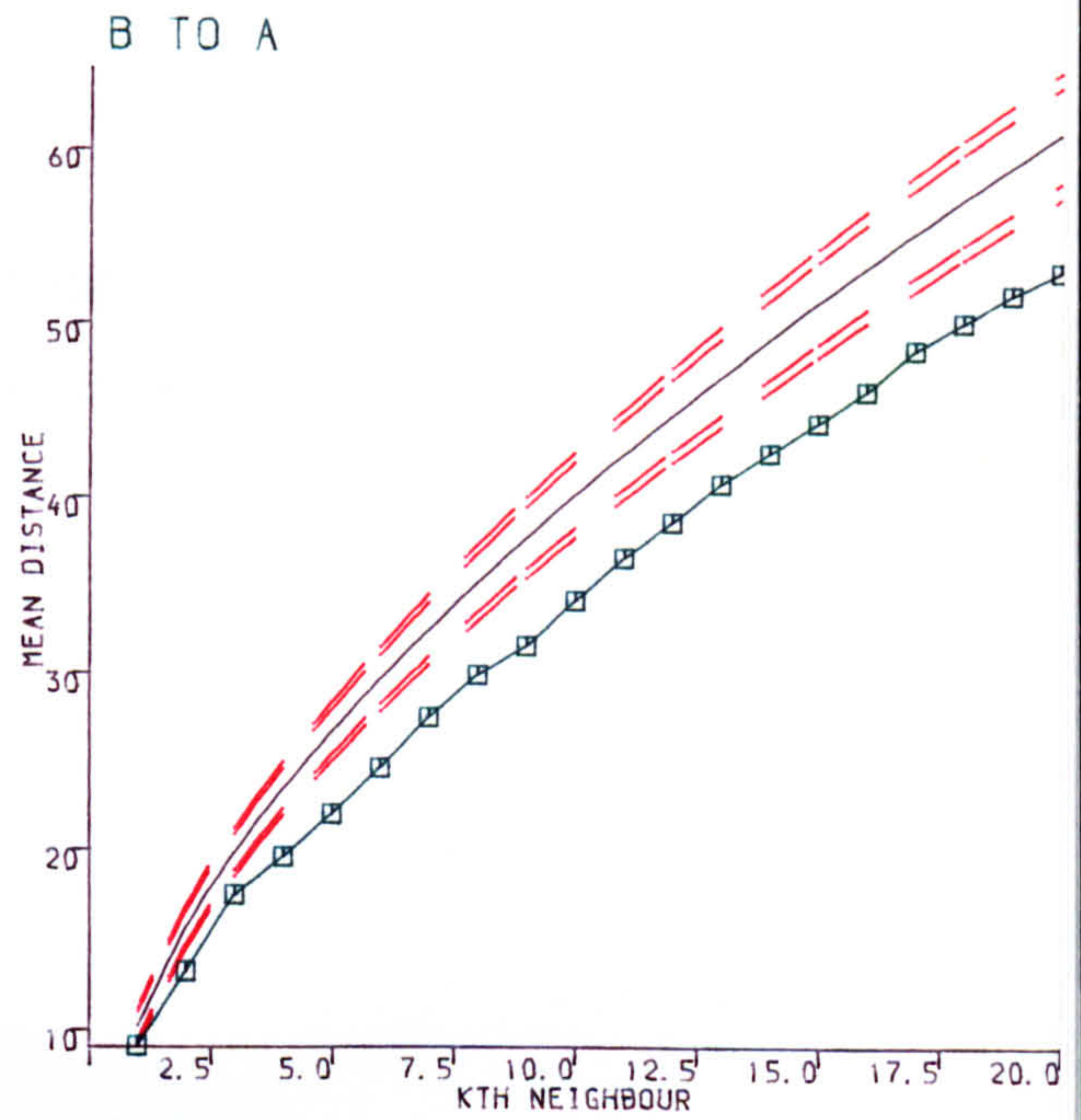
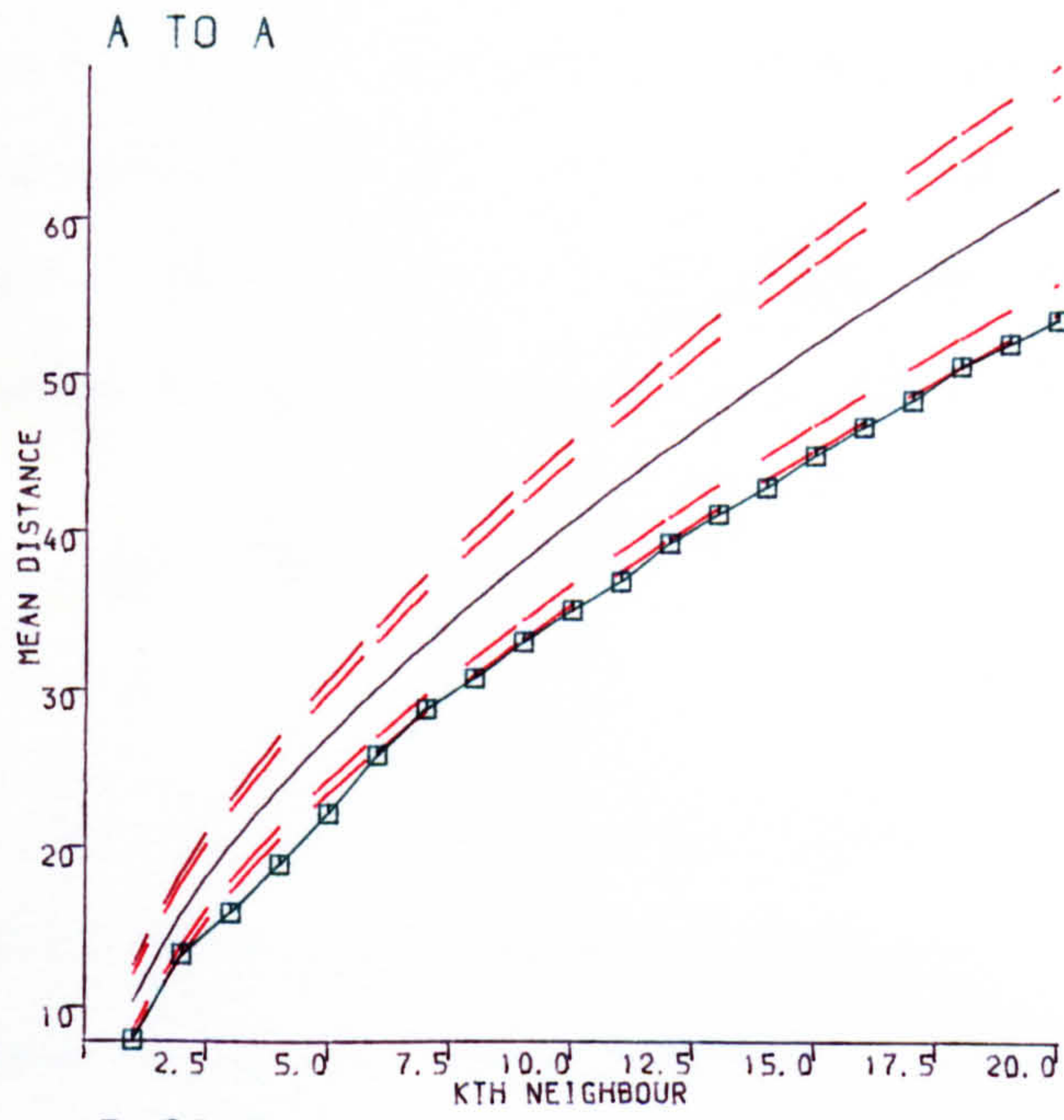
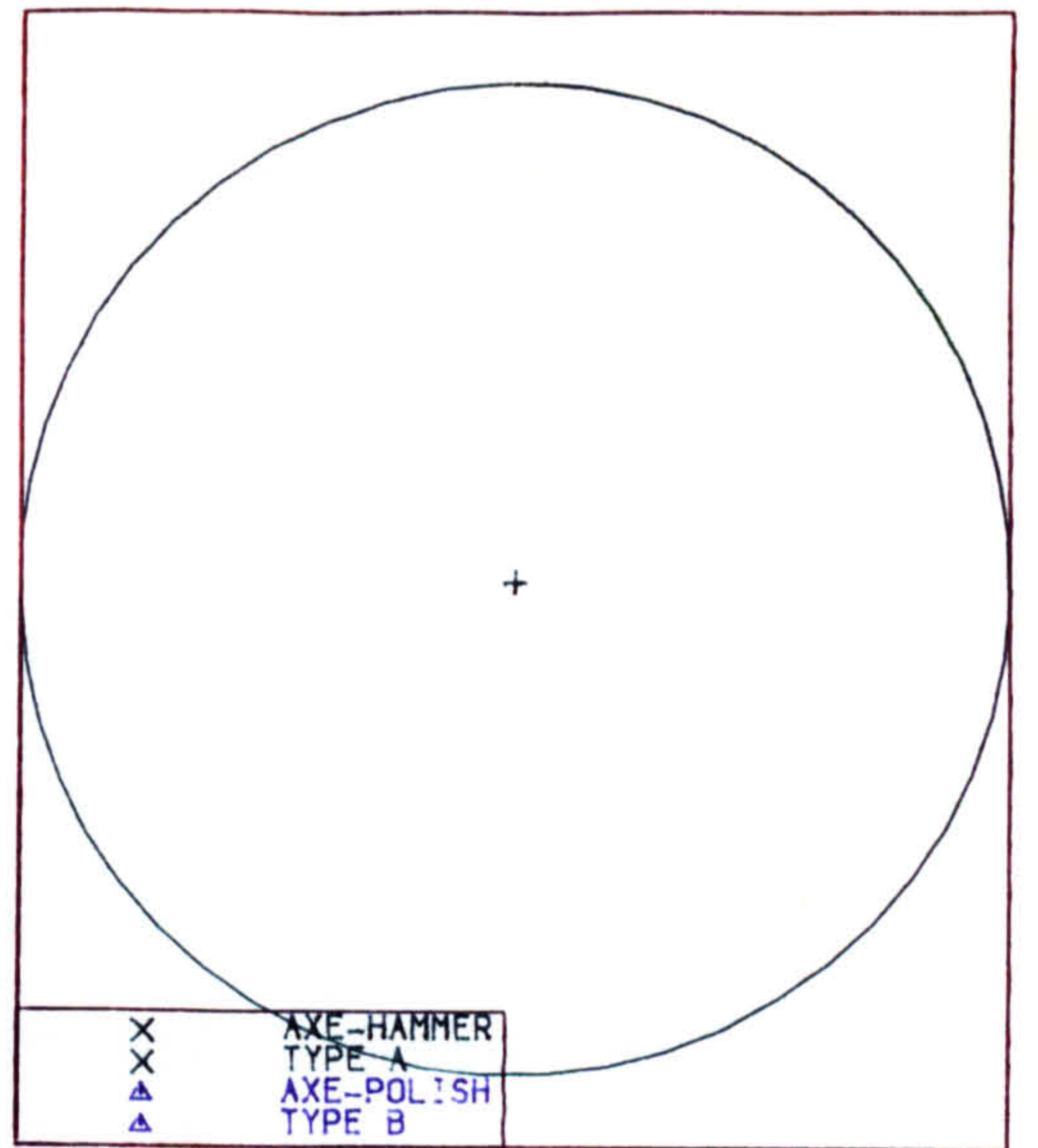
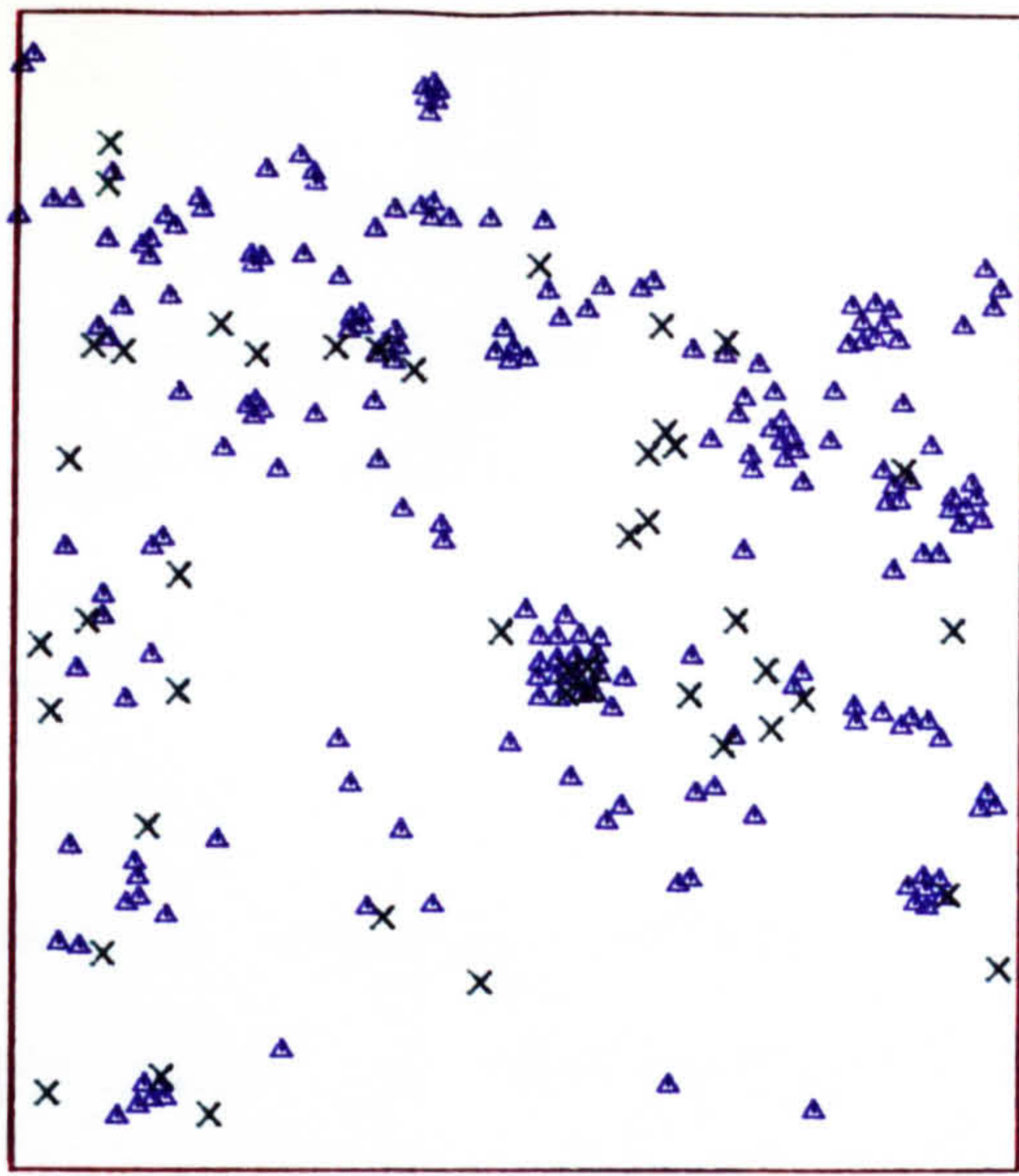


