The automated generalisation of small scale topographic maps, with particular reference to the Ordnance Survey 1:50,000 scale maps

Smith, N. S.

Smith, N. S. (1980) *The automated generalisation of small scale topographic maps, with particular reference to the Ordnance Survey 1:50,000 scale maps*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/7435/

The automated generalisation of small scale topographic maps, with particular reference to the Ordnance Survey 1:50,000 scale maps

N. S. Smith, B.A., A.R.I.C.S

Dissertation submitted in partial
requirement for the degree of
Master of Science in Spatial Data
Analysis in Geography, 1979-80.

CONTENTS

Chapter 1 Introduction

Chapter 2 The elements of generalisation and the requirements of an automated system

Chapter 3 Ordnance Survey conventional small scale map series, their specifications and the methods used in their derivation

Chapter 4 Automated generalisation of the 1:50,000scale maps

Chapter 5 Experiments in generalising 1:50,000 scale map data

Chapter 5 continued.

CHAPTER 1

1.1 INTRODUCTION

In October 1979 the committee set up under the chairmanship of Sir David Serpell to review the Ordnance Survey presented its report to the Secretary of State for the Environment. One of the proposals put forward by the committee was:

> 'a......research and development programme would be set up (within Ordnance Survey)
>
> i) to resolve uncertainties about the detailed nature of user needs for smaller scale data and the appropriateness of existing Ordnance Survey digitising procedures, data structures and software, both for map production and for the provision of data suitable for other computer based analyses.
>
> ii) subject to the results of (i) to develop the appropriate software for the creation of a topographic database from the 1:50,000 series to facilitate the conversion of the 1:50,000 First Series to Second Series; and to produce derived mapping at smaller scales'.

Central to both (1) and (2) is the concept of using generalisation of data to produce subsets of the databases for specific purposes including the production of derived maps at smaller scales. The objective of this dissertation is to consider the processes involved in the manual generalisation of maps and to find ways in which these might be minimised by automated means in the context of the topographic map series at 1:50,000 and 1:250,000 scales produced by Ordnance Survey.

1.2 Automated Generalisation

The reasons for the Serpell Committee's proposals lie in the increasing awareness of the potential use of cartographic data in computer-readable form. Much of the growing body of literature on digital cartography is concerned with the computer science details - the data structure, algorithms etc. or the cartographic-capabilities, achievements and uses of various systems. A considerable proportion of the latter category is concerned with generalisation (Uitermark 1980) but in almost all cases, generalisation of map data is for the specific purpose of emulating manual generalising techniques, producing a graphic output on paper. Digital techniques are rarely used to produce a generalised digital map dataset other than plotter drive tapes.

Generalisation by automated methods is worthy of study for purposes other than that of producing maps. There is evidence (Royal Society 1978) to suggest that geographical information systems will be in widespread use in the near future. These information systems will probably include data overlaid on 'base map' data in a manner analogous to transparent paper overlays on a topographic map. How these systems might work is not considered here (see Rhind 1976) but this dissertation is concerned with the content of the 'base map' and how it is derived.

There is a school of thought that a 'scale free' topographic database (Royal Society 1978) would provide any information that is required by an accessing system, but this assumes no finite restrictions on computer processing power, response time or storage capacity. In conventional maps 'scale dependancy' results in part from the physical and perceptual limitations of humans in interpreting printed information on a map. In a computer system the constraints are not derived so much from resolution as from the limitations on facilities available and the cost of using them. In general these costs are related to data volume (and storage medium) and processing complexity, so digital representation of a map must therefore aim to minimise data quantity while retaining maximum information content.

1.3 Outline of work.

The work that this dissertation describes was broken down into three major parts. The first was to use published sources and consider the processes involved in generalisation of maps with regard to digital processing. Secondly, Ordnance Survey manuals and discussion with Ordnance Survey draughtsmen provided information for an analysis of the production of conventional 1:50,000 scale topographic maps and the maps derived from them. Last, some particular aspects of generalisation were considered in detail and experiments used to demonstrate the techniques, costs and associated benefits of undertaking the generalisation by automated, rather than manual methods.

A number of assumptions were made. The type of maps considered was topographical of a style in which lines are used to represent the majority of features. From this premise, the likely means of digital encoding was taken to be by vectors in which each individual line is represented by co-ordinates along its length (or by some means of generating such co-ordinates). The alternative method of encoding maps, by a raster array is more suited to certain aspects of map data manipulation (Peuquet 1979) but not with regard to many aspects of handling lines. Furthermore, the data available for experimentation were representing only linear objects (at 1:50,000 scale). For these reasons the emphasis is on the processing of vector encoded maps.

CHAPTER 2

THE ELEMENTS OF GENERALISATION AND THE REQUIREMENTS OF AN AUTOMATED SYSTEM.

2.1 Definitions: the difference in concept between automated and conventional generalisations.

The processes involved in generalisation of maps are complex and definitions difficult. Steward (1974) has tabulated the variety of terms used and the variety of meanings that can be applied to any one term and come to no universal definitions. No agreement can be reached even on the description of map scales - a 'small' scale in one country (eg 1:50,000 in the United Kingdom) may be the largest national scale in another. The difficulties in producing objective descriptions of generalisation processes has also been discussed by Uitermark (1980) who has drawn on the works of Robinson and Sale (1969 and 1978) and Topfer (1974) to list:

- selection of map content	('selektie van de kaartinhoud')
- simplification of detail	('vereenvouding')
- feature displacement	('verplaatsing')
- amalgamation	('klassifikatie')
- symbolisation	
- text positioning.	

Lichtner (1979) cites Hake (1975) and gives a more technical description of seven 'elementary procedures' of two main types:

i. pure geometric generalisation
- simplification
- enlargement
- displacement

ii. geometric/conceptual generalisation
- combination
- selection/elimination
- classification/typification (symbolisation)
- valuation (enhancement)

These procedures may also be classified on the basis of function into:

- those concerned with controlled reduction of information (simplification, combination, symbolisation, selection)
- those concerned with cosmetic enhancement for direct viewing (enlargement, displacement)

Moellering (1980) has discussed the differences between 'virtual' and 'real' maps (Fig.2.1) and this division of generalisation procedures can be accommodated into Moellering's classification. The first functional class of procedures is internal to a non-viewable type of map - for example removal of minor names from a gazeteer or filtering of line data are processes internal to a virtual map. The second class can be considered as a transformation between 'virtual' maps and a 'real' map or CRT display.

Directly viewable as a map image ?

		YES	NO
Permanent Tangible Reality ?	Yes	Real map Map sheets (drawn) Machine Drawn maps	Virtual Map II Gazeteer Survey data
	No	Virtual Map I CRT display Cognitive map (in two dimensions)	Virtual map III Map data on magnetic media Cognitive map (multidimensional)

Figure 2.1 Classes of 'maps' (after Moellering 1980)

In the following sections the differences between real and virtual mapping will be discussed. The fundamental difference is that the real map is dependant on the relationship between the physical space available on the medium of representation, whether it be screen, paper or film,and the interpretative powers of the human user. Virtual maps, on the other hand, have almost no limits in spatial resolution, and furthermore, are capable of allowing a much greater amount of qualitative description than a real map in terms of the number of classification of map entities, and relationships between them.

Terms that will be used in the following sections include 'features' and 'objects' that are represented. These are difficult, if not impossible to define, and are used for convenience. A map feature is taken as being an entity on a map that would be interpreted by a general user. More rational alternatives will be discussed.

2.2 Selection of map content

The selection of map content has been considered as having two components (Rhind 1973) - between and within features. The latter is feature simplification and will be considered separately.

Selection of map content includes two closely related factors, namely reduction of the number of features by density and by type. The number of features that can be depicted within a given area of ground on a change of map scale was considered by Topfer and Pillewizer (1966) who proposed 'radical laws of map generalisation'.

In their simplest form these are:

$$n_f = C_e . C_i . n_a . \qquad \sqrt{m_a \quad m_f}$$

where

n_f	is the number of items to be depicted at derived scale
n_a	is the number of items at source scale
m_a	is the denominator of the representative fraction of the source
m_f	is the denominator of the representative fraction of the derived scale
C_e	is a constant related to exaggeration of symbols
C_i	is a constant related to emphasis of areal or linear features ('symbolic form')

As Robinson and Sale (1969) point out, this 'law' gives no indication of the changes in classification or the simplification of detail required to conform to it. Similarly the effects of feature importance and clustering are not considered. It is of limited practical use and may be more appropriate at scales smaller than 1:50,000. Stenhouse (1979) has attempted to compare predicted numbers of towns depicted with the number of towns observed on atlas maps of the United Kingdom and found broad agreement on scales of 1:2,000,000 to 1:20,000,000.

Kadmon (1972) has attempted to use objective analysis to select settlements for map depiction using sixteen ranked and weighted descriptors including population, administrative status and a number of other socio-economic and communication variables. These descriptors were then used to produce settlement maps with varying degrees of generalisation. One of the variables used by Kadmon was a 'remoteness' index which he weighted heavily in order to ensure that sparse areas of map contained some detail. This perpetuates a conventional map practice undertaken to maintain style: cartographers have traditionally resented blank spaces, sometimes filling them with imaginative detail when they had no information, and minor detail when they had. Anisotropic selection of features is complex to emulate in automated cartography and it is questionable whether it is a valid primary generalisation technique, or whether it is cosmetic and valid only for graphic output, if at all.

Data selection on the basis of the attributes of each feature can only be undertaken if the structure of the data permits it. Current practices in automated cartography using vector encoding normally allow for some degree of attribute coding to be attached to a feature. This may range from the highly complex system employed by the United States Geological Survey National Mapping Service which allows for each feature to be coded with up to 24 codes from a maximum of 9999999 (of which only about 1000 are used) through that used for the Ordnance Survey large scales mapping program (one code from about 110) to simple systems differentiating only between line weights or graphical attribute of the line.