Figure 10.9 Vale of York analysis III.

association. These results generally support Radley's interpretation, although the relationships involving the axe hammers are slightly anomalous.

Overall these analyses do seem to support the general assertion that during the prehistoric period the majority of the occupation in the Vale of York was confined to the higher well drained soils. However there is some indication that within this pattern there is some differing depositional processes for some classes of material, i.e. bronze implements and axe hammers. It should also be noted that the accuracy of the published data on which these analyses were based is in doubt, specifically the clusters of points at York (at the centre of the distribution maps) are very uniform.

10.7 DISCUSSION

The role of spatial association of lattice and point processes in archaeology is usually somewhat different than for spatial association of surface processes. As described in 8.0 and 9.0 surface processes are often found being used as either prospecting or survey tools to define the extent, form and type of archaeological information subsequent to further analysis. In contrast lattice and point processes are often found being used for detailed archaeological interpretation of established data. For example the most well known use of association of point data is in the determination of tool kits on Palaeolithic sites. Here the point data represents the spatial positions of different types of stone tools. A tool kit is a group of stone tools that were used together and hence would be expected to be spatially associated.

From these observations it would be easy to argue that overall lattice and point processes play a more important role than surface processes in archaeology as they are concerned with primary archaeological interpretation. However it is more important to see all these processes not in terms of their specific applications but in terms of their generic usefulness. By tying processes and analyses down to particular applications, the possibility of developing more generalised and hence more powerful analyses is lost.

PART 4

11.0 CONCLUSION

11.1 SUMMARY

The main objective of this thesis, the investigation of the problems of spatial association within the context of site survey and regional distribution analysis, has been illustrated with data drawn from the prehistoric record of North Yorkshire and collected from the Romano-British ladder settlement at Heslerton. Part one provides very brief summaries of the geomorphological and archaeological backgrounds to these sources.

Part two describes the survey techniques used at Heslerton; geophysical surveying, surface collection and aerial photography. In chapter four the problems of data presentation are considered and contouring is justified as the best method of display of a systematic sample of a continuous function defined on a plane. Contouring is flexible and robust analytically and helps to define the outline of archaeological features; The BAGS software was set up as a complete and sophisticated system of producing contours.

Of the geophysical surveying techniques considered in chapter five, only resistance surveying proved suitable and even then conditions were marginal. Conventional resistance surveying is problematic as errors occur in manual recording and it is difficult to process the readings in the field. The automatic field logging system, based on a portable microcomputer, solves these problems. Using a specially designed interface to the resistance meter the system allows data to be captured automatically in the field,

stored, processed as dot densities for initial interpretation and transferred to other computer systems for further processing. The Heslerton study area survey results contained a high noise to signal ratio due to the effects of modern ploughing. Filtering techniques were ineffective because, as with most archaeological data, the sampling density was too low.

In chapter six the phenomenon of surface distributions was explored. A major problem lies in determining whether a direct relationship exists between surface distribution patterns and buried archaeology. The previous research in this area indicates that no general laws can be formulated but that a pragmatic approach must be taken. For the Heslerton study area various field experiments and observations showed that surface material reflected quite accurately the location of buried archaeological features. Using a sampling collection scheme, within contiguous quadrats, a record was made of the surface distributions. There was tight control and monitoring of fieldworkers to minimise collection biases. The spatial distribution of the collected Romano-British pottery showed definite structure in the area of occupation associated with the ladder settlement.

The aerial photography evidence for the Heslerton study area was evaluated in chapter seven. Aerial photography is a widely used survey method in archaeology but presents difficult problems of interpretation. There are also technical problems of geometric transformation from photograph to map plane. The aerial photography features visible within the Heslerton study area are all archaeological and were transformed onto an ordnance survey map with the use of a microcomputer system. They define part of the actual

ladder settlement alongwith adjacent field boundaries which indicate areas of cultivation.

Part three develops various concepts within spatial association. Chapter eight models the data and methods of spatial association for site survey as operands and operators. These allow both complex data structures and processing strategies to be built. For simple association operands need to be type and structure compatible. Measures of density and complexity converted the aerial photography data to a gridded form. Sampling densities were then adjusted to a common value. These operations represent transformations of operands to equivalent types or spaces. The direct association of data within disjoint data spaces is probably impossible. However the movement between these spaces, using transformations and inverse transformations, has exciting potential. An inverse data transform could be used on geophysical data to define actual feature outlines. Unfortunately such an inverse transform does not conform to proper mathematical definition. Operand transformations can alternatively be viewed as an internal restructuring of a data space. The data space for pottery quantification is often formed by data collected on a basis of ease of access rather than archaeological importance. By defining a measure of pottery brokenness for the Heslerton material a more archaeologically relevant structure, than other more common methods of quantification, was formed. The archaeological value of data should be measured in terms of the range of procedures that can be applied, a measure of the generality of structure, rather than simple intrinsic worth.

Chapter nine develops spatial association operators for the survey data. By equating the data with the results of surface

processes, they can be associated by using either global or local operators. Global operators included the use of bivariate correlation, multivariate correlation, bivariate linear regression and multivariate linear regression. Although easy to apply they do not detect local heterogeneity in the data. Only the aerial photography density measure and the pottery density showed any strong relationships with these global operators. In contrast to global operators, local operators try to define some structure within the data. Similarity maps do this at the individual reading scale, while local association maps allow analysis of coarser partitions of data. When applied to the Heslerton survey data sets the aerial photography data tended to define features which had high concentrations of pottery internally.

The analysis of regional lattice and point processes is investigated in chapter ten. Regional lattice data is often treated as the result of an irregular sampling of a surface process. To illustrate such data a contouring method for scattered point data is used. Furthermore the methods of local association described in chapter nine are, with care, also applicable. Point processes have received a lot of attention within quantitative archaeology, most notably with the use of nearest neighbour analysis. This technique has however several theoretical problems associated with its use on archaeological data. The two most important problems are edge effects and the non-sampled nature of most archaeological data. Various solutions have been attempted, but all are limited to first order neighbours only. A new correction for K^{th} order neighbours, developed initially formulating a general analytical correction and then using computer simulation to form empirical correction

constants, provides a powerful solution to these theoretical problems. In comparison to previous corrections for first order neighbours this new correction compares well. It can be used as a basis for a description of K^{th} order association for point data, as illustrated by the analysis of several prehistoric artifact distributions in the Vale of York.

11.2 DISCUSSION

Throughout the thesis emphasis has been firmly placed on the exploration and development of techniques and their practical applications; Illustrative material being taken from the prehistory of North Yorkshire and native Roman settlement in the Vale of Pickering. In this context the relevance of the thesis can be judged not only on the number of problems tackled but also on the number of new questions raised.

Reviewing the contents of the preceding text several central themes emerge which have a wider relevance than just spatial association. The first of these concerns the relationship between what could be termed pure archaeology and formalised methods of investigation, as represented by statistical and mathematical techniques. A major problem in this area is to institutionalize the concept of models within archaeology. A model is a simplification of reality that allows a consistent working solution, to an often complex problem, to be formed. Many archaeologists either reject models because they constantly want to go outside of them, or else try to produce such large all embracing models that they are unattainable and therefore of no practical use. In science there is

an observation, model building, model testing sequence. This is very hard to relate to archaeology as many observations are non-repeatable. Excavation is the often quoted 'unrepeatable experiment', while for economic reasons potentially repeatable fieldwork is often a once only exercise. Traditional spatial analysis basically sets up a model of randomness. Statistical tests can then be used to judge if observations tend to deviate from this model towards clustering or regularity. For a lot of archaeology this modelling of global spatial structure has very limited use. For example given several distributions of tool types across a mesolithic occupation area, the question of whether they are dispersed or aggregated is not as archaeologically interesting as determining where possible centres of aggregation lie. The model itself is not wrong, but it is the wrong model to answer many significant archaeological questions. These same comments apply to global models of association.

At the other extreme there are methods of purely descriptive use, for example contouring. These offer only subjective interpretation, therefore there is a need for a compromise in the form of local models for spatial analysis and association. However as should be evident from the preceding text there is no easy definition of how local, local should be. There is no obvious solution to this problem from the mathematical/statistical side. Many of the techniques presented in the thesis can be validly used on any arbitrary partitioning of the data. It is up to archaeology to define the scale at which modelling is relevant. This may seem an arbitrary choice and can be. But the same is true for all statistical methods, they are all open to abuse. What is

fundamentally needed is a maturity in the approach of archaeology to its methods and techniques. A willingness to define formal goals or objectives would allow a more satisfactory use and development of techniques in spatial association and analysis.

Even if such changes could be effected, there still remain several philosophical problems regarding spatial association in archaeology. Throughout the thesis a distinction has been made between cause and effect. In any form of spatial association, the data being analysed are a static pattern. Any analysis results are therefore intrinsic to the pattern. This forms another major criticism of spatial methods in archaeology, that although we can attempt to analyse patterns we can not get directly at the processes that produce them. To some extent this objection is based on a simplified view of how spatial patterns develop within archaeological data. This view can be divided into two opposing arguments, the deterministic and the stochastic. The deterministic argument states that all human behaviour is deterministic in execution if not in intent. Thus spatial patterning produced by human behaviour should show discernible structure. This structure may however be disrupted by subsequent processes, acting on the spatial patterning, that have no direct relationship with the original human behaviour. The effect of these subsequent processes is both to destroy existing structure and introduce new spurious patterning. In contrast the stochastic argument states that although at the individual scale human behaviour is deterministic, when viewed at a larger scale it becomes stochastic. Thus spatial patterning produced by human behaviour should often seem random.

In perspective both arguments are equally valid, although they

both over simplify the relationship between cause and effect. However, the arguments are irrelevant to present techniques of spatial analysis as even the most sophisticated are based on comparatively simple probabilistic models. These allow only, what some archaeologists would term naive hypotheses, to be posed. It is very unlikely that more powerful techniques, if developed, would allow more archaeologically significant hypotheses to be formulated. At best spatial association and spatial analysis within archaeology, should be able to offer objective testing of relatively simple yet relevant hypotheses about the spatial structure of archaeological data. This information should form only a part of the final 'archaeological interpretation' of that data.

APPENDIX A
SOFTWARE OUTLINES

A.1
Program outline for contour plot

contour_plot

BEGIN

open and read REPORT file (contains grid names and locations);
copy data files to binary direct access scratch files, adding
the bottom row of top neighbour and left column of right
neighbour;

read contouring parameters;

initialise plotting;

FOR each data grid DO

BEGIN

calculate relative origin;

contour_grid

END;

close plotting

END.

contour_grid

BEGIN

remove degenerate nodes by adding 1% to node value;

FOR each contour DO

BEGIN

set up markers for intersections;

draw_open_contours;

draw_close_contours

END

END;

draw_open_contours

BEGIN

trace_contours from left edge of grid;

trace_contours from top edge of grid;

trace_contours from right edge of grid;

trace_contours from bottom edge of grid

END;

draw_close_contours

BEGIN

WHILE any remaining markers for intersections DO

trace_contours

END;

```
trace_contours
BEGIN
  WHILE not reached edge of grid or not reached starting point DO
    BEGIN
      eliminate marker for intersection;
      plot contour through cell using bilinear spline;
      calculate next cell entered
    END
  END;
END;
```

A.2
System outline for BAGS

BAGS is a CCL procedure. It turns pagination off. Gets an object program START. Runs it. This creates a CCL procedure called FIRST. FIRST is executed. Pagination is turned on.

START is an object program. The source program is written in Fortran 77 with inclusions from the Bradford Subroutine Library. It allows the selection of the store data or process data option. Creates CCL procedure FIRST.

FIRST is a CCL procedure which calls

Either:

INDATA is a CCL procedure. Gets an object program ENTER and system catalogue BAGSCAT. ENTER is run. This creates a CCL procedure called SECOND. SECOND is executed.

ENTER is an object program. The source program is written in Fortran 77 with inclusions from the Bradford Subroutine Library. It allows the selection of catalogue entry, index file entry and data grid entry. Creates CCL procedure SECOND.

SECOND is a CCL procedure which executes data management tasks as defined by ENTER.

Or:

PROSS is a CCL procedure. Gets an object program MAINC and system catalogue BAGSCAT. MAINC is run. This creates a CCL procedure called SECOND. SECOND is executed.

MAINC is an object program. The source program is written in Fortran 77 with inclusions from the Bradford Subroutine Library. It allows the viewing or printing of the catalogue. It allows the selection of an index file, and the printing or use of that file. Creates CCL procedure SECOND.

SECOND is a CCL procedure which

Either:

prints the catalogue on the default lineprinter.

Or gets an index file renames it INDEX and:

prints the index file on the default lineprinter.

Or calls:

SELECT is a CCL procedure. Gets an object program SETUP. Runs it. This creates a CCL procedure called THIRD. THIRD is executed.

SETUP is an object program. The source program is written in Fortran 77 with inclusions from the Bradford Subroutine Library. It allows the selection of data files and processing tasks. Creates a data file REPORT with data file selection. Creates CCL procedure THIRD with processing selection.

THIRD is a CCL procedure which calls

Either:

PRTREP is an object program. The source program is written in Fortran 77. It prints the REPORT files to the default lineprinter.

And/or:

FFP is an object program. It executes fourier transform filtering. Implemented by S.S.Ipson.

And/or:

CONT is a CCL procedure. Gets an object program CONSET. Runs it. This creates a CCL procedure called FOURTH. FOURTH is executed.

CONSET is an object program. The source program is written in Fortran 77 with inclusions from the Bradford Subroutine Library. It allows the selection of contouring parameters which are placed in a data file XXXCON. Creates CCL procedure FOURTH.

FOURTH is a CCL procedure. Gets an object program CONTOR. Runs it. This creates a CCL procedure called FIFTH. FIFTH is executed. This sequence of calls may be submitted as a batch job.

CONTOR is an object program. The source program is written in Fortran 77 with inclusions from the NAG and GHOST libraries. It is a contouring program. It reads its parameters from XXXCON and writes its plotting instructions to GRID80. Creates CCL procedure FIFTH.

FIFTH is a CCL procedure. It plots the GRID80 file.

A.3
Program outline for RSCS

```
RSCS
BEGIN
  initialise system;
  define function keys;
  REPEAT
    WAIT for function key;
    CASE function key DO
      BEGIN
        f1 : automatic_log;
        f2 : load_grid;
        f3 : list_catalogue;
        f4 : display;
        f5 : hardcopy;
        f6 : manual_log;
        f7 : save_grid;
        f8 : set_up_catalogue;
        f9 : editor;
        f10: utilities
      END
    UNTIL computer switched off
  END.
```

```
automatic_log
BEGIN
  clear RAM file;
  display grid;
  set for auto log;
  log
END;
```

```
load_grid
BEGIN
  select file;
  read catalogue;
  read file;
  report any errors
END;
```

```
list_catalogue
BEGIN
  read catalogue;
  print catalogue
END;
```

```

display
BEGIN
  select display mode;
  CASE display mode DO
    BEGIN
      dml : statistics;
      dm2 : dot_density;
      dm3 : interval_codes
    END
  END;

```

```

hardcopy
BEGIN
  print RAM file
END;

```

```

manual_log
BEGIN
  clear RAM file;
  display grid;
  set for manual log;
  log
END;

```

```

save_grid
BEGIN
  input file name;
  read catalogue;
  write_file;
  report any errors
END;

```

```

set_up_catalogue
BEGIN
  initialise catalogue;
  write catalogue
END;

```

```

editor
BEGIN
  select edit mode;
  CASE edit mode DO
    BEGIN
      em1 : station_edit;
      em2 : grid_edit
    END
  END;

```

utilities

```
BEGIN
  select utility mode;
  CASE utility mode DO
    BEGIN
      uml : transfer
    END
  END;
END;
```

log

```
BEGIN
  REPEAT
    display logging position;
    WAIT for logging command;
    IF move command
      THEN process move
    ELSE IF in auto log
      THEN
        BEGIN
          access machine code routine to read meter;
          update RAM file;
        END
      ELSE { in manual log }
        update RAM file from keyboard input
    UNTIL logging complete;
    save_grid
  END;
END;
```

statistics

```
BEGIN
  calculate descriptive statistics;
  IF errors
    THEN report errors
  ELSE
    BEGIN
      print statistics;
      calculate histogram;
      print histogram
    END
  END;
END;
```

dot_density

```
BEGIN
  input range values;
  FOR each quarter of grid DO
    BEGIN
      plot dot densities on screen;
      dump screen to printer
    END
  END;
END;
```

```

interval_codes
BEGIN
    input range values;
    print interval codes
END;

write_file
BEGIN
    write file to next free space on tape;
    IF no errors
        THEN
            BEGIN
                update catalogue;
                write catalogue
            END
        ELSE
            report errors
    END;

station_edit
BEGIN
    display grid;
    REPEAT
        select station;
        edit station
    UNTIL end of edit
END;

grid_edit
BEGIN
    input scaling constant;
    add to RAM file
END;

transfer
BEGIN
    select computer;
    read catalogue;
    FOR each file DO
        IF file selected for transfer
            THEN transfer
    END;

```

A.4
Program outline for apgrid

```
apgrid
BEGIN
  clear the density matrix;
  clear the complexity matrix;
  FOR each quadrat DO
    calculate_density_complexity;
  write out density and complexity matrices
END.

calculate_density_complexity
BEGIN
  read number of features;
  clear pixel mesh;
  FOR each feature DO
    BEGIN
      read in feature data;
      polygon clip and determine seed points;
      pixel convert outlines using simple digital
      differential analyser
    END;
  complexity equals proportion of pixels set;
  flood fill mesh from seed pixels;
  density equals proportion of pixels set
END;
```

A.5
Program outline for f extract

f_extract

BEGIN

read data grid;

set array of station markers to not used;

read significance level and background level;

search_grid for high valued features;

search_grid for low valued features;

FOR each station marker DO

IF not used

THEN set corresponding result to undisturbed;

write results

END.

search_grid

BEGIN

FOR each possible seed DO

IF not used

THEN

IF t-test significant

THEN

BEGIN

set seed station markers to used;

set corresponding results;

group_stations

END

END;

group_stations

BEGIN

IF station at x,y not used

THEN

BEGIN

add station to group;

IF t-test significant

THEN

BEGIN

group_stations at x,y+1;

group_stations at x+1,y+1;

group_stations at x+1,y;

group_stations at x+1,y-1;

group_stations at x,y-1;

group_stations at x-1,y-1;

group_stations at x-1,y;

group_stations at x-1,y+1

END

ELSE remove station from grid

END

END;

A.6
Program outline for local a map

```
local_a_map
BEGIN
  read operand 1;
  read operand 2;
  read cut off distance;
  FOR each grid point DO
    BEGIN
      least_squares_quadratic_fit(operand 1);
      write coefficients to scratch file 1;
      least_squares_quadratic_fit(operand 2);
      write coefficients to scratch file 2;
      calculate taxonomic distance;
      store taxonomic distance in matrix
    END;
  write out taxonomic distance matrix;
  FOR 1000 simulations DO
    BEGIN
      choose set of coefficients at random from scratch
      file 1;
      choose set of coefficients at random from scratch
      file 2;
      calculate taxonomic distance;
      add taxonomic distance to distribution
    END;
  calculate mean and standard deviation of distribution;
  standardise taxonomic distance matrix;
  write out standardised taxonomic distance matrix
END.
```


A.7
Program outline for k nna

```
k_nna
BEGIN
  read in standard point data file;
  select type A data;
  select type B data;
  read in maximum neighbour for analysis;
  check_kmax;
  calculate nna_stats for A-A distances;
  calculate nna_stats for A-B distances;
  calculate nna_stats for B-B distances;
  calculate nna_stats for B-A distances;
  plot_graphs
END.

check_kmax
BEGIN
  calculate centroid of analysis area;
  calculate shortest distance from centroid to boundary;
  calculate complexity limit;
  IF calculated maxk < kmax given
    THEN kmax := maxk
END;

nna_stats
BEGIN
  FOR k = 1 TO kmax DO
    BEGIN
      calculate observed mean k'th nearest neighbour
      distance;
      calculate expected mean k'th nearest neighbour
      distance using new correction;
      calculate sampling variance using new correction
    END
  END;
END;

plot_graphs
BEGIN
  plot distribution map;
  plot complexity range and data titles;
  plot_means for A-A distances;
  plot_means for B-A distances;
  plot_means for B-B distances;
  plot_means for A-B distances
END;
```

```
plot_means
BEGIN
  plot observed means;
  plot expected means;
  plot 5% significance levels;
  plot 1% significance levels
END;
```

APPENDIX B
BAGS USER MANUAL

BRADFORD
ARCHAEOLOGICAL
GEOPHYSICS
SYSTEM
USER MANUAL

RELEASE 1.1

M. A. KELLY 1984

The software this document describes and the document itself is
the copyright of:

M. A. KELLY

School of Archaeological Sciences,

University of Bradford,

Richmond Rd,

BRADFORD,

West Yorkshire,

BD7 1DP,

ENGLAND.

The documentation and software is freely available to members of
the archaeological community under the following conditions:

1) Software and documentation are obtained directly from
the author. This ensures access to the latest
release number.

2) Any use of the software or documentation presented in a
published form is duly acknowledged.

The author admits no responsibility for the consequences of any
errors in the software or documentation.

Software and documentation copies to:

Signed:

1.0 INTRODUCTION

This document describes the use of the software package BAGS (Bradford Archaeological Geophysics System). The package is intended primarily for the data management and contour plotting of geophysical data. The use of the package can however be extended to cover data grids of any origin.

The package is implemented on the C.D.C. Cyber 170-720 mainframe at the University of Bradford Computer Centre. Please ensure you are familiar with the use of this computer before reading further.