A common practice in conventional cartography is to distinguish only between major types of features directly - by varying line width, style or colour, or by different symbols - and then to apply some sort of secondary identification - a name (descriptive or proper), or a modifier (such as the 'conifer' or 'deciduous' symbol on the OS One Inch maps 'depiction of woodland'). This system is necessary to reduce the number of different types of line on a map and reflects the need to avoid visual over-complexity. An example might be two 'public buildings', say, a hospital and a town hall: many mapmakers (including Ordnance Survey) enhance these, using a thicker line or more prominent infill than for other buildings, but would differentiate between them by means of a name or textual description. It appears that users have little problem in differentiating between them on the map, relating the text to the outline, even if these are separated with other data intervening. To automatically link linear and textual information by means of spatial searching is complex and methods of coding have been devised which overcome this problem. Such methods include the hierarchical and relational systems which might encode the public building examples as:

BUILDING		line
PUBLIC	or	enhanced .....
municipal hospital		codeword (name)

In the first case, three code fields can be used to differentiate between the objects, but the name must be stored separately (if necessary), while in the second the inherent relationship between an object and its name is retained. A combined system in which the name is held as a 'modifier' to the code of the object might be used, but the example above illustrates the difference between coding for information purposes and that for cartographic purposes. The current Ordnance Survey large scale digital data contains codes with either no information about the object type (e.g. Object shown by dot) or nothing about the cartographic style (e.g. Railway). In neither case can text be related, so any generalisation involving feature type selection must be carried out on all members of a feature type and if only 'cartographic' coding is used, features of different types cannot be represented.

Simple object coding may be enhanced by codes containing topological information which present further problems in generalisation in that the topology must be maintained, or modified as necessary. For example, in Figure 2.2, the line BC might be coded so that the division between lake and wood (coniferous) is maintained. If generalisation takes place in such a way that the wood polygon BCXY is not to be shown then coding of segment BC has to be changed.

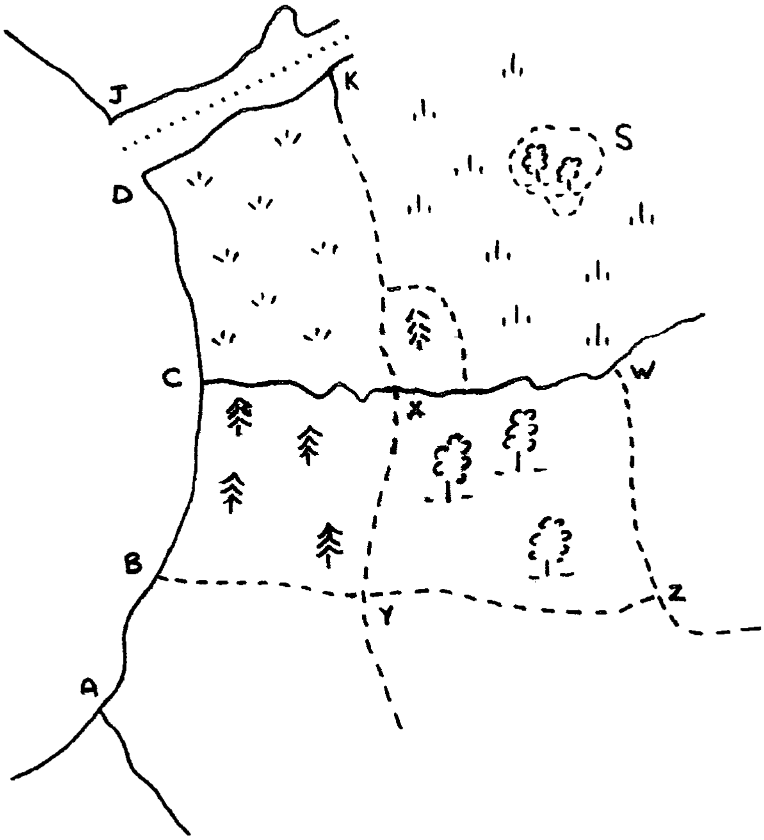


Figure 2.2 Fictitious example of a small scale topographic map

In order to reduce the complexity and effort involved in data capture, the number of codes available is normally finite and predefined, just as the specification for a survey limits the number of types of objects to be surveyed. The result of this must be some degree of classification which will influence the possible ways in which reclassification or selection will take place, the concept of a 'classification free' database being as impractical for most cases as a scale free one. Rhind and Hudson have discussed this with regard to land use mapping, giving examples of the problems of defining, capturing and processing classified data which are identical to those of topographic data in many respects. (Rhind and Hudson, 1981)

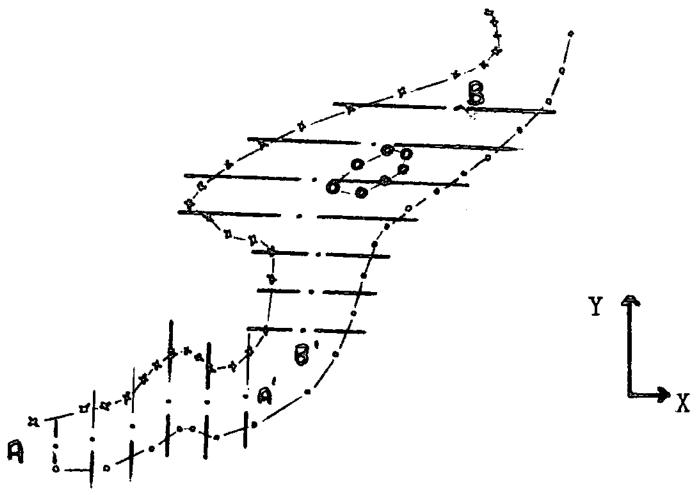
2.3 Amalgamation

Feature amalgamation is a semi-subjective part of the generalisation process in that decisions have to be made which may not be consistent across a map. As an example, the woodland parcel NPQX in figure 2.2 might be considered too small to show as a separate parcel on a given reduction of scale (or information). Conventional cartographic practices would probably result in it either being ignored or amalgamated with an adjacent parcel such as XWZY, provided the classification is appropriate - as it is in this case, both parcels being of woodland type. Parcel S is in a different situation, being detached and surrounded by a grassland parcel. Depending on the required emphasis of the map it might either be ignored, amalgamated with the grassland parcel in effect, or if the purpose if the map were to emphasise woodland, it might be combined with the larger parcel of woodland to the southwest, resulting in great exaggeration of its size and considerable movement in the apparent centroid.

This example serves to demonstrate that there are three major controlling factors of area amalgamation in generalisation; the minimum depictable area, the maximum distance over which amalgamation can take place, and the classification. Of these, the minimum depictable area and classification are determined by the specification of the map or dataset, but the definition of the distance between parcels is less easy to define, as it could be between centroids or between the closest points on the respective perimeters. These parameters can be applied in computer processing of map data and Lichtner (1979) has demonstrated that buildings on an urban map can be amalgamated successfully.

The ODYSSEY system developed by the Laboratory for Computer Graphics at Harvard University is also claimed to have the ability to amalgamate area features using the idea of a 'tolerance band' around each candidate feature. When these oversized polygons intersect, amalgamation takes place. (Chrisman, lecture at Ordnance Survey 1980.)

Linear amalgamation presents similar problems. As shown in the example of segment BC of Fig.2.2 reclassification may produce redundancy in nodes and provided that the coding is compatible, nodes B and C could be eliminated. A more complex problem is posed by amalgamation of lines such as those between D and K and J and H on Fig. 2.2 which might be required to be depicted by a single line. As figure 2.3 shows, lines such as these rarely have either the same number of points or similar spacing and so determination of a 'mean' position is non-trivial. Shiryayev (1977) has given a way of dealing with this in which each point on one side of the double feature is compared with those on the other. Methods in which an interpolated value for an increment in X or Y (Fig 2.3) is generated involve a considerable amount of computation and logical dexterity to be able to deal with highly sinuous features. Raster encoding the feature reduces this problem but this in itself may cause problems in available computer store. A local raster method is used in the Laserscan line following device in which software generates a centre line from the detected position of the line's outside edges (OS unpublished paper). In this a scan is made at right angles to the approximate alignment of a sub-parallel pair of features and removes the need to swap between interpolation along the axes of a universal raster as shown in figure 2.3



A-A' Centre line derived from interpolation in Y
B-B' Centre line derived from interpolation in X

A' - B' could use either interpolation - or a locally orientated grid

Figure 2.3 Creation of centre line from two adjacent lines by raster.

2.4 Symbolisation

While all maps involve the representation of 'real' objects by a symbol, there is a fundamental distinction between representation by a 'symbol' and a 'factual' or realistic depiction. As map scales decrease, the physical space available for a real object on a map sheet also decreases and the use of 'symbolic' rather than 'factual' representation increases. This process covers a number of procedures undertaken directly as a conscious effort to reduce one, two or three dimensional objects to a pre-specified depiction. The symbolisation may only take place in one dimension of a two dimensional object - a road is normally a two dimensional feature on a map at scales larger than 1:10,000 (in UK) but at smaller scales is scribed with a standard width although the length may be only marginally simplified. Symbolisation may be undertaken because at a particular scale a factual depiction is not needed, or because the physical size of the real depiction would not be in accord with the emphasis that the feature is adjudged to require. In conventional cartography symbolisation is almost always associated with exaggeration of size - particularly at small scales, but in a vector encoded digital representation of topographical information the symbol itself means nothing, and symbolisation is a matter of replacing a high order of dimensionality with a lower one. In a raster encoded dataset the symbols will occupy space, but pattern recognition techniques could be used to interpret tham and reduce them to a point or 'line' as appropriate.