2.0 RUNNING THE PACKAGE

To run BAGS follow these instructions:

- 1) Log on to the mainframe.
- 2) Issue the following NOS command:

OLD,BAGS/UN=PYS400 [RETURN]

- 3) BAGS is now ready to run. To run it type:

BAGS [RETURN]

When BAGS has finished the following message will appear:

REVERT. BAGS FINISHED.

BAGS can be run again by repeating instruction 3.

BAGS is a menu driven package. It will present you with menus that contain a series of options labelled with numbers. You must select an option and indicate your choice by typing the option number followed by [RETURN]. In other contexts BAGS will ask you for numbers or text as input. You should type these in followed by [RETURN]. In response to any input prompt you may press [RETURN] on its own. This will terminate BAGS.

BAGS creates two files called:

XXXCON

REPORT

BRADFORD ARCHAEOLOGICAL GEOPHYSICS SYSTEM USER MANUAL.

which are saved in your store. These contain information only of interest to BAGS and should therefore be left alone.

3.0 MAIN MENU

When you first enter BAGS you will see the following menu:

BRADFORD ARCHAEOLOGICAL GEOPHYSICS SYSTEM.

MENU:

STORE DATA 0

PROCESS DATA 1

SELECT OPTION ?

If you type 0 [RETURN], you will select the store data option. If you type 1 [RETURN], you will select the process data option. If you type [RETURN], you will terminate BAGS. Any other input will be ignored.

The store data option is used to store data grids permanently on the computer. Once stored these grids may be processed by the process data option.

4.0 STORE DATA OPTION

The following menu will appear for the store data option:

DATA STORAGE.

MENU:

CATALOGUE ENTRY 0

INDEX FILE ENTRY 1

KEYBOARD DATA FILE ENTRY 2

SELECT OPTION ?

If you type 0 [RETURN], you will select the catalogue entry option. If you type 1 [RETURN], you will select the index file entry option. If you type 2 [RETURN], you will select the keyboard data file entry option. If you type [RETURN], you will terminate BAGS. Any other input will be ignored.

Data grids are formed by a block of 20 X 20 readings. A group of data grids comprises a site. There must be an index file for each site. This contains the names of all the associated data grids and their positions relative to one another. The site must finally be entered into the catalogue. This contains a list of all the sites accessible to BAGS.

4.1 CATALOGUE ENTRY OPTION

The following will appear for the catalogue entry option:

INDEX FILE NAME ?

USER CODE ?

DESCRIPTION ??

??

Type in the index file name followed by [RETURN]. This can be composed of letters or digits. It must start with a letter and not exceed seven characters in length.

Type in the user code for the store in which the index file is saved (normally this will be your own user code), followed by [RETURN].

Type in a description of the site between the two sets of question marks, followed by [RETURN]. E.g. location, date etc.... This description appears only in the BAGS catalogue.

You may press [RETURN] in response to any of the questions, this will terminate BAGS.

If you do not terminate BAGS you will be returned to the store data option menu on completion of the questions.

4.2 INDEX FILE ENTRY OPTION

The following will appear for the index file entry option:

INDEX FILE NAME ?

SITE DESCRIPTION ??

??

STATION INTERVAL (METRES) ?

ORIENTATION (DEGREES) ?

Type in the index file name followed by [RETURN] (see CATALOGUE ENTRY OPTION).

Type in a description of the site between the two sets of question marks, followed by [RETURN] (see CATALOGUE ENTRY OPTION). This description appears on any output produced by BAGS for the site.

Type in the distance between grid points (station interval) in metres, followed by [RETURN]. The distance will be truncated to integer metres. If the result is less than 1 metre or greater than 100 metres the distance will be rejected. Non-numeric input will also be rejected.

Type in the angle of geographical north with respect to site arbitrary north (orientation) in degrees, followed by [RETURN]. The angle should be measured in a clockwise direction. The orientation is truncated to integer degrees. If the result is less than 0 degrees or greater than 360 degrees the orientation will be rejected. Non-numeric input will also be rejected.

BRADFORD ARCHAEOLOGICAL GEOPHYSICS SYSTEM USER MANUAL.

You may press [RETURN] in response to any of the questions, this will terminate BAGS.

If you do not terminate BAGS the following will appear:

INDEX GRID LIMITS:

XMIN (1-20) ?

XMAX (1-20) ?

YMIN (1-20) ?

YMAX (1-20) ?

The data grids within a site will be arranged in a pattern. An example is shown below where # represents a grid.

```
  #  
  # #  
  # #
```

This pattern needs to be described to BAGS. To do this image imposing a mesh onto the pattern of grids and numbering each mesh cell with x and y coordinates.

```
  3 #  
y 2 # #  
  1 # #  
    1 2  
    x
```

The mesh limits i.e. the minimum and maximum values of x and y determine the size and shape of the mesh. These values can in fact

BRADFORD ARCHAEOLOGICAL GEOPHYSICS SYSTEM USER MANUAL.

range from 1 to 20. This gives a minimum mesh size of 1 data grid and a maximum mesh size of 400 data grids.

Type in the minimum and maximum mesh limits for the x and y coordinates, followed by [RETURN]s as indicated. Any values that are outside the input range will be rejected. Non-numeric input will also be rejected.

You may press [RETURN] in response to any of the questions, this will terminate BAGS.

If you do not terminate BAGS the following will appear:

INPUT \$ TO SKIP OR TO USE LAST USER CODE

X 1

Y 1

FILE NAME ?

USER CODE ?

For each cell of the site mesh you have defined, you will be offered the opportunity of assigning a data grid file name. Each mesh cell is identified by its x and y coordinates which are shown on the display. The display will be repeated for each possible mesh cell.

Type the data grid file name followed by [RETURN]. In some instances there will be no data grid to associate with the site mesh cell. To indicate this type the \$ character, followed by [RETURN].

Type the user code of the store in which the data grid is saved

(normally this will be your own user code), followed by [RETURN]. As data grids for a site are usually saved under the same user code it can become tedious to keep typing this many times. You can type the \$ character, followed by [RETURN]. This uses the last user code typed in. It is important when using this facility that you initially do type a user code. If you do not and try to use the \$ facility your input will be rejected.

You may press [RETURN] in response to any of the questions, this will terminate BAGS.

If you do not terminate BAGS then an index file with the file name given will be created and saved in your store. It is important therefore that the file name does not coincide with a file already saved in your store. You will then be returned to the store data option menu.

4.3 KEYBOARD DATA FILE ENTRY OPTION

Data grids may be entered by typing them at a terminal using this option. It is probably more efficient however to record the data grids on the Epson HX-20 system (see Resistance Survey Computer System User Manual) and transfer them to the mainframe.

The following will appear for the keyboard data file entry option:

DATA FILE NAME

?

Type in the data grid file name followed by [RETURN]. The file name must follow the same rules as for index files (see CATALOGUE ENTRY OPTION). You may press [RETURN] only, this will terminate BAGS.

If you do not terminate BAGS the following will appear:

INPUT DATA

?

A prompt will appear for each row of data to be entered. The prompts start at row one. Type in the row of readings making sure each reading is separated by at least one space. Twenty readings must be entered for each row even if there are null readings. Null readings should be indicated by the value -900 (this is consistent with the Epson HX-20 system). When you have typed a row press [RETURN]. The next prompt will appear. You may press [RETURN] in response to any of the prompts, this will terminate BAGS.

If you do not terminate BAGS the data grid file will be created and saved in your store. It is important therefore that the file name does not coincide with a file already saved in your store. You will then be returned to the store data option menu.

5.0 PROCESS DATA OPTION

The following menu will appear for the process data option:

DATA CATALOGUE INTERROGATION.

#####

#####

MENU:

VIEW CATALOGUE 0

SELECT INDEX FILE 1

PRINT CATALOGUE 2

SELECT OPTION ?

If you type 0 [RETURN], you will select the view catalogue option. If you type 1 [RETURN], you will select the select index file option. If you type 2 [RETURN], you will select the print catalogue option. If you type [RETURN], you will terminate BAGS. Any other input will be ignored.

5.1 VIEW CATALOGUE OPTION

The BAGS catalogue keeps an alphabetical list of all the sites, as described by index files, accessible to the package. The view

catalogue option will initially display the first page of this catalogue. Successive uses of this option will cycle through the catalogue pages in sequence. After the last page the sequence will be repeated. The information displayed consists of three columns:

index file name : user code : site description

5.2 SELECT INDEX FILE OPTION

The following will appear for the select index file option:

INDEX FILE NAME ?

Type in the site index file name you wish to process followed by [RETURN]. If the file name is not recognised you will be returned to the process data option menu. You may press [RETURN] only, this will terminate BAGS.

BRADFORD ARCHAEOLOGICAL GEOPHYSICS SYSTEM USER MANUAL.

If you do not terminate BAGS the following menu will appear:

#####

TEST PYR337 THIS IS A TEST SITE.

#####

MENU:

USE INDEX FILE 0

PRINT INDEX FILE 1

SELECT OPTION ?

If you type 0 [RETURN], you will select the use index file option. If you type 1 [RETURN], you will select the print index file option. If you type [RETURN], you will terminate BAGS. Any other input will be ignored.

The catalogue entry for the site you have selected is displayed above the menu.

5.2.1 USE INDEX FILE OPTION

The following menu will appear for the use index file option:

FILE AND PROCESSING SELECTION.

3 4 5

1 2

SELECT FILE [ALL:A END:E] ?

A plan is drawn of the data grids within the site selected. Each grid is given a number to identify it from the rest. If you wish to process all the data grids within the site type A [RETURN]. If you only want to process a selection of the data grids then type a data grid number followed by [RETURN]. The data grid number will be overwritten on the display. Repeat the selection of data grids until you have identified all the data grids you wish to process. Now type E [RETURN]. If you type a data grid number that is not displayed, or one that you have already selected, your input will be rejected. Non-numeric input will also be rejected. If you press [RETURN] only, you will terminate BAGS.

If you do not terminate BAGS the following menu will appear:

MENU:

PRINT REPORT FILES 0

CONTOUR 1

FILTERING 2

SELECT OPTIONS [E:END] ?

The data grids you have selected for processing are now called report files. Report files may be processed by any combination of options as displayed above. The processing options may therefore be chained together in a sequence, each process option working on the results of the previous process option. To specify the sequence type the option number for the first process followed by [RETURN]. Repeat for the other options in the sequence (if any). When complete type E [RETURN]. The typing of non-existent option numbers or non-numeric input will be rejected. You may press [RETURN] only, this will terminate BAGS.

Note: BAGS is designed for the easy incorporation of new processing options. These may be developed totally independently of BAGS and incorporated later.

5.2.1.1 Print report files option

The report files can be copied to a line printer using this option. The line printer used is the default one at the computer centre.

5.2.1.2 Contour option

The following will appear for the contour option:

CONTOURING.

RELATIVE SCALE [≤ 1] ?

NUMBER OF CONTOURS [≤ 20] ?

CONTOURS START AT ?

CONTOUR INTERVAL ?

Type the relative scale of the contour plot followed by [RETURN]. This should be less than or equal to 1. If it is greater than 1, less than or equal to 0 or non-numeric it will be rejected. Relative scale refers to the maximum sized plot that can be obtained on a particular plotting device. A value of one will therefore give you the largest possible plot on a device.

Type the number of contours wanted followed by [RETURN]. Input will be truncated to an integer. If the result is greater than 20 or less than 1 it will be rejected. Non-numeric input will also be rejected.

Type the value at which the contours will start, followed by [RETURN]. Non-numeric input will be rejected.

Type the interval over which contours will be plotted, followed by [RETURN]. Non-numeric input will be rejected.

You may press [RETURN] in response to any of the questions, this

will terminate BAGS.

If you do not terminate BAGS the following will appear:

COLOUR (ON/OFF) ?

NUMBERING (ON/OFF) ?

GRID LINES (ON/OFF) ?

Type either ON [RETURN] or OFF [RETURN] to plot contours in colour or black and white. Any other input will be rejected.

To plot identification numbers on the contours type ON [RETURN], otherwise type OFF [RETURN]. Any other input will be rejected.

Type ON [RETURN] to plot grid lines between data grids, otherwise type OFF [RETURN]. Any other input will be rejected.

You may press [RETURN] in response to any of the questions, this will terminate BAGS.

BRADFORD ARCHAEOLOGICAL GEOPHYSICS SYSTEM USER MANUAL.