The problems associated with symbolisation as a generalisation process relate to the physical size of a symbol in relation to the space available on the map. To overcome them cartographers resort to the cosmetic techniques of displacement, fragmentation and anisotropic selection.

Further aspects of symbolisation will not be discussed - they will be considered under other topics.

2.5 Simplification

Map content is frequently predetermined by the purpose of the map, hence simplification of the features on a map can affect the overall 'character' of the map in a number of ways.

First, simplification inevitably affects the cartometry - reducing the number of bends in a line, alters its length, or the size of the contained areas. Secondly, simplification, if carried too far, becomes caricature which may eventually cause gross distortion in the overall information content of the map, in addition to distortion of the spatial content: the corollary of this is that a good simplification reduces the data without reducing the message that the data is conveying.

The effect of simplification on cartometry has been considered by Kishimoto (1967) as reviewed by Maling (1968) who points out that logically a coefficient of generalisation could be determined by relating cartometry to scale and extrapolating to a unit scale and hence deduce 'actual' distances. As Mandelbrot (1967 and 1975) has demonstrated, lengths can be considered to be indeterminate at a 1:1 scale and all littoral landmasses of infinite fragmentation. Baugh (1975) and Baugh and Boreham (1976) had sufficient problem in defining the best estimate of the length of the Scottish coastline without even involving these concepts! Map users have to accept that maps can never fully represent reality so far as making measurements from them is concerned.

Mandelbrot's concepts of fractional dimensionality can be used to define what is happening when simplification takes place. His measure is one of the space filling ability of an object, giving it a dimensionality somewhat greater than the Euclidean - the coastline of an island being of the order of 1.3. Line simplification may be considered to be the preservation of this fractal quantity while reducing the amount of information. In an earlier paper, Pannekoek (1962) has stressed the need to retain the <u>character</u> of a depicted area. His example of the depiction of the Rhine delta on various published maps demonstrates the differences in portrayed sinuosity, a rough measure of the space filling ability and hence the fractal value.

Like Topfer's laws, Mandelbrot's ideas can give no clue as to how to approach an objective system of simplification. In discussion with the draughtsmen who produce the OS 1:250,000 scale map series it was found impossible to define rules that were used - every line had to be considered 'on its own merit' and in relation to the rest of the map. They stressed the need to ensure that sinuous features at 1:50,000 scale were sinuous at 1:250,000, again to retain the 'character'.

Rhind (1973) has discussed the various methods of automatically reducing line sinuosity, and concludes that 'in most cases high frequency filtering of the spectra present appears to occur, modified by the retention of accurate representation or deliberate caricatures of salient shapes'. He also mentions that re-chaining (amalgamation) of lines is frequently necessary and that the length of line segments may influence the generalisation process with certain algorithms. Rhind considers salients as a separate problem suggesting that there are three levels of sinuosity reduction:

- the lowest level is a caricature in which position and 'sinuosity index' are not retained
- an intermediate level in which position <u>or</u> sinuosity index are retained
- the highest level in which position and sinuosity are retained and other line features maintained in their correct relative positions.

The majority of work to date appears to have been directed towards the intermediate level of reduction. (Keir 1976; Douglas and Peucker, 1973: Page 1978;

Baxter, 1977; OS 1979; Vanicek & Woolnough 1975) The reason for this approach rather than the higher level one is probably because of the computing complications involved in ensuring that adjacent and contiguous features are in harmony. (Brassel 1977)

Rhind (1973) lists seven classifications of methods by which lines may be simplified, ranging from simple selection to complex techniques involving assessment of the importance of each point used to describe a line. None of the methods described can be used to relate features while simplifying; such an operation is considered to be a post-processing operation, probably undertaken interactively.

Methods involving 'intelligent' data reduction tend to have concentrated on filtering existing data rather than replacing it by parameters derived from frequency analysis. Vanicek and Woolnough (1975) have described a method and the Mapping and Charting Establishment, Royal Engineers, have used parameter-based data storage methods but found that extraction of the data for further processing was cumbersome and slow.

Douglas and Peucker (1973) described three methods based on deviation distance one of which has also been described by Lang (1969). In these methods the deviation of the orientation of a line segment from the previous segment or segments is tested against preset tolerances. Other writers have used similar ideas, Baxter (1977) using area of three consecutive points, followed by an angular deviation check against a tolerance; Page (1978) and Hunting Surveys (Keir 1976) using a 'sagittal' distance method. Ordnance Survey (1979) have developed a method combining both offset distance and angular deviation approaches.

Other proposals to aid simplification have included the flagging of important points at data capture stage (Koeman and van der Weiden, 1970) but this could present considerable problems in coding, especially when using automated or semi-automated capture methods. In principle, all possible stages of simplification would have to be taken into account at that stage. It is clear that simplification of line information cannot be expected to produce an identical product to that undertaken by a draughtsman who will tend to emphasise or suppress according to convention or intuition. For example, if a well known but relatively insignificant feature is to be shown on the map, it is likely that it will be slightly emphasised to ensure that its position and extent are unambiguous. To maintain spatial integrity, compensatory changes may have to be made in other features: to undertake this process automatically would be difficult.

2.6 Feature displacement and Fragmentation

Rhind (1973) has likened feature displacement to avoid overlap as variable local mapping, and notes that permissibility of shifting features depends on the purpose of the map. Fragmentation is a more extreme technique than shifting, resulting where there is no physical space on the map sheet to show all the features that should be present. Examples of the gradation between displacement and fragmentation are manifold: the breaking of contours on steep slopes is a good example, but less detectable cases can be noted where any feature intersects with another, especially at acute angles.

Like simplification, displacement is a skilled job undertaken by draughtsmen who are unable to describe how they do it other than by giving a hierarchy of archetypal situations - a road, railway and river in a narrow valley being the usual example. In almost all cases on a topographic map, networks are enhanced at the expense of areas - a road at 1:250,000 scale being exaggerated some 10-20 times in width. Other linear features may be contracted, but areas on a small scale printed map will inevitably be considerably less than their true value.

Lichtner (1979) has demonstrated that the displacement process can be undertaken relatively easily by a computer - for a single line, and at fairly large scales. Johannsen (1973) has described a quasi-interactive method for small scales, without showing results. Both writers emphasise that operator intervention is essential, and give no comparative figures for the amount of time and effort required to undertake a particular task compared with a skilled draughtsman working in a conventional manner.

Fragmentation of features is most common on conventional maps when another feature of higher position in a hierarchy must be shown. Normally text is given the highest position and all features are broken to show it. On a conventional map this is relatively easy - if a text overlay is made then photo-mechanical means, such as 'unsharp' masks can be used to blank out other detail. In an automated system 'box masking' of detail to allow text to be 'dropped in' is possible and available in many graphics packages, but at considerable cost in processing. However, the masking of detail on a conventional map is not a box but a halo around each character. To reproduce this technique automatically is feasible but would also add considerably to the cost of processing.

2.7 Text

On any map textual information occupies a special category of feature. Like contours, text does not actually exist on the ground in its own right; yet unlike contours it has no fixed planimetric position. It is related to features on the ground which may be of ill-defined extent, e.g. the Pennines.

On a virtual map descriptive text, (i.e. text other than proper names) might be replaced by a system of attribute coding which could then be used to produce descriptive text as required. Although this has problems in increasing the complexity of data capture, it is a more flexible approach as any real map subsequently produced at a different scale from the source will require changes in text position.

From the point of view of map generalisation in which a real map is to be produced, the explicit linking of features to associated proper names in some way has great advantages. If any selection or amalgamation of features is to take place, then the associated text must be selected or rejected. Positional movement may also be required as a result of this, but then text must be considered as a 'super-feature' in its own right, subjected to selection because of its requirement for space on a map. Simplification or an analogy may be undertaken - examples being the rejection of specific names in favour of general and the use of abbreviated forms of text. Displacement and fragmentation of the text information itself must also be considered.

2.8 Conclusion

In order to produce some criteria upon which automation of map generalisation can be based, a scheme for generalisation must be proposed. Such a scheme is shown in general at Figure 2.4, which emphasises the division between the information reduction aspects and the cosmetic.

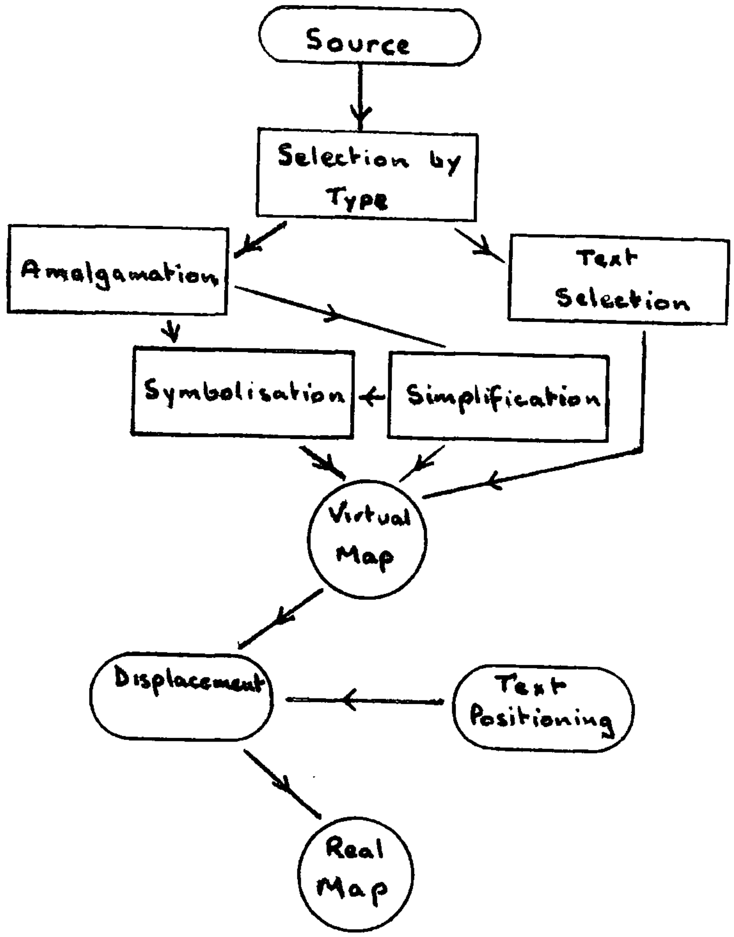


Figure 2.4 Schema for map generalisation

From the discussion of the individual aspects of generalisation in this chapter it can be seen that there are few major technical problems in producing datasets and maps which could correspond closely with those produced by manual methods. There is a difference however, between the technical feasibility of map generalisation and the economic feasibility. As Thompson (1978) has shown, the viability of computer assisted large scale mapping rests on the ability to use the data produced for other purposes, including production of derived (i.e. generalised) maps. Uitermark (1980) has concluded that of the generalisation processes discussed above, automated text emplacement is not likely to be economically viable, that simplification and displacement/fragmentation are likely to be so only in some circumstances and that symbolisation, qualitative amalgamation and selection are simple and likely to be undertaken. He also makes reference to only two production systems for 1:25,000 scale map production - at Ordnance Survey, and at Institut für Angewandte Geodasie, West Germany. Neither of these systems is using the more sophisticated techniques.

There is therefore scope for considering generalisation with regard to production processes. This will be done in subsequent chapters.

CHAPTER 3.

ORDNANCE SURVEY CONVENTIONAL SMALL SCALE MAP SERIES, THEIR SPECIFICATION AND THE METHODS USED IN THEIR DERIVATION.

3.1 Introduction

The Ordnance Survey produces a variety of maps ranging is scale from 1:1,250 to 1:1,125,000. Maps at scales down to 1:625,000 are arranged in two families: (OS 1980 d)

- the Basic scales of 1:1,250, 1:2,500, and 1:10,000 which are the largest scales at which the whole country is mapped.
- the Small scales which constitute the 1:25,000, 1:50,000, 1:63,360, 1:250,000 1:625,000 and 1:1,125,000 derived maps (plus some other special maps at intermediate scales).

The division of maps into families used in this dissertation is that commonly used internally within Ordnance Survey and which relates to the source of the information. However, as the 1:25,000 scale map series are presently derived by photo-mechanical reduction of 1:10,000 scale maps, which in turn are derived (when appropriate) from other basic scale maps, it is logical to include the 1:25,000 series within the Large Scales family. The 1:50,000 scale draws relatively little information directly from the large scale maps; revision, for example, is carried out by a team of surveyors working solely for this map family. The reasons for this dichotomy are mainly in the number of large scale plans that would have to be simultaneously revised and subsequently referenced to produce one 40km square 1:50,000 scale sheet; in the case of a densely populated sheet covering an urban area mapped at

1:1,250 scale this could amount to nearly 6,400 basic scale sheets.

This factor, along with the need for rapid national coverage and the requirement for features to be included which are not shown on the large scale maps led the Serpell Committee to recommend a dual approach to producing national topographic databases, one of which would be based on the large scale map data and another, (probably interim) one that would be based on the 1:50,000 scale maps.

3.2 1:50,000 series

3.2.1 Introduction

The 1:50,000 scale maps (now designated the 'Landranger' series) are the parent series for a family of maps extending to 1:625,000 scale. The information sources accessed in producing the 204 sheets in the current series are also utilized in producing these derivatives although the smaller scale popular maps (i.e. the 1:250,000 Routemaster series and 1:625,000 Routeplanner map) contain some additional information and are produced on a more frequent revision cycle than the parent series.

The specifications used within OS are laid down in detailed documentation giving both a description of how a feature should be depicted and examples. This is to maintain continuity within the map series and to ensure that different sheets drawn at different times and by different staff have the same appearance. The specification depends in part on what is technically and economically feasible to produce and also on:

previous specifications - i.e. tradition and what the user expects
map user conferences - in which representatives of users give opinions on content and style.
market survey (Drewitt 1973)

Ordnance Survey internal research, directive and opinion.

The present coverage of the United Kingdom at 1:50,000 scale is on two series, each of two editions: the First and Second Series in the Coloured and the Outline editions. The outline edition is essentially the same as the coloured edition, printed in monochrome without colour fillings for roads, without contours or tourist information and with Civil Parish boundaries added.

The First Series maps are the result of photomechanical enlargement and recompilation of the Seventh Series 1:63,360 scale maps and the specification is very similar to that of the latter, after allowing for the effect of enlargement (Harley 1975). The current catalogue (OS 198C) lists 73 Second Series and 131 First Series sheets. Within the Second Series, a further division is made on the basis of those with 'rational' (i.e. 10 metre) contour interval. This division is approximately equal (OS 1980b) and maps are currently being converted from the 'Imperial' to rationalised metric contours at a rate of about seven sheets a year.

The Second Series specification allows for a rescribed version of the existing information held on either the existing First Series sheets or, in the case of some sheets in Scotland, London and North Wales, directly from enlarged 1:63,360 keys. Apart from revision information, major new information to be added to sheets includes the metric contours for the 1:10,000 scale map series. Since as many as 64 sheets at 1:10,000 scale cover one sheet at 1:50,000 scale the conversion process is a slow one although metric contouring of 1:10,000 scale maps is scheduled to be completed by 1984. Tourist information showing places of interest, information centres, car parks and other items has also been added to the specification.

A certain amount of information portrayed on the

First Series was deleted from the Second Series specification. Much of this was considered by OS to be of little importance, based on market research (Drewitt 1973, HMSO 1979) and was expensive or difficult to maintain. Such information included items such as woodland type (coniferous or deciduous), differentiation of rough grazing land from improved land, bathymetric contours and civil parish boundaries (on the coloured edition).

The reaction from users has been mixed but (OS 1980b) refers to a forthcoming experiment to reinstate one of the contentious items, that of woodland type. In the case of submarine information the OS view was that the maps might be used as a substitute for proper charts and thus be potentially dangerous. Other losses of information will be discussed below.

The items to be depicted on the Second Series specification are listed at Appendix A.

3.2.2 Point Features

The point features depicted on the maps are of two main varieties. A point feature may be the symbolic representation of an item of finite size, plotted as such either because the area covered by the feature would be too small at the publication scale to be seen easily (e.g. chimneys, milestones, telephone boxes) or to enhance and standardise the depiction of items that could be correctly shown in outline (e.g. churches, railway stations) In some instances, the choice between factual and symbolic representation is arbitrary and may exhibit an implicit hierarchical structure as in the case of churches and cathedrals. Cathedrals are normally shown as a public building (i.e. 'factually', excluding minor juts and recesses, drawn in a thicker line than 'normal' buildings) and annotated. Small cathedrals may be shown by the appropriate church symbols. The converse, with large churches being shown 'factually' is not permitted in the specification.

The second type of point feature is that where a symbol is used to show a feature which is not tangible on the ground. This category includes some antiquities (e.g. battlefields, site of ....) and spot heights which may or may not coincide with the position of a surveyed mark.

The actual geographical position of a point is normally at either the centre of the symbol (e.g. triangulation pillar, or church without spire or tower +) or at the centre of the base of the symbol if it is a pictorial one (e.g. radio mast, coastal beacon). The true position as depicted may not relate directly to the geographical position since a point is frequently moved to accommodate generalisation of other nearby features.

3.2.3 Line Features

As Appendix A shows, line features comprise the majority of the items shown on the Landranger maps and of these over 50% of feature types depicted are in the communications groups, although contours account for the greatest line length (Howard 1967). Like point features, they may be divided into the intangible and the symbolic representations of areas of highly elongate shape. The first category includes two groups of major items on the maps, boundaries and contours. Special cases which form a transition between the abstract and the finite also exist. These include the lines depicting features which may require special conditions to occur - high water mark, bankfull stream limits and some historical information (e.g. course of old river, road or railway which may no longer be visible).

Administrative boundaries provide one example of a problem resulting from the adjustment of linear features to take account of the width of map symbols. On all large scale plans, boundaries are shown as 'mered' to an adjacent feature such as a wall, fence stream or hedge, or as the centre line of a wide feature, or as 'undefined'. Yet many of the features to which the boundary is mered are not shown at 1:50,000 scale, which results in the haphazard appearance of many boundaries on the maps. Alternatively, when a boundary is mered to the side of a feature that is to be depicted by a symbolic line, the position of the boundary line will be moved from its true geographic position. Over large areas this will have little effect but in some small civil parishes it will cause systematic errors in any measurements of area or perimeter distance from the map. The boundary is, however, maintained in its true topologic position with respect to the mered feature - which is probably more important in most cases.

Like any other map series that uses both contours and other annotations for surface representation, the 1:50,000 scale map contours are frequently fragmented. This fragmentation is considered essential to maintain a cartographic style but forms illogical breaks in the representation of a continuous surface. Discontinuities in the contours occur:

where contour values are inserted

at embankments, cuttings, quarries and wide water features.

in areas where adjacent contours would be closer than the specification allows.