If you do not terminate BAGS the following menu will appear:

PLOTTING DEVICES SELECTION.

MENU:

LINE PRINTER 0
GRAPHICS TERMINAL 1
CALCOMP NARROW PLOT 2
CALCOMP SPECIAL NARROW PLOT 3
CALCOMP WIDE PLOT 4
CALCOMP SPECIAL WIDE PLOT 5

SELECT OPTIONS [END:E] ?

You can select one or several plotting devices on which the contour plots will be produced. Type the device number followed by [RETURN]. Repeat this for as many devices as required. When finished type E [RETURN]. Any other input will be rejected. You may press [RETURN] only, this will terminate BAGS.

If you do not terminate BAGS the following will appear:

INTERACTIVE (I) OR BATCH (B) JOB ?

Type I [RETURN] to produce contour plots interactively. Type B [RETURN] to produce the contour plots in batch mode. Any other input will be rejected. You may press [RETURN] only, this will terminate BAGS.

If you are contouring a large number of data grids it will probably be more efficient to use batch mode. If you have specified the graphics terminal as the plotting device you must use

interactive mode.

5.2.1.3 Filtering option

The filtering option uses fourier transforms to filter signals of undesirable wavelength from your data. Refer to external documentation to use this option.

5.2.2 PRINT INDEX FILE OPTION

The contents of the site index file can be copied to a line printer using this option. The line printer used is the default one at the main computer centre. On completion of this option BAGS will be terminated.

5.3 PRINT CATALOGUE OPTION

The whole of the BAGS catalogue can be copied to a line printer using this option. The line printer used is the default one at the main computer centre. On completion of this option BAGS will be terminated.

APPENDIX C
RSCS USER MANUAL

RESISTANCE SURVEY

COMPUTER SYSTEM

USER MANUAL

RELEASE 1.2

M. A. KELLY 1984

The software this document describes and the document itself is
the copyright of:

M. A. KELLY

School of Archaeological Sciences,

University of Bradford,

Richmond Rd,

BRADFORD,

West Yorkshire,

BD7 1DP,

ENGLAND.

The documentation and software is freely available to members of
the archaeological community under the following conditions:

1) Software and documentation are obtained directly from
the author. This ensures access to the latest
release number.

2) Any use of the software or documentation presented in a
published form is duly acknowledged.

The author admits no responsibility for the consequences of any
errors in the software or documentation.

Software and documentation copies to:

Signed:

1.0 INTRODUCTION

This document describes the use of a computer system for the logging and processing of resistance survey data. The system is based on an Epson HX-20 portable computer and is intended to log readings from a '2-probe' survey configuration. Please ensure you are familiar with the operation of the survey rig and the HX-20 before reading further.

The two most important characteristics of the system are its portability and automatic logging facility. These allow the system to be used at any of four levels of increasing sophistication:

- 1) To process data already logged.
- 2) To manually log data from record sheets and process them.
- 3) To manually log data directly in the field and process them.
- 4) To automatically log data in the field and process them.

Resistance survey data is recorded in the form of grids of 20 X 20 readings. Each grid should have a unique name to identify it within the survey. The computer system can only work on one grid at any time, but many grids can be stored on the microcassette drive. Each side of a microcassette can in fact store up to 10 grids (giving a total of 20 grids per microcassette). As well as storing the grids the computer also stores a catalogue at the beginning of each microcassette side. This contains the names of the grids stored and helps the system find a particular grid when

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

needed. As you will see later the system is controlled by commands which are given by pressing the function keys on the computer.

There is a command CSET (Catalogue SET) that tells the computer to set up an empty catalogue of grid names on a microcassette side. This catalogue is used by the two commands SAVE and LOAD which respectively transfer a grid from the computer to the microcassette drive, and transfer a grid from the microcassette drive back to the computer. There is also a command CLST (Catalogue LiST) which lists on the printer the contents of the catalogue i.e. the grids stored on a microcassette side.

There are two logging commands available ALOG (Automatic LOGging) and MLOG (Manual LOGging). With automatic logging the resistance readings are directly captured from the resistance meter. With manual logging the resistance readings must be typed in using the keyboard. If readings are incorrectly recorded they can be corrected using the EDIT (EDITor) command.

The resistance readings within a grid can be listed on the printer using the HCOPY (Hard COPY) command. A grid can also be processed to produce a dot density display using the DISP (DISPlay) command.

Please read the next two sections carefully. They tell you how to load the program and run it. If you have any problems with the program then simply reload it. Next read section 4.0. This briefly explains how the system is used. You should then read the more detailed descriptions of the individual commands contained in section 5.0 onwards.

2.0 LOADING THE PROGRAM

Once loaded into the HX-20, the program will remain in memory even when the computer is switched off. To load the program follow these instructions:

- 1) Switch the computer on.
- 2) Make sure the computer does not contain any programs or data you do not want to lose.
- 3) Press the menu key.

```
CTRL/@ Initialize
1 MONITOR
2 BASIC
```

- 4) Press the CTRL key and the @ key at the same time.

```
Enter DATE and TIME
MMDDYYHHMMSSCR
=
Press BREAK to abort
```

- 5) Enter the month, day, year, hour, minute and second followed by RETURN.

```
Enter DATE and TIME
MMDDYYHHMMSSCR
=062084085430          [RETURN]
Press BREAK to abort
```

```
CTRL/@ Initialize
1 MONITOR
2 BASIC
```

6) Press 2.

Copyright 1982 by
Microsoft & EPSON
Pl: 0 Bytes
>

7) Insert the program microcassette.

8) Type: WIND [RETURN]

9) When the microcassette has been rewound type: LOAD [RETURN]

10) When the program has been loaded type: RUN [RETURN]

11) Follow the printed instructions:

KEY IN:
'K2RUN' RETURN CTRL@
THEN:
'B' RETURN

The rest of the program will then be loaded. This will take several minutes.

12) Switch the computer off.

13) Switch the computer on. The program should now automatically run. Attach the labels for the function keys.

3.0 RUNNING THE PROGRAM

To run the program switch the computer on. After a few seconds you will see:

```
Resistance survey
computer system.
M.A.Kelly 1984.
```

followed by a tone and:

```
Select command.
```

```
Break In 75
```

```
>
```

The program is now waiting for a command. A command is issued by pressing one of the function keys at the top of the keyboard. They should be labelled as follows:

```
MLOG  SAVE  CSET  EDIT  UTIL
```

```
====  ====  ====  ====  =====
```

```
ALOG  LOAD  CLST  DISP  HCOPY
```

To select a command on the bottom row just press the appropriate function key. To select a command on the top row press SHIFT and the appropriate function key at the same time.

4.0 USING THE SYSTEM

The following is a brief description of how the system may be used to log and process resistance survey data. The individual commands are described in greater detail in the following sections.

Any microcassettes that are to be used to store survey grids must be initialised by the CSET (Catalogue SET) command. It is probably a good idea to initialise a whole batch of microcassettes at once and keep them as a stock of ready data microcassettes. Data grids may be logged either manually or automatically using the MLOG or ALOG commands. Once logged a grid is saved automatically on the microcassette drive. Any grid may be transferred back to the computer from the microcassette drive by using the LOAD command. Any mistakes made and not corrected when logging the data can be corrected using the EDIT (EDITor) command. The EDIT command also allows you to scale a whole data grid either up or down for grid matching purposes.

The CLST (Catalogue LIST) command provides a useful contents list that can be kept with the microcassette. A printout of the raw data can be obtained by the HCOPY (Hard COPY) command. The DISP (DISPlay) command allows you to process a grid. Descriptive statistics can be produced as well as dot density and interval code displays, all copied automatically onto the printer.

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

Finally, data grids may be transferred using the UTIL (UTILities) command from the HX-20 to other computers where more sophisticated processing may be carried out.

5.0 CSET (CATALOGUE SET)

This command sets up an empty catalogue on a microcassette side. This must be done before any grids can be stored on a microcassette side.

Press CSET.

Set up catalogue.
If correct cassette
then press any key
else press RETURN.

If you wish to abort then press RETURN. Otherwise insert your data microcassette. Make sure that you have inserted the right microcassette and have selected the right side. If there are any programs or data on the microcassette, they will be lost when the catalogue is set up. When ready press any key other than RETURN to set up the catalogue.

ERRORS.

1) High tone with message:

then press any key
else press RETURN.
IO Error In 2505
>

There is no microcassette in the drive. Place an appropriate microcassette in the drive and press CSET again.

6.0 SAVE

This command saves a grid from the computer to the microcassette drive (this command is automatically called at the end of ALOG and MLOG).

Press SAVE.

Save grid.

Grid name? _

If you wish to abort then press RETURN. Otherwise type in a grid name followed by RETURN. The grid name may be composed of any characters but it is sensible only to use letters and digits and to start with a letter. Although any length of name can be typed only the first eight characters will be used.

ERRORS.

1) Microcassette drive continuously winds.

There is either no microcassette in the drive or a microcassette without a catalogue. Press BREAK.

Grid name? FIGG2
Abort In 2005
>

Place an appropriate microcassette in the drive and press SAVE again.

2) Error printout.

```
-----  
                ** ERROR **  
  
Grid name FIG2  
already used.  
-----
```

Press SAVE again and change the grid name.

3) Error printout.

```
-----  
                ** ERROR **  
  
Catalogue full.  
-----
```

Select a new microcassette side with a non-full catalogue. Press
SAVE again.

7.0 LOAD

This command loads a grid from the microcassette drive to the computer.

Press LOAD.

Load grid.

Grid name? _

If you wish to abort then press RETURN. Otherwise type in the grid name followed by RETURN.

ERRORS.

1) Microcassette drive continuously winds.

There is either no microcassette in the drive or a microcassette without a catalogue. Press BREAK.

```
Grid name? FIG2
Abort In 2005
>
```

Place an appropriate microcassette in the drive and press LOAD again.

2) Error printout.

```
-----
** ERROR **
Grid FIG3
not found.
-----
```

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

Press LOAD again. Make sure you have the right microcassette in the drive and / or spelt the grid name correctly.

8.0 CLST (CATALOGUE LIST)

This command lists the catalogue of the microcassette side in the microcassette drive onto the printer.

Press CLST.

List catalogue.

Printout:

Cassette Side	
1	HS11
2	HS12
3	HS13
4	HS14
5	HS15
6	HS16
7	HS17
8	HS18
9	HS19
10	HS20

ERRORS.

1) Microcassette drive continuously winds.

There is either no microcassette in the drive or a microcassette without a catalogue. Press BREAK.

List catalogue.

Abort In 2005

>

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

Place an appropriate microcassette in the drive and press CLST again.

9.0 EDIT (EDITOR)

This command calls the editor. The current grid in the computer is edited.

Press EDIT.

```
Editor.  
1-station edit  
2-grid edit
```

If you wish to abort then press RETURN. Station edit allows you to edit individual station readings. If you want station edit press 1. Grid edit allows you to add or subtract a constant value to all of the readings in the current grid. If you want grid edit press 2.

9.1 STATION EDIT

A picture of a grid is drawn. The columns and rows 5, 10, 15 and 20 are indicated by markers on the sides of the grid. Stations that contain a reading are filled, stations that do not contain a reading (null stations) are left blank.

```
#### r= 1 c= 1  
#### r old 187ohms  
#### new ohms  
c
```

A small cursor can be seen flashing in the top left of the grid. The position of the cursor is indicated by row and column at the top of the display. The present or old reading at that station is indicated in the middle of the display.

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

To edit a station reading, move the cursor to the station. This is done with the cursor keys which are located just above the RETURN key.

SHIFT + <^ key moves the cursor up

SHIFT + >v key moves the cursor down

<^ key moves the cursor left

>v key moves the cursor right

If you try to move the cursor beyond the edge of the grid you will simply move onto the opposite edge.

When you have reached the station you want to edit press the space bar.

```
#### r= 1 c= 5
#### r old 173ohms
#### new? ohms
c Edit mode
```

If you now press RETURN you will not alter the station reading, otherwise type the new reading followed by RETURN.

```
#### r= 1 c= 5
#### r old 173ohms
#### new? 110 ohms
c Edit mode [RETURN]
```

```
#### r= 1 c= 5
#### r old 110ohms
#### new ohms
c
```

If you type a negative number then the station reading will be set to a null reading. Typed readings are rounded to integer values. Non numeric characters give a station reading of 0.

To finish editing press RETURN.

9.2 GRID EDIT

Grid edit.

Value? -

If you wish to abort then press RETURN. Otherwise enter a constant value (positive or negative) followed by RETURN. This will then be added to all of the readings of the current grid. Non numeric characters give a constant value of 0.

10.0 HCOPY (HARD COPY)

This command prints the resistance readings for the current grid in the computer onto the printer.

Press HCOPI.

Hard copy.

A printout is shown below:

Grid HS11

1

.	.	153	173	173
.	.	170	157	153
177	173	156	152	159
177	170	142	156	150
178	168	164	157	159
149	174	144	173	168
182	190	157	165	173
166	151	156	164	190
168	133	166	172	176
173	165	161	173	197
174	165	163	179	175
167	175	155	171	177
189	177	165	181	181
199	179	166	175	178
163	165	169	179	187
95	194	157	163	178
192	193	169	176	179
176	178	173	159	185
178	168	157	164	184
163	185	176	185	219

6

178 149 167 141 175
171 177 153 165 153
146 180 145 150 157
154 166 165 149 137
152 170 153 167 146
173 171 165 137 149
178 184 162 152 134
170 193 180 153 156
185 199 168 164 149
170 173 172 163 137
171 187 190 154 161
178 191 169 141 138
185 186 215 147 164
188 187 181 161 111
204 182 154 146 181
177 184 177 159 185
182 180 165 188 169
191 197 178 173 142
181 199 185 171 167
194 199 185 162 153

11

168 167 148 162 173
163 146 151 167 175
142 150 144 165 174
147 143 146 173 172
161 156 152 168 172
158 149 133 173 175
157 161 146 161 179
158 165 168 173 177
150 156 149 176 184
150 164 157 177 179
157 156 160 183 175
163 160 157 179 177
169 160 167 179 183
95 155 169 174 169
172 162 167 169 179
183 173 167 177 170
153 162 157 175 175
166 160 169 185 175
183 185 167 181 168
189 165 169 170 174

16

178	167	173	173	159
173	178	181	179	160
175	168	169	163	156
163	179	170	173	168
162	171	167	177	161
163	173	152	162	145
160	181	169	167	169
164	173	162	165	165
173	170	165	156	162
169	185	183	166	161
173	173	173	171	164
165	184	177	173	155
170	180	181	192	161
172	182	181	186	171
141	177	181	178	160
166	167	172	164	162
169	162	171	176	160
145	185	177	190	170
153	189	169	179	161

The resistance readings are printed out in four blocks, the starting column number for each block is indicated. Null stations (i.e. stations without a reading) are displayed as dots. All other readings are displayed as integer ohms. The printout can be cut and assembled into a record sheet if desired.

11.0 DISP (DISPLAY)

This command allows the processing of the current grid in the computer.

Press DISP.

Display grid.
1-Statistics
2-Dot density
3-Interval codes

If you wish to abort then press RETURN. Statistics calculates and prints useful information about the current grid. If you want statistics press 1. Dot density produces a dot density display of the current grid. If you want dot density display press 2. Interval codes produces a simple but rapid display of the current grid. If you want interval codes display press 3.

11.1 STATISTICS

Statistics.

A printout is shown below:

Grid HS11
Statistics

Number of readings

400

Minimum value

95 ohms

Maximum value

224 ohms

Mean

172.795 ohms

Standard deviation

15.5462 ohms

Histogram of values in
the range mean +/- 2 s.d.

Frequency relative to
modal class frequency.

0 I**
1 I*****
2 I*****
3 I*****
4 I*****
5 I*****
6 I*****
7 I*****
8 I*****
9 I*

Mid class marks

0 144.812 ohms
1 151.05 ohms
2 157.249 ohms
3 163.467 ohms
4 169.686 ohms
5 175.904 ohms
6 182.123 ohms
7 188.341 ohms
8 194.56 ohms
9 200.778 ohms

ERRORS.

1) Error printout.

```
-----  
** ERROR **  
Insufficient data values  
-----
```

There are less than two readings in the grid, therefore statistics cannot be calculated.

11.2 DOT DENSITY

```
Dot density  
min ?_
```

If you wish to abort then press RETURN. Otherwise enter the minimum resistance value for the dot density display followed by RETURN.

```
Dot density  
min?164  
max?_
```

If you wish to abort then press RETURN. Otherwise enter the maximum resistance value for the dot density display followed by RETURN. For both minimum and maximum values if non numeric characters are typed a value of 0 will be used.

A dot density display is produced on the screen in four sections. These are copied onto the printer. The printer output may be cut (the middle white band removed) and assembled into a survey display.

11.3 INTERVAL CODES

Interval codes
min ?_

If you wish to abort then press RETURN. Otherwise enter the minimum resistance value for the interval codes display followed by RETURN.

Interval codes
min ?164
max ?_

If you wish to abort then press RETURN. Otherwise enter the maximum resistance value for the interval codes display followed by RETURN. For both minimum and maximum values if non numeric characters are typed a value of 0 will be used.

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

A printout is shown below:

```
-----  
Grid HS11  
Interval codes  
min= 164 ohms  
max= 190 ohms  
  
* 56 7 3 633 1573555  
 43 46 2 1 3657870  
75 87 256331  
643 22 5417453  
7312 4 3 0 3414370  
 45 53542 56105 1  
8* 257910 0 0708333  
2 01*4*8 235615122  
32 246**31 69542 1  
5205*4541 4567 8934  
521764** 0 09655541  
36 467*3 10 772975  
*6288*** 130379488*0  
*7267**80 353488*4  
1237**8 841337 7870  
 * 176977*953642*897  
**367882*3 1 6613411  
6755***75 2039631460  
73 198*94399383 96*4  
1*69****1 *2345 *370  
-----
```

Readings that are less than the minimum resistance value are left blank. Readings that are greater than or equal to the maximum resistance value are printed as blocks. Readings between the minimum and maximum resistance values are scaled from 0 to 9 and printed as such. The printer output can be cut and assembled as for dot densities.

12.0 MLOG (MANUAL LOGGING)

This command selects manual logging. Resistance readings must be typed in.

Press MLOG.

Manual logging.

The current grid is wiped clean. All station readings are set to null readings. It is important therefore that you should never enter MLOG if the current grid has not been saved on microcassette.

A picture of a grid is drawn. The columns and rows 5, 10, 15 and 20 are indicated by markers on the sides of the grid. Stations that contain a reading are filled, stations that do not contain a reading are left blank.

```
#### r=20 c= 1
#### r old      ohms
#### new       ohms
c      Mlog mode
```

A small cursor can be seen flashing in the bottom lefthand corner of the grid. The position of the cursor is indicated by row and column at the top of the display. The present or old reading at that station is indicated in the middle of the display. The cursor indicates the station at which a reading can be taken. The cursor will move in the form shown below:

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

```

                column
                1 2 3 4 . . . . . 20
1             ^ V ^ V
2             ^ V ^ V
3             ^ V ^ V
row 4         ^ V ^ V
.             . . . .
.             . . . .
20            ^ V ^ V
    
```

To record a reading at the present station type the reading followed by RETURN.

```

##### r=20 c= 1
##### r old      ohms
##### new? 67 ohms
c      Mlog mode           [RETURN]

##### r=19 c= 1
##### r old      ohms
##### new      ohms
c      Mlog mode
    
```

If you type a negative number, a null reading will be given to the station. If you type non numeric characters a value of 0 will be given to the station.

The cursor will automatically move on to the next station. To move forward without altering station values press RETURN. To move backwards without altering station values press the INS/DEL key. To move forward and set stations you pass to null readings press the HOME/CLR key.

You can use the RETURN key to skip unwanted stations. If you make a mistake and want to correct a station reading use the RETURN key or INS/DEL key to get to the station and then record the reading again.

When you have logged all the readings for the grid (or as many as

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

you intend to) press the TAB key, this will automatically call SAVE
(see 6.0 SAVE).

13.0 ALOG (AUTOMATIC LOGGING)

This command selects automatic logging. Resistance readings must be logged via the A.D.C. (analogue to digital converter).

Press ALOG.

Automatic logging.

The current grid is wiped clean. All station readings are set to null readings. It is important therefore that you should never enter ALOG if the current grid has not been saved on microcassette.

```
#### r=20 c= 1
#### r old ohms
#### new ohms
c Alog mode
```

The operation of ALOG is exactly the same as MLOG (see 12.0 MLOG) except ALOG has an automatic reading key. To automatically read the resistance meter press the SCRN key. The resistance reading will be displayed for a short interval.

To connect and operate the A.D.C. please refer to the operation manual supplied with it.

14.0 UTIL (UTILITIES)

This command allows access to various utilities. This is where new facilities may be added to the system.

Press UTIL.

Utilities.
1-transfer
2-
3-

Press the space bar.

4-
5-
6-
7-

Pressing the space bar again will bring you round to the first display. If you wish to abort then press RETURN. Transfer allows you to transfer data grids from the HX-20 to other computers. If you want transfer press 1.

14.1 TRANSFER

Transfer grids.
1-RML
2-HP
3-CYBER

This utility allows you to transfer data grids to:

- 1) RML 380-ZD microcomputer
- 2) HP minicomputer
- 3) CYBER 170-720 mainframe

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

If you wish to abort then press RETURN. Otherwise insert the microcassette with the grids you want to transfer, make sure it is the right microcassette and the right side. Select the appropriate computer.

ERRORS.

1) The microcassette drive continuously winds.

There is either no microcassette in the drive or the wrong microcassette. Press BREAK.

2-HP
3-CYBER
Abort In 2005
>

Place an appropriate microcassette in the drive and press UTIL again.

For each data grid on the microcassette side the following sequence will appear. Assume the HX-20 is transmitting data as soon as you enter transfer. The other computer should be listening from then onwards.

Transfer FIG2

Y/N

If you press N for no, the next data grid will be skipped to. If you press Y for yes the present data grid will be transmitted.

RESISTANCE SURVEY COMPUTER SYSTEM USER MANUAL.

Press any key
to transfer

Press any key.

TRANSFER ACTIVE

The next data grid will then be presented for transmission.
Transfer ends when all the data grids on a microcassette side have
been offered for transfer.

APPENDIX D
STANDARD POINT DATA FILE

STANDARD POINT DATA FILE (SPDF)

A standard point data file consists of a collection of two types of record:

- 1) Key records
- 2) Data records

A key record identifies the type of data records that follow it, until another key record is encountered. Every record has a key field, record positions 1 to 3, which identifies the record type.

1 INFORMATION

This is descriptive text information. The key record is:

```
INFORMATION
.....
1 3
```

The following data records contain free format strings of up to 20 characters. The key field must be left blank.

```
'THIS IS AN'
.....
1 3
```

```
'EXAMPLE'
.....
1 3
```

2 BOUNDARY

This is the x and y coordinates of the boundary points. The key record is:

```
BOUNDARY
.....
1 3
```

The following data records contain the x and y coordinates of the

boundary points as free format reals. The key field must be left blank. The boundary points should be listed in sequence as enumerated around the boundary line.

```
      4.56 2.18  
.....  
1 3
```

```
      3.74 18.94  
.....  
1 3
```

```
      2.11 2.17  
.....  
1 3
```

3 CODEBOOK

This relates integer codes to generic point type names. The key record is:

```
CODEBOOK  
.....  
1 3
```

The following data records contain free format strings of up to 10 characters defining point type names and associated free format integer codes. The key field must be left blank.

```
      'SHERDS' 1  
.....  
1 3
```

```
      'FLINTS' 2  
.....  
1 3
```

```
      'BONE' 17  
.....  
1 3
```

4 COORDINATES

This specifies the x and y coordinates of the point data. The key record is:

```
COORDINATES
.....
1 3
```

The following data records contain the x and y coordinates as free format reals, the generic code as a free format integer and an optional free format string of up to 10 characters that gives a specific point name. The key field must be left blank.

```
4.78 2.31 1 'POT73/1'
.....
1 3
```

```
3.21 7.11 1 'POT73/9'
.....
1 3
```

```
5.61 2.01 17 'JAW27/4'
.....
1 3
```

5 SET

This toggles an automatic generic code generating facility on and off. The key record is:

```
SET
.....
1 3
```

On the first call the next record should specify the generic code to be generated as a free format integer. For subsequent coordinate data records this generic code will be associated even if the records already contain generic codes of their own. This facility may be switched off by another call of the key record. This facility is global, i.e. operative throughout the whole file. To

set up a new generating generic code, the facility must be switched on again and the new generic code given.