The majority of symbols depicted on the Ordnance

Survey 1:50,000 maps are relatively straightforward and represent features that cannot be adequately depicted 'factually' at the scale. However, a number of anomalies exist in the specification particularly with regard to communication and water features. In certain areas of Scotland which have single track roads as classified by the Department of Transport, the differentiation between these and normal roads is on the basis of a pecked infill. Thus these roads cannot be differentiated on the Outline editions of the maps. This is probably of little significance. A more serious logical problem concerning roads relates to minor roads where the surrounding area dictates the symbolic representation. In urban areas, all minor (i.e. unclassified) roads are shown with a narrow gauge unfilled double line. When the area surrounding the road is not filled by the tint used to depict built-up areas, only roads which are untarred, drives and tracks are shown by this symbol and minor roads with tarred surfaces have separate symbols related to road width. Thus road width cannot be inferred from the maps in built-up areas and anomalies may occur in successive editions of the maps when new building is shown and roads which may have been upgraded are in fact shown by a narrower symbol. Criticism of road depiction on official maps of Great Britain has come from Morrison (1971) who maintains that extra information such as road condition or potential speed might be shown.

The width of the symbol used to depict streams depends not only on the width of the feature but on the position along the feature; an 'atlas' style is used to depict narrow streams in which a line constantly tapers from the source of the stream to the mouth. In the Second Series specification, the depicted width of a stream at the source is always 0.15 millimetres (7.5 metres at ground scale) and may be up to 0.6 millimetres at the mouth (30 metres).

However all streams between 4 and 8 metres wide on the gound must be scribed with lines not less than 0.3 millimetres wide and streams over 8 metres are shown as double line symbols with a minimum internal width of 0.3 millimetres. Thus a line 0.5 millimetres wide which apparently represents a stream 25 metres wide in practice represents one between 4 and 8 metres wide and a 10 metre wide stream is represented by a line with an internal width of 0.3 millimetres, giving an overall width between the centre of the casing line of 0.45 millimetres. Detailed analysis of stream features is further complicated by the atlas style of depiction because a stream under 8 metres wide which narrows cannot be shown as doing so.

A further example of the 'insensitivity' of some symbols is the case of railway lines in which it is no longer possible to differentiate between single and multiple track routes. Yet within the areas of sidings, considerable detail may be shown (e.g. Teeside sidings NZ 470 190). Although this detail may be of little consequence to the average map user, it may render the map unsuitable for use as a base for a map prepared (at smaller scale) for say, aviation purposes. (Current civil aviation pilotage charts still distinguish between single and multiple track railways).

Thus in the analysis of linear features on the 1:50,000 scale, there are several areas of ambiguity and uncertainty which might cause an over-ambitious user of the data in a form more easily manipulated than a map to make spurious inferences. In addition to the interpretation problems discussed above, there are the problems of geographical displacement of data which have been discussed in greater detail in an earlier chapter.

3.2.4. Areal features

Keates (1972) criticising the Ordnance Survey One Inch Series wrote:

> 'only the exceptional is represented so far as the land surface is concerned... if there is no symbol the user is expected to realise that there is some sort of cultivated or improved land present'.

This statement still pertains to the 1:50,000 scale maps. Indeed what he terms the 'negative information approach' has gone further in that it is no longer possible to determine whether the land is improved or not since the symbol representing areas of 'Bracken, Heath and Rough Grassland' has been discontinued.

Currently, explicitly described area features are relatively few in number and are shown on the maps as areas with a screened infill (e.g. woodland, orchard, large areas of glasshouses, bodies of water, built-up areas). Implicit representation of areas is more common and the usual method is by a bounding line (which may form part of another feature such as a road casing), and a symbol or a descriptive or proper name. In some instances, such as golf courses and danger areas, there is not necessarily any bounding line and the symbol is 'placed centrally to the area' (Ordnance Survey undated). Supplementary coding is frequently carried out by names although this may be limited to the proper name or the most basic descriptive name (for example hospitals are not described as mental, general etc., neither are they distinguished as being private or belonging to the National Health Service). In such cases, only the extent of building itself is usually shown and the grounds - which may be subject to the same authority and access restrictions - are not distinguishable.

Areal features may also be subjected to movement if required to fit in with adjacent features. More usually, they form buffer areas between links of the road network and are thus represented slightly smaller than is geometrically correct as roads are enhanced in width for clarity, to increase the usefulness of the map for road users and the need to preserve a minimum width of colour line for printing purposes.

3.2.5. Names

The names on any topographic map not only provide the user with information on the name of a feature but also allow a user to form some idea of the importance of the item to which they refer. In common with most other map series of this scale descriptive names are used extensively to distinguish between features of slightly different type which are represented by the same symbol and therefore form an extremely important component of the 1:50,000 scale maps.

Draughtsmen are given very precise instructions as to the fount and size of lettering to be used to describe a particular feature, but very little information on position. The basic rule is that 'the map user can easily identify the feature, or extent of the feature they describe.......Names will be placed so that the minimum of detail is broken. Motorway detail will not be broken'. Other rules cover the use of conjections, Celtic versions of Anglicised names and spelling.

Thus the names to be seen on a conventional 1:50,000 scale map are positioned not by any consistant rule, but according to the way the draughtsman considers that the <u>user</u> would be able to make unambiguous decisions as to reference.

3.3 1:250,000 Scale maps

3.3.1 Introduction

The 1:250,000 scale maps produced by the Ordnance Survey are currently the sixth series of such maps (including the 'Quarter Inch' maps at the closely related scale of 1:253,440), but are now designated the 'Routemaster Series'. This name accurately describes the emphasis of the maps which is on roads and items of use to the cross-country motorist. Appendix A lists the features depicted on the maps and demonstrates how a very large amount of feature selection has taken place from the 1:50,000 source maps. Some extra features are added, such as hill shading, layer tinting and road distances. A considerable amount of information is added in the form of marginal or inset boxes - but this is outside the scope of this dissertation.

The Routemaster maps are produced by a small team of very experienced draughtsmen and is unique within the Ordnance Survey in being the only map series to be still drawn with pen and ink rather than scribed. Names and symbols are provided from standard phototypesetter or symbol sheets, but all line information is hand drawn. It is also the largest scale Ordnance Survey map to be in direct competition with commercially produced maps of the same scale over the whole country, many of which use the Ordnance Survey map as a base, and it generates more revenue from copyright in this role than in the sale of paper maps. (OS Annual Reports 1975-1980).

A number of other products are made at this scale, including repayment tasks contracted to OS by the Civil Aviation Authority (CAA) and an outline edition for administration and planning. The CAA sheets are based on the ¼ inch specification and differ in a number of ways from the Routemaster - railways are distinguished by number of tracks for example and information pertinent to aerial navigation is enhanced.

Many of the techniques used to depict information on the 1:250,000 scale maps are identical to those on the 1:50,000 after allowing for the difference in drawing technique and the reduction in number of features shown.

3.3.2. Content

Because it is primarily designed to be a road map, great emphasis is made on road depiction, all roads being exaggerated more than any other feature type. This has created problems in congestion of detail especially in urban areas and so selection of through routes only is made. This contrasts with the situation in rural areas where very minor roads (including some private ones) are shown to depict access to major named buildings and small hamlets.

A rapid experiment to demonstrate anisotropic names distribution is demonstrated in figure 3.1 which shows a histogram of the number of names in 53 10 kilometre grid squares in England and Wales. These were chosen at random from areas in England and Wales selected for being Urban (U) agricultural (R) and mountain (M) in the proportion of 13:20:20:, using population data as the basis of selection

U - 'urban' grid squares
M - 'mountain' - " -
R - 'rural' - " -

	1 - 3	
	4 - 6	U M R
Names/ 10 km.	7 - 9	U U M M M R
	10 - 12	U U U M M M M R R R R R R
	13 - 15	U U U U M M M M M R R R R R R
	16 - 18	U U M M M M R
	19 - 21	M M M R R R
	22 - 24	U R
	25 - 27	

Figure 3.1 Histogram showing frequency of names in 10km grid squares for urban, rural and mountain areas.

This variation in names approximates to a normal distribution, but the stratification of the sample shows no preference. A small sample such as this (approximately 0.05%) of the area of Britain with coarse stratification can do little more than suggest that there is a tendency to name more minor features in rural areas. An analysis of name types confirms this (Figure 3.2)

	Urban(%)	Rural(%)	Mountain(%)
Spot height	-	20(7)	43(16)
Archaeology	-	28(10)	31(12)
Village (Suburbs)	94(57)	130(48)	55(21)
Towns	40(24)	30(11)	2(1)
Water features	12(7)	27(10)	83(31)
Major areas	2(1)	3(1)	3(1)
Others	18(11)	34(13)	52(20)
	166	272	266
Mean	12.8	13.6	13.3

Figure 3.2. Analysis of sample of names on Routemaster maps

Such an analysis, although again based on the small sample confirms the subjective nature of generalisation, at least in names.

To quantify feature displacement is more complex. An analysis of one sheet of the 1:50,000 Second Series (no. 93, Teeside and Darlington) with the equivalent

1:250,000 (No. 5 Northern England) was undertaken for a number of features. Over the whole 1:50,000 scale map sheet the distribution of vector differences demonstrates very little difference between the two, allowing for the precision of measurement (a good quality scale was used and allowance made for the differences in map stretch by measuring from grid lines). On a road map, railways might be expected to be displaced more than road features, but this is apparently not so in this case.

In order to test the effect of differential movement on features, twenty well-defined road intersections around Darlington were measured (Table 3.2), and the displacements noted. It can be seen that there is a constant shift between the two maps in eastings of about 120 metres, and in northings of about 70 metres, (Fig.3.3). Removing these shifts, the residual vector difference would be expected to be the result of random errors, which is more or less what occurs (Fig.3.4), when considered overall. The small circles on Fig. 3.4 show the relative enlargement of roundabouts at 1:250,000 scale, and when considering the overlap of these in the central cluster of points (the Darlington inner ring road) it is possible to see how the draughtsman has shifted the points to prevent overlapping roundabouts.