6 ***

By inserting asterisks in the key field a data record is turned into a null record, i.e. a record that is skipped at input.

The key records and associated data records (called blocks) may be sequenced in any order. A specific block type may appear more than once in a file.

BIBLIOGRAPHY

- Adams, R.McC. and Nissen, H.J. (1972). The Uruk Countryside. Chicago: University of Chicago Press.
- Aitken, M.J. (1958). Magnetic prospecting I- The Water Newton Survey. *Archaeometry*, vol 1, p 24-29.
- Aitken, M.J. (1974). *Physics and Archaeology* (2nd edn). Oxford: Clarendon Press.
- Ammerman, A.J. and Feldman, M.W. (1978). Replicated collection of site surfaces. *American Antiquity*, vol 43, p 734-740.
- Andersen, J. (1969). A new technique for archaeological field measuring. *Norwegian Archaeological Review*, vol 2, p 68-75.
- Ash, R. (1965). *Information Theory*. New York: John Wiley and Sons.
- Aspinall, A. and Lynam, J.T. (1970). An induced polarisation instrument for the detection of near-surface features. *Prospezioni Archeologiche*, vol 5, p 67-75.
- Aspinall, A. and Walker, A.R. (1975). The earth resistivity instrument and its application to shallow earth surveys. *Underground Services*, vol 3, no 1, p 12-15.
- Aston, M. and Rowley, T. (1974). *Landscape Archaeology*. Newton Abbot: David and Charles.
- Baker, C.M. (1978). The size effect: an explanation of variability in surface artifact assemblage content. *American Antiquity*, vol 43, p 285-293.
- Barker, P. (1977). *The Techniques of Archaeological Excavation*. London: Batsford.
- Barnett, S. (1975). *Introduction to Mathematical Control Theory*. Oxford: Oxford University Press.
- Bastwood, T. (1953). *British Regional Geology, Northern England*. London: HMSO.
- Becker, H. (1978). Magnetic prospecting with a difference-proton magnetometer with an automatic data-record on a digital cassette. *Archaeo-Physika*, vol 10, p 633-637.
- Bezier, P. (1974). Mathematical and practical possibilities of UNISURF. In (R.E. Barnhill and R.F. Riesenfeld, Eds) *Computer Aided Geometric Design*. New York: Academic Press.
- Biddle, M. and Kjolbye-Biddle, B. (1969). Metres, areas and robbing. *World Archaeology*, vol 1, p 208-219.

- Binford, L.R. (1978). Dimensional analysis of behavior and site structure: learning from an Eskimo hunting stand. *American Antiquity*, vol 43, p 330-336.
- Bowers, P.M., Bonnichsen, R. and Hoch, D.M. (1983). Flake dispersal experiments: noncultural transformation of the archaeological record. *American Antiquity*, vol 48, p 553-572.
- Bowman, K.O. and Shenton, L.R. (1968). Properties of Estimators for the Gamma Distribution. Report CTC-1. Oak Ridge: Union Carbide Corp.
- Bradley, R. (1984). Sieving for soil marks. *Oxford Journal of Archaeology*, vol 3, p 71-76.
- Brodie, L. (1981). *Starting FORTH*. Englewood Cliffs: Prentice-Hall.
- Brodlie, K.W. (1980). A review of methods for curve and function drawing. In (K.W. Brodlie, Ed) *Mathematical Methods in Computer Graphics and Design*. London: Academic Press, p 1-37.
- Butland, J. (1980). A method of interpolating reasonably-shaped curves through any data. *Proceedings of Computer Graphics 80*. London: Online Publications, p 409-422.
- Butland, J. (1984). *SIMPLEPLOT Mark 2: Section 2. Plotting 3-Dimensional Data*. Bradford: Bradford University Research Ltd.
- CDC. (1982). *NOS Version 2: Guide to System Usage. Vol 2*. Sunnyvale: Control Data Corporation.
- Chamberlain, M.P. and Haigh, J.G.B. (1982). A microcomputer system for practical photogrammetry. *Computer Applications in Archaeology*. Birmingham: University of Birmingham, p 142-149.
- Clark, A.J. (1975). Archaeological prospecting: a progress report. *Journal of Archaeological Science*, vol 2, p 297-314.
- Clark, A.J. (1983). The testimony of the top soil. In (G.S. Maxwell, Ed) *The Impact of Aerial Reconnaissance on Archaeology*. CBA Research Report 49, p 128-135.
- Clark, G.A. (1976). More on contingency table analysis, decision making criteria and the use of log linear models. *American Antiquity*, vol 41, p 259-273.
- Clark, G.A., Effland, R.W. and Johnstone, J.C. (1977). Quantitative spatial analysis. *Computer Applications in Archaeology*. Birmingham: University of Birmingham, p 27-44.
- Clark, P.J. and Evans, F.C. (1954). Distance to nearest neighbour as a measure of spatial relationship in populations. *Ecology*, vol 35, p 443-453.
- Clarke, D.L. (1968). *Analytical Archaeology*. London: Methuen.

- Clarke, D.V. (1978). Excavation and volunteers: a cautionary tale. *World Archaeology*, vol 10, p 63-70.
- Cliff, A.D. and Ord, J.K. (1973). *Spatial Autocorrelation*. London: Pion.
- Coles, J. (1972). *Field Archaeology in Britain*. London: Methuen.
- Corder, P. (1928). *The Roman Pottery at Crambeck, Castle Howard*. York: Roman Malton and District Reprint no. 1.
- Corder, P. and Birley, M. (1937). Fourth century Romano-British kilns, near Crambeck. *Antiquaries Journal*, vol 1, p 398-413.
- Corder, P. and Kirk, J.L. (1932). *A Roman Villa at Langton, near Malton, E. Yorkshire*. Leeds: Roman Malton and District Reprint no. 5.
- Dacey, M.F. (1963). Order neighbour statistics for a class of random patterns in multidimensional space. *Annals of the Association of American Geographers*, vol 53, p 505-515.
- Dacey, M.F. (1973). Statistical tests of spatial association in the locations of tool types. *American Antiquity*, vol 38, p 320-328.
- Davis, J.C. (1973). *Statistics and Data Analysis in Geology*. New York: John Wiley and Sons.
- De Boor, C. (1978). *A Practical Guide to Splines*. New York: Springer-Verlag.
- De Vos, S. (1973). The use of nearest neighbor methods. *Tijdschrift voor Econ. en Soc. Geografie*, vol 64, p 307-319.
- Deuel, L. (1969). *Flights into Yesterday*. London: Macdonald.
- Dixon, W.J. and Massey, F.J. (1983). *Introduction to Statistical Analysis*. Tokyo: McGraw-Hill.
- Donnelly, K.P. (1978). Simulations to determine the variance and edge effect of total nearest-neighbour distance. In (I. Hodder, Ed) *Simulation Studies in Archaeology*. Cambridge: Cambridge University Press, p 91-95.
- Dunn, C.J. (1976). Ring ditches in the Derwent valley. Leeds: Yorkshire Philosophical Society, p 60.
- Dwight, H.B. (1961). *Tables of Integrals and Other Mathematical Data* (4th edn). New York: Macmillan.
- Earle, T.K. (1976). A nearest neighbor analysis of two formative settlement systems. In (K.V. Flannery, Ed) *The Early Mesoamerican Village*. New York: Academic Press, p 196-223.
- Ebdon, D. (1976). On the underestimation inherent in the commonly-used formulae. *Area*, vol 8, p 165-169.

- Evans, R. and Jones, R.J.A. (1977). Crop marks and soils at two archaeological sites in Britain. *Journal of Archaeological Science*, vol 4, p 63-76.
- Everitt, B.S. (1977). *The Analysis of Contingency Tables*. London: Chapman and Hall.
- Fisher, R.A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
- Foard, G. (1977). Systematic fieldwalking and the investigation of Saxon settlement in Northamptonshire. *World Archaeology*, vol 9, p 357-374.
- Foley, R. (1981). Off-site archaeology: an alternative approach for the short-sited. In (I. Hodder, G. Isacc and M. Hammond, Eds) *Pattern of the Past*. Cambridge: Cambridge University Press, p 157-183.
- Freeman, D.T. (1981). *On Site Visualisation and Data Processing Techniques of Survey Data in Archaeological Prospection*. Unpublished M.Phil. thesis. Bradford: University of Bradford.
- Frink, D.S. (1984). Artifact behavior within the plow zone. *Journal of Field Archaeology*, vol 11, p 356-363.
- Fuchs, C., Kaufman, D. and Ronen, A. (1977). Erosion and artifact distribution in open-air Epi-Paleolithic sites on the coastal plain of Israel. *Journal of Field Archaeology*, vol 7, p 171-179.
- Fulford, M.G. and Hodder, I.R. (1974). A regression analysis of some late Romano-British fine pottery: a case study. *Oxoniensia*, vol 39, p 26-33.
- Graham, I. (1980). Spectral analysis and distance methods in the study of archaeological distributions. *Journal of Archaeological Science*, vol 7, p 105-129.
- Green, P.G. and Sibson, R. (1978). Computing Dirichlet tessellations in the plane. *Computer Journal*, vol 21, p 168-173.
- Greenwood, J.A. and Durand, D. (1960). Aids for fitting the gamma distribution by maximum likelihood. *Technometrics*, vol 2, p 55-65.
- Greig-Smith, P. (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. *Annals of Botany*, vol 16, p 293-316.
- Greig-Smith, P. (1964). *Quantitative Plant Ecology* (2nd edn). London: Methuen.
- Griffiths, M. (1982). *Archaeological Record System*. Northallerton: North Yorkshire County Council Computer Services.
- Hammond, G. and Hammond, N. (1981). Child's play: a distorting factor in archaeological distribution. *American Antiquity*, vol 46, p 634-636.

- Hammond, N. (1974). The distribution of late Classic Maya major ceremonial centres in the central area. In (N. Hammond, Ed) Mesoamerican Archaeology New Approaches. Austin: University of Texas Press, p 313-334.
- Hawkes, J. (1968). The proper study of mankind. *Antiquity*, vol 42, p 255-262.
- Hayes, A. and Osborne, D. (1961). Fixing site locations by radio-direction finder at Mesa Verde. *American Antiquity*, vol 27, p 110-112.
- Heathcote, C. (1983). Applications of Magnetic and Pulsed Induction Methods to Geophysical Prospection at Shallow Depth: a Theoretical and Practical Study of the Responses of Some Survey Systems. Unpublished Ph.D. thesis. Bradford: University of Bradford.
- Heathcote, C. and Kelly, M.A. (1979). The Oakes Tiddington, a Geophysical Survey. Bradford geophysical reports. Bradford: University of Bradford.
- Hietala, H.J. and Larson, P.A. (1979). SYMAP analysis in archaeology: intrasite assumptions and a comparison with TREND analysis. *Norwegian Archaeological Review*, vol 12, p 57-64.
- Hinchliffe, J. and Schadla-Hall, R.T. (Eds) (1980). *The Past Under the Plough*. Occasional papers no 3. London: Department of the Environment.
- Hirth, K.G. (1978). Problems in data recovery and measurement in settlement archaeology. *Journal of Field Archaeology*, vol 5, p 125-131.
- Hodder, I. (1972). Locational models and the study of Romano-British settlement. In (D.L. Clarke, Ed) *Models in Archaeology*. London: Methuen, p 887-909.
- Hodder, I. and Hassall, M. (1971). The non-random spacing of Romano-British walled towns. *Man*, vol 6, p 391-407.
- Hodder, I. and Okell, E. (1978). An index for assessing the association between distributions of points in archaeology. In (I. Hodder, Ed) *Simulation Studies in Archaeology*. Cambridge: Cambridge University Press, p 97-107.
- Hodder, I. and Orton, C.R. (1976). *Spatial Analysis in Archaeology*. Cambridge: Cambridge University Press.
- Hodson, F.R., Sneath, P.H.A. and Doran, J.E. (1966). Some experiments in the numerical analysis of archaeological data. *Biometrika*, vol 53, p 311-324.
- Hogg, A.H.A. (1980). *Surveying for Archaeologists and Other Fieldworkers*. London: Croom Helm.

- Hotelling, H. (1940). The selection of variates for use in prediction with some comments on the general problem of nuisance parameters. *Annals of Mathematical Statistics*, vol 11, p 271-283.
- Houlder, P.J. (1983). Some Studies of the Underlying Theoretical Considerations Associated with Resistivity Surveying for Archaeology. Unpublished M.A. dissertation. Bradford: University of Bradford.
- Johnson, I. (1977). Local density analysis. *Computer Applications in Archaeology*. Birmingham: University of Birmingham, p 90-98.
- Johnson, N.L. and Kotz, S. (1970). Continuous Univariate Distributions-1. Boston: Houghton Mifflin Company.
- Kelly, M.A, Dale, P. and Haigh, J.G.B. (1984). A microcomputer system for data logging in geophysical surveying. *Archaeometry*, vol 26, p 183-191.
- Kelly, M.A. and Haigh, J.G.B. (1983). Determination of archaeological structure by spatial analysis and association. *Computer Applications in Archaeology*. Bradford: University of Bradford, p 181-188.
- Kelly, M.A. and Haigh, J.G.B. (1984). Automatic data logging for resistance surveying and subsequent data processing. *Computer Applications in Archaeology*. Birmingham: University of Birmingham, p 161-169.
- King, L.J. (1969). *Statistical Analysis in Geography*. Englewood Cliffs: Prentice-Hall.
- Kintigh, K.W. and Ammerman, A.J. (1982). Heuristic approaches to spatial analysis in archaeology. *American Antiquity*, vol 47, p 31-63.
- Koch, G.S. and Link, R.F. (1971). *Statistical Analysis of Geological Data. Volume II*. New York: John Wiley and Sons.
- Larson, P. (1975). Trend analysis in archaeology: a preliminary study of intrasite patterning. *Norwegian Archaeological Review*, vol 8, p 75-80.
- Lindeman, R.H., Merenda, P.F. and Gold, R.Z. (1980). *Introduction to Bivariate and Multivariate Analysis*. Glenview: Scott, Foresman and Company.
- Linnington, R.E. (1970). A first use of linear filtering techniques on archaeological prospecting results. *Prospezioni Archeologiche*, vol 5, p 43-54.
- Lynam, J.T. (1970). *Techniques of Geophysical Prospection as Applied to Near Surface Structure Determination*. Unpublished Ph.D. thesis. Bradford: University of Bradford.
- Mack, R.N. and Pyke, D.A. (1979). Mapping individual plants with a field-portable digitizer. *Ecology*, vol 60, p 459-461.

- Mandelbrot, B. (1982). *The Fractal Geometry of Nature*. New York: W.H. Freeman and Co.
- McConalogue, D.J. (1970). A quasi-intrinsic scheme for passing a smooth curve through a discrete set of points. *Computer Journal*, vol 13, p 392-396.
- McLain, D.H. (1974). Drawing contours from arbitrary data points. *Computer Journal*, vol 17, p 318-324.
- McNutt, C.H. (1981). Nearest neighbors, boundary effect, and the old flag trick: a general solution. *American Antiquity*, vol 46, p 571-592.
- Mead, R. (1974). A test for spatial pattern at several scales using data from a grid of contiguous quadrats. *Biometrics*, vol 30, p 295-307.
- Merriam, D.F. and Sneath, P.H.A. (1966). Quantitative comparison of contour maps. *Journal of Geophysical Research*, vol 71, p 1105-1115.
- Mood, A.M., Graybill, F.A. and Boes, D.C. (1974). *Introduction to the Theory of Statistics*. New York: McGraw-Hill.
- Moore, P.G. (1954). Spacing in plant populations. *Ecology*, vol 35, p 222-227.
- Mueller, J.W. (1974). The use of sampling in archaeological survey. *Society for American Archaeology, Memoirs* 28.
- NAG. (1981). *NAG Graphical Supplement: a Simple Plotting System for use with the NAG Fortran Library, Mark 1*. Oxford: Numerical Algorithms Group.
- NAG. (1982). *The NAG Fortran Library Manuals*. Oxford: Numerical Algorithms Group.
- Newman, W.M. and Sproull, R.F. (1981). *Principles of Interactive Computer Graphics*. Tokyo: McGraw-Hill.
- Noland, G.S. (1979). *The GHOST Graphical Output System (Fortran version)*. Manchester: University of Manchester Regional Computer Centre.
- O'Brien, M.J., Beets, J.L., Warren, R.E., Hotrabhavanda, T., Barney, T.W. and Voigt, E.E. (1982). Digital enhancement and grey-level slicing of aerial photographs: techniques for archaeological analysis of intrasite variability. *World Archaeology*, vol 14, p 171-190.
- Orton, C.R. (1975). Quantitative pottery studies some progress, problems and prospects. *Science and Archaeology*, vol 16, p 30-35.
- Orton, C.R. (1980). *Mathematics in Archaeology*. Cambridge: Cambridge University Press.

- Orton, C.R. (1982a). Computer simulation experiments to assess the performance of measures of quantity of pottery. *World Archaeology*, vol 14, p 1-20.
- Orton, C.R. (1982b). Stochastic process and archaeological mechanism in spatial analysis. *Journal of Archaeological Science*, vol 9, p 1-23.
- Palmer, R. (1976). A method of transcribing archaeological sites from oblique aerial photographs. *Journal of Archaeological Science*, vol 3, p 391-394.
- Palmer, R. (1977). A computer method for transcribing information graphically from oblique aerial photographs to maps. *Journal of Archaeological Science*, vol 4, p 283-290.
- Palmer, R. (1978). Aerial archaeology and sampling. In (J.F. Cherry, C. Gamble and S. Shennan, Eds) *Sampling in Contemporary British Archaeology*. *British Archaeological Reports*, vol 50, p 129-148.
- Pavlidis, T. (1982). *Algorithms for Graphics and Image Processing*. Berlin: Springer-Verlag.
- Perry, J.N. and Mead, R. (1979). On the power of the index of dispersion test to detect spatial pattern. *Biometrics*, vol 35, p 613-622.
- Pielou, E.C. (1961). Segregation and symmetry in two species populations as studied by nearest neighbor methods. *Journal of Ecology*, vol 49, p 255-269.
- Pielou, E.C. (1969). *An Introduction to Mathematical Ecology*. New York: Wiley Interscience.
- Pinder, S., Shimada, I. and Gregory, D. (1979). The nearest-neighbour statistic: archaeological application and new developments. *American Antiquity*, vol 44, p 430-445.
- Plog, F. (1974). Settlement patterns and social history. In (M.J. Leaf, Ed) *Frontiers of Anthropology*. New York: Van Nostrand, p 68-91.
- Plog, S. (1978). Sampling in archaeological surveys: a critique. *American Antiquity*, vol 43, p 280-285.
- Powlesland, D. (1982). *Heslerton Parish Project, 1982-92*. Yedingham: Powlesland.
- Rackerby, F. (1973). A statistical determination of the black sand occupation at the Macoupin site, Jersey Co., Illinois. *American Antiquity*, vol 38, p 96-101.
- Radley, J. (1974). The prehistory of the Vale of York. *Yorkshire Archaeological Journal*, vol 46, p 10-22.

- Redman, C.L. and Watson, P.J. (1970). Systematic, intensive surface collection. *American Antiquity*, vol 35, p 279-291.
- Ramm, H. (1978). *The Parisi*. London: Duckworth.
- Reichlin, S. (1974). A mechanical device for drawing excavation sections. *American Antiquity*, vol 39, p 373-375.
- Rich, E. (1983). *Artificial Intelligence*. Singapore: McGraw-Hill.
- Rick, J.W. (1976). Downslope movement and archaeological intrasite spatial analysis. *American Antiquity*, vol 41, p 133-144.
- Rogers, A.R. (1982). Data collection and information loss in the study of spatial pattern. *World Archaeology*, vol 14, p 249-258.
- Rolfsen, P. (1980). Disturbance of archaeological layers by processes in the soil. *Norwegian Archaeological Review*, vol 13, p 110-118.
- Roper, D.C. (1976). Lateral displacement of artifacts due to plowing. *American Antiquity*, vol 41, p 372-375.
- Rosenfeld, A. and Kak, A.C. (1982). *Digital Picture processing* (2nd edn). New York: Academic Press.
- Sabin, M.A. (1980). Contouring - a review of methods for scattered data. In (K.W. Brodlie, Ed) *Mathematical Methods in Computer Graphics and Design*. London: Academic Press, p 63-86.
- Schagen, J.P. (1982). Automatic contouring from scattered data points. *Computer Journal*, vol 25, p 7-11.
- Scollar, I. (1970). Fourier transform methods for the evaluation of magnetic maps. *Prospezioni Archeologiche*, vol 5, p 9-41.
- Scollar, I. (1975). Transformation of extreme oblique aerial photographs to maps or plans by conventional means or by computer. In (D.R. Wilson, Ed) *Aerial Reconnaissance for Archaeology*. CBA research report 12, p 52-59.
- Scollar, I. (1978). Computer image processing for archaeological air photographs. *World Archaeology*, vol 10, p 71-87.
- Scollar, I. and Kruckeberg, F. (1966). Computer treatment of magnetic measurements from archaeological sites. *Archaeometry*, vol 9, p 61-71.
- Simek, J.F. and Larick, R.R. (1983). The recognition of multiple spatial patterns: a case study from the French upper Paleolithic. *Journal of Archaeological Science*, vol 10, p 165-180.
- Sorensen, P.R. (1984). Fractals. *Byte*, vol 9, no 10, p 157-172.
- Sowerbutts, W.T.C. and Mason, R.W.I. (forthcoming). A microcomputer based system for small-scale land geophysical surveys.

- Star, J.L. (1985). Introduction to image processing. *Byte*, vol 10, no 2, p 163-170.
- Stead I.M. (1968). An Iron Age hillfort at Grimthorpe, Yorkshire, England. *Proceedings of the Prehistoric Society*, vol 34, p 148.
- Sterud, E.L. and Pratt, P.P. (1975). Archaeological intra-site recording with photography. *Journal of Field Archaeology*, vol 2, p 151-167.
- Sutcliffe, D.C. (1980). Contouring over rectangular and skewed rectangular grids - an introduction. In (K.W. Brodlie, Ed) *Mathematical Methods in Computer Graphics and Design*. London: Academic Press, p 39-62.
- Sutherland, I.E. and Hodgman, G.W. (1974). Reentrant polygon clipping. *Communications of the Association for Computing Machinery*, vol 17, p32.
- Taylor, T.P. (1979). Soil mark studies near Winchester, Hampshire. *Journal of Archaeological Science*, vol 6, p 93-100.
- Thompson, H.R. (1956). Distribution of distance to n-th neighbour in a population of randomly distributed individuals. *Ecology*, vol 37, p 391-394.
- Tolstoy, P. and Fish, J.K. (1975). Surface and subsurface evidence for community size at Coapexco, Mexico. *Journal of Field Archaeology*, vol 2, p 97-104.
- Trueman, A.E. (1967). *Geology and Scenery in England and Wales*. London: Penguin Books.
- Varley, W.J. (1968). Barmston and the Holderness Crannogs. *East Riding Archaeology*, vol 1, p 12.
- Washburn, D.K. (1974). Nearest neighbor analysis of Pueblo I-III settlement patterns along the Rio Puerco of the East, New Mexico. *American Antiquity*, vol 39, p 315-335.
- Whallon, R. (1974). Spatial analysis of occupation floors II, the application of nearest neighbour analysis. *American Antiquity*, vol 39, p 16-34.
- Wilcock, J.D. (1970). Some developments in the portrayal of magnetic anomalies by digital incremental plotter. *Prospezioni Archeologiche*, vol 5, p 55-58.
- Wilson, D.R. (1982). *Air Photo Interpretation for Archaeologists*. London: Batsford.
- Wobst, H.M. (1983). We can't see the forest for the trees: sampling and the shapes of archaeological distributions. In (J.A. Moore and A.S. Keene, Eds) *Archaeological Hammers and Theories*. New York: Academic Press, p 37-85.

Zahl, S. (1974). Application of the S-method to the analysis of spatial pattern. *Biometrics*, vol 30, p 513-524.

Zahl, S. (1977). A comparison of three methods for the analysis of spatial pattern. *Biometrics*, vol 33, p 681-692.

Zubrow, E. (1971). Carrying capacity and dynamic equilibrium in the prehistoric Southwest. *American Antiquity*, vol 36, p 127-138.