It must be emphasised that this experiment is to demonstrate what happens in one particular case; it is not meant to be an analysis of the generalisation. To do this, very many other cases would have to be examined, using high precision measuring devices and stable materials. Furthermore the main purpose of maps such as the Routemaster series is not the overall spatial accuracy, but the accuracy of the relative information. Nonetheless this example serves to show the problems that might arise in trying to describe objectively a fairly straightforward technique. It is not a complex case, but on British maps it is a common one - the East Midlands and Yorkshire sheet (No. 6) has at least forty similar situations of closely spaced roundabouts at which similar decisions have to be made about displacement. Many of these are much more complex than the example given - those in Manchester, Birmingham and Leeds are probably the most complex, and while a draughtsman has to experiment, shuffling roundabouts by repenning and erasing, the greatest possible advantage of an automated system is not to be able to undertake it wholly automatically, but to be able to remove the need for a physical working drawing at any scale convenient to a draughtsman. This then is the premise on which the next chapter will be written, that an interactive system is essential for the automation of the generalisation of the 1:50,000 scale maps.

Table 3.2 Analysis of variation in planimetric position between 1:50,000 and 1:250,000 scale maps.

Junction	13	14	15	16	17	18
13		36	68	81	99	75
14	25		32	48	51	64
15	77	46		24	38	68
16	92	30	43		19	56
17	81	65	60	32		38
18	78	76	95	76	45	
Scale		1:250,000				

Mean Inter-junction Distance 531 metres

1:50,000

Mean inter-junction distance 614 metres

Note: The above table is an extract from a comparison of inter junction distances derived from 1:50,000 and 1:250,000 scale maps of Darlington. The relative displacements of all 20 junctions in the sample are shown in Figures 3.3 and 3.4

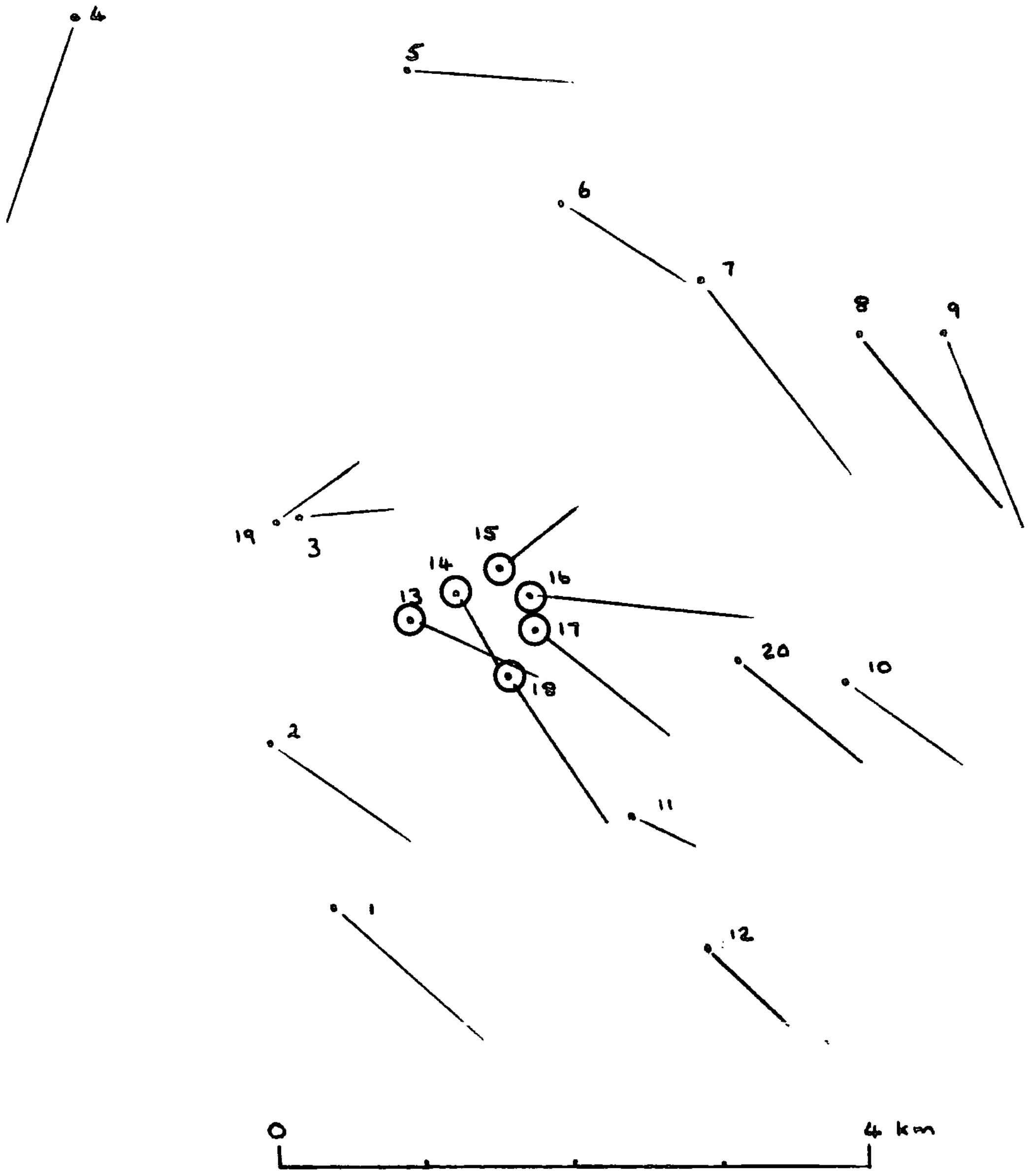


Figure 3.3 Gross vector differences between the representation of road junctions around Darlington

Scale of plan position - 1:50,000
Scale of vectors - 1:10,000

○ Representation of roundabout symbol at 1:250,000 scale enlarged to 1:50,000 scale.

Figure 3.4 Normalised vector differences between the representation of road junction around Darlington at 1:50,000 and 1:250,000 scales

Scale of plan position - 1:50,000
Scale of vectors - 1:10,000

CHAPTER 4

AUTOMATED GENERALISATION OF THE 1:50,000 SCALE MAPS

4.1 Introduction

It has been shown in previous chapters that in any automation of generalisation to produce 'real' maps there is bound to be a compromise between the purely geometric routines, carried out by a computer in 'background' modes, and the 'geometric-conceptual' routines carried out interactively, yet in considering the practical application of this the question of cost arises. To carry out all the tasks of a manual generalising system, the information to be built into the source data would have to be extremely detailed, and the situation would arise where it is totally uneconomic to gather data in the form required and it is cheaper to redigitise at smaller scales, rather than automatically generalise from larger scale uncoded material.

With this in mind, it is therefore essential to balance the requirements for any digital map data. The supplying agency (OS) must consider the costs and practicalities of input, its own internal use (such as generalisation) and also the users' requirements, which will vary from simple graphic production to interfacing with data base management systems.

On the data capture aspect, cost and time are of the essence. The more complex the data set, the more difficult it is to capture, encode, validate, edit and store. There is a minimum threshold of timing and cost, that of capturing a basic data set without attributes (known as 'spaghetti', a term originating possibly in the Laboratory for Computer Graphics in Harvard). Simple attribute coding can add considerably to the threshold cost; complex information gathering could make the cost rise explosively.

This chapter considers first aspects of an ideal data set and system, and secondly, what might be feasible.

4.2 Information on a map

The 1:50,000 series was described in the previous chapter to be a hybrid product much of which is blank space overlaid by a network of lines representing roads, railways etc. All of these interact and using the 'phenomenon-based approach' (Mark 1979), a model can be built up of these interactions on a logical basis. It is known that rivers and roads cross only at fords, bridges, tunnels or ferries, so a relationship is established between roads, rivers and the crossing features. That built-up land is mutually exclusive to water or forest, railway station exist only on railways in use, (they revert to being 'buildings' upon closure), power lines may cross any other feature, are all examples of easily formed explicit relationships. Others may be more subtle and probabilistic, such as that coniferous forests are unlikely to be contained within urban areas; deciduous woodland is more likely in such areas, but neither is particularly common compared with 'ornamental land'. Other subtle examples include the fact that human habitation rarely occurs above 700 metres (in Britain), that motorways and railways have gentle gradients compared with, say, minor roads, air and sea ports must have land access. All these are trivial and obvious to a skilled human map interpreter, but at present a machine has to be explicitly informed of the existance of any relationships if it is to consider them during data processing.

In order to represent these relationships conceptually, techniques introduced in data analysis (in the sytems useage rather than the statistical) can be used. One such technique is entity modelling, an example of which is shown at Figure 4.1. This example is of a general over view of some of the major relationships between major feature types; for more detailed planning of data structures a finer resolution of feature types is required.

Such a diagram as Figure 4.1 can be used within generalisation to assess some of the affects of undertaking a particular procedure. As an example, in the situation where a road crosses a river at a bridge, removing the river would leave an illogical bridge (over nothing) so it too should be removed. Similarly, displacing or smoothing the course of the stream may involve moving the position of the bridge.

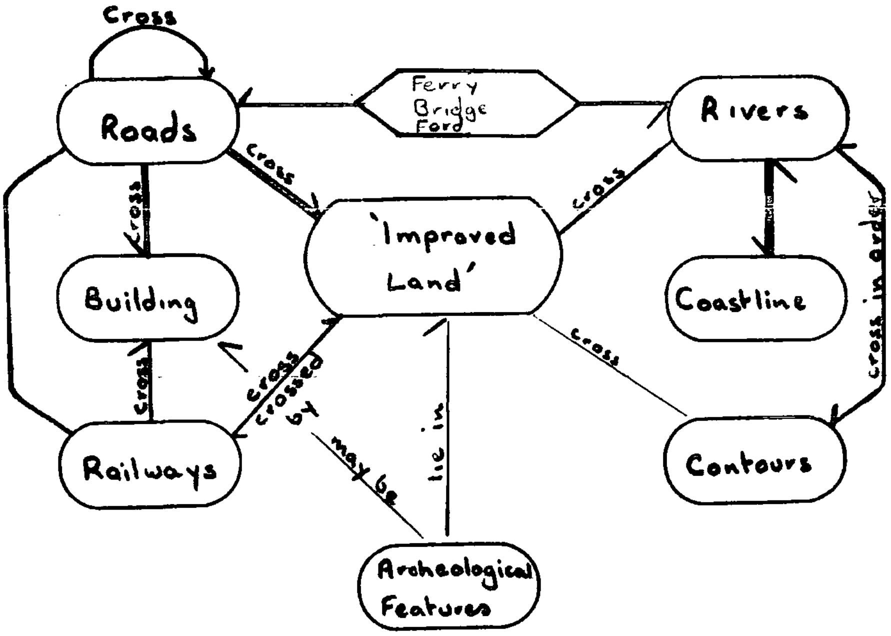


Figure 4.1 Simplified relational model of 1:50,000 scale map data. Thickness of line indicates importance of relationship.

This modelling approach does not, however, describe all the conditions of feature interaction. In most cases, a road may cross many rivers and a river be crossed by many roads, but a bridge only carries one road across one river (Figure4.2) This is an example of 'many to one' interactions -

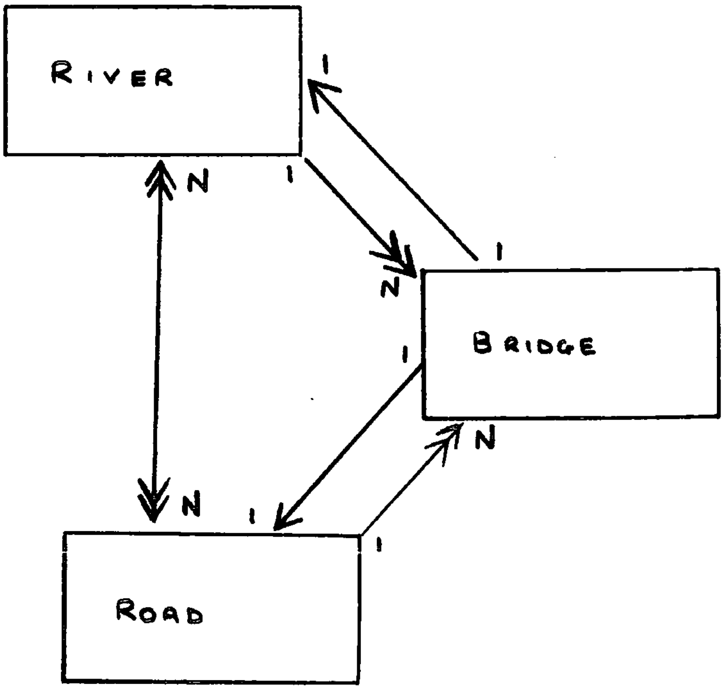


Figure 4.2 Interactions between features - example of 'many to one'

rivers are crossed by many bridges but each bridge crosses only one river, and the number of such interactions can be quantified either intuitively or empirically. This is relatively straightforward to include in the design of a data structure to represent the information. More difficulty is experienced when the 'neighbourhood' has to be considered. For example, the interactions between contour lines and streams are on a 'many to many' basis, but a stream cannot cross the same contour line more than once - it must progressively cross lines of decreasing value (or increasing depending on the direction of consideration) and, most complex, the shape of the contour line where it crosses the stream may be of importance. To incorporate this into data is no doubt possible, but whether it is feasible on a production basis is questionable.

4.3 The form of the data for generalisation of the 1:50,000 scale maps

There is a fairly well-defined trichotomy of existing and conceptual cartographic data:

- unstructured 'spaghetti'
- topological
- polygon

To some extent this can be applied both to raster and vector encoded data, but is more clearly defined in the latter.

Chrisman (personal communication) has outlined the relative advantages and disadvantages of each type in a number of respects when encoded as vectors as in table 4.1.

Requirement	'Spaghetti'	Topological	Polygon
retrieve line	easy	easy	
retrieve polygon	X	simple	easy
contiguity verification	X	easy	messy
polygon overlay	X	practical	hard

For raster encoding the position is different; polygon manipulation becomes easier (for example polygon overlay becomes simple) and line handling more complex. (Peuquet 1979).

Polygon encoding of data was developed as a result of the need to display thematic cartographic images based on polygons, the basic concept being that each polygon is self-contained and represented as one unit. Hence all lines except those on the outer hull are represented at least twice (depending on polygon nesting) and problems are frequently experienced in unintended 'slivers' of overlap or voids. As the Ordnance Survey map data is not polygon-based, but line-based, this method of encoding cannot be considered practical.

Unstructured data is essentially for graphic purposes and will not be considered further, other than to mention that it is the basic level of data in many systems - USGS DLG-1 data, OS databank, ODYSSEY input etc.

Topologically described data are probably the major area of research in computer-assisted cartography at present (Dutton ed., 1977, Cox and Rhind 1978). The reason for this interest is that such data combine the relative simplicity of unstructured data with the complete description of the map surface in the polygon systems without problems of double representation of boundaries. Fundamental topological systems such as DIME (Cox and Rhind 1978) are based on graph theory with associated algebraic operators.

The fundamental topological systems use descriptors on the edges to describe the regions lying on either side, and in DIME only the nodes possess geometric information, i.e. all edges are straight lines. This is unacceptable for many cartographic purposes and DIME variants, such as POLYVRT, have allowed for edges to possess sinuosity.

The basic topological description is a,b;c,d, (Figure 4.3) of from , to ; left, right. This allows for complete description of connectivity and adjacency in a map but, as has been outlined in Chapter 2, maintaining this description when generalising can be complex. A conceptually higher level of topological description, based on the connectivity of adjacent two dimensional objects can also be considered (Figure 4.4) which could provide an ability to undertake rapid spatial searching for such purposes as finding potential neighbours for amalgamation.

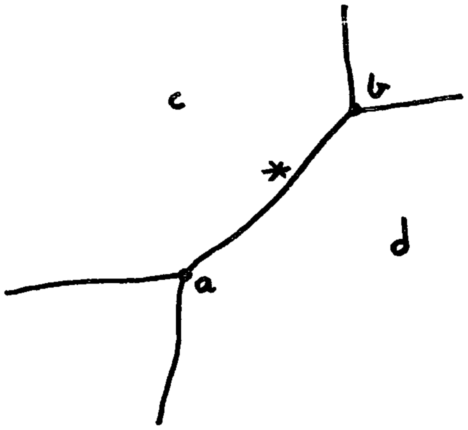
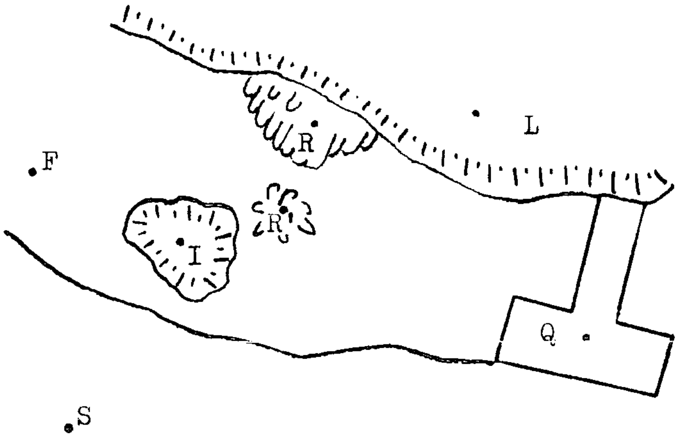


Figure 4.3 Topological description of edge * is a,b;c,d (after Cox and Rhind 1978).

Some methods of coding might use an implied multi-level system of adjacency coding attached to the edges. For example a line coded as high water mark which actually forms a land/foreshore boundary might be a land/sea boundary if the foreshore is deleted during generalisation. This relationship could be implicit - all land/foreshore boundaries could be considered as land/sea at a particular level of generalisation or explicit, which would allow anomalies such as offshore tidal islands to be taken into account (Fig. 4.4)

Complex coding of boundary lines to represent areas has been used in a number of topographical datasets (Cox and Rhind 1978) but in most cases at one level only. Using a concept such as an 'area node' which carries pointers to both current and potential neighbouring polygons could, it is believed, be of use in the generalisation of vector encoded maps.



Area	Neighbours	
	Actual	Potential
S	QF	LR
L	RQF	SIR'
F	RLQIR'	
R	LF	R'
R'	F	IR
I	F	LS
Q	SFL	

Figure 4.4 Example of actual and potential neighbouring polygons and of the use of 'area nodes' to indicate these relationships.

In conclusion, it is considered that a data set created from source maps such as the Ordnance Survey 1:50,000 scale series

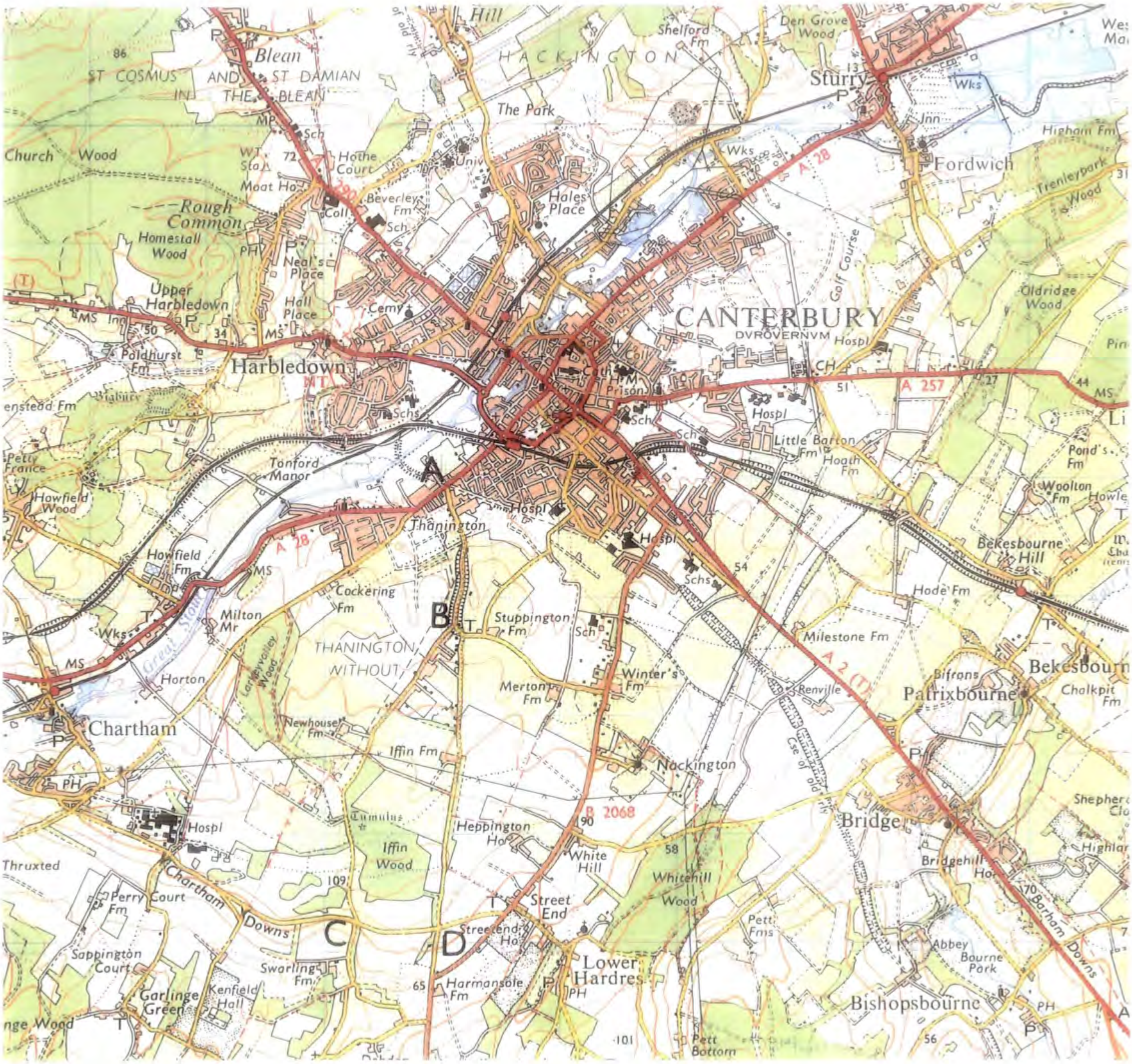


Figure 4.5

Extract from 1:50000 scale map 179 (First Series) information encoded with a view to generalisation should include:

- linear and point features encoded in a rational form, using multiple hierarchical qualitative descriptions of the feature. This fulfils the graphic and 'topographic' requirement.
- topological representation of the network constituting the map so that all explicitly described areas are directly represented.
- information concerning logical connection (or separation) between individual parts of features and other features. This information is aspatial and may be conceived as features being attributes of other features.
- text used as a modifier to an attribute of a feature and also as a feature in its own right (in the case of indeterminate features).

4.4 Would it work, and is it really necessary?

The arguments above are made with only general reference to the maps themselves. To demonstrate the requirement for the complex contents of a data set, consider Figure 4.5, an extract from the Ordnance Survey First Series sheet 179 (a First Series sheet was chosen because it contains more information than the Second Series).

To demonstrate the requirement for 'multiple, hierarchical, qualitative descriptors consider the roads ABC and ABD - 'normal' minor roads.

At A, the road running south from the A28 is a normal unclassified road more than 4 metres wide - thus it has a simple descriptive attribute with a width modifier which could be applied to the centre line. After 400 metres it becomes the parish boundary - the secondary attribute. 100 metres further on the road enters a cutting - another modifier (to the road modifier as far as the boundary part is concerned, a cutting is insignificant) and the road is also a footpath (the third main attribute). At B it forks, leaves the cutting and so goes on with a variety of changing circumstances (unfenced on one side, narrows, parish boundary, footpath etc.,) until C and D. Thus on this short road section there are 14 changes in the description of the roads alone.

Had the parish boundary been a county boundary, then districts and parishes bounded would also have been implied, demonstrating the direct hierarchical nature of <u>inclusive</u> features. Roads, on the other hand, are mutually <u>exclusive</u> of space unless elevated. (for this exceptional case a modifier to the normal case would be required). Within the road network, these two minor roads have another 15 nodal points at which another road or track joins, and two more occur where a power line crosses.

Hence to describe this seven kilometre length of road a maximum of three main attributes and six modifiers are required and the connectivity has to be described by some 20 nodes where either a road joins or the qualitative description changes. Furthermore, to describe the areas on either side of the roads there are some 30 changes of 'vegetation'.

This exercise demonstrates the complexity of information that the map contains - A to B is 1.5 km out of hundreds of thousands of kilometres of similar roads in the country.

For an example of the advantages of logical connections between features, the river Great Stour is close to the threshold below which (in Ordnance Survey terms) it would be represented by a single line. Generation of this line involves the 'averaging' of the two outer casings and on the extract shown the only connection is at the edge of the sheet. To allow the averaging to take place, the two banks need to be identified and so a spatial search could be made or one bank braced to the edge of the 'sheet', along the edge of the sheet and so to the other bank. However, it is simpler in theory to simply code one as 'left bank' with a pointer to the appropriate right bank. At river junctions the pointers would alter - but the coding would not.

Finally, the linking text to other features must be considered. On the Canterbury extract, text is used to differentiate between the cathedral, prison and college, and there is some ambiguity. Text linking in the digital map data would prevent this and would have the added advantage of reducing the number of classification types required.

4.5 Conclusions.

The discussion of the information content of a 1:50,000 scale map, together with analysis of an example demonstrates the importance of relationships in any complex interpretation of maps, including those for generalisation.

Similar problems have been identified in topographic data sets elsewhere, and two main strategies have been adopted in the United States of America where the United States Geological Survey Digital Line Graph System uses a very complex system of software-generated pointers to indicate relationships between features (USGS 1980 Smith pers. comm.) This requires manual editing to be carried out at a lower level because of the complexity of handling these relationships between features. Hence the system is relatively inflexible, but users do not require particularly complex software to extract information.

The alternative approach, used by the Laboratory for Computer Graphics at Harvard in the ODYSSEY system is to use a relatively simple spatially sorted file structure, and rely on complex processing. Dutton (1977) has described how this system is 'navigated' using a series of program modules to convert data from an unstructured to structured form, using intermediate files which may be of use in their own right. Hence data handling and editing is relatively straightforward, but to undertake simple tasks such as plotting or more complex ones such as the creation of polygons, a considerable amount of processing is required. However as Dutton (1977) points out, the data storage requirements have been minimised at the expense of complex programs.

Both these systems have procedures for generalisation, including feature selection and line smoothing. It is difficult to judge how cost-effective they are in use, but the ODYSSEY system has been recommended as a unified approach for mapping for the US Army (Sharpley, Lieserson, Schmidt, 1978). Whether such a large system could be used effectively in an environment where single user maps is required is, however, open to question.

CHAPTER 5

Experiments in generalising 1:50,000 scale digital map data.

5.1 INTRODUCTION

The previous chapters have described theoretical aspects of generalisation, this chapter describes a series of experiments made on data derived from 1:50,000 scale Ordnance Survey maps.

The experiments involved the following stages:

1. Obtaining and reading suitable data.
2. Performing an initial exploratory analysis
3. Definition of possible generalisation processes
4. Setting up and carrying out experiments
5. Forming conclusions from the above stages and deciding upon suitable processes for generalisation of data of this type.

The facilities available for computer experimentation using map data within the Department of Geography at the University of Durham included a PDP-11/34 mini-computer (192 k byte memory, floating point processor running under the RT-11 operating system with 16 k BASIC+ and assembler languages), and access to the NUMAC (Northumberland Universities Multi-Access Computing) System IBM 370/168 computer. Terminals were available for online access to both these machines and included a Tektronix 4014 storage tube graphical display unit. A CIL drum plotter could be used for producing plots of graphics data files. The NUMAC system supports most high level languages and has extensive packages of programs for a number of applications, including the display of graphics.

As a preliminary to the study a number of programs were written in conjunction with T.A. Adams to allow data to be transferred from the NUMAC system to the Department of Geography mini-computer. Programs were also written to allow the PDP-11 to drive the Tektronix terminal in graphics mode and to edit data and program files. The PDP-11 was used for program development and algorithm testing but the NUMAC system was used where large amounts of processing were required.

5.2 Obtaining and reading data

It was decided that there was insufficient time to develop a system to digitise maps and in any case digitising facilities were only available in Newcastle. Experimentation was limited to data available from other sources.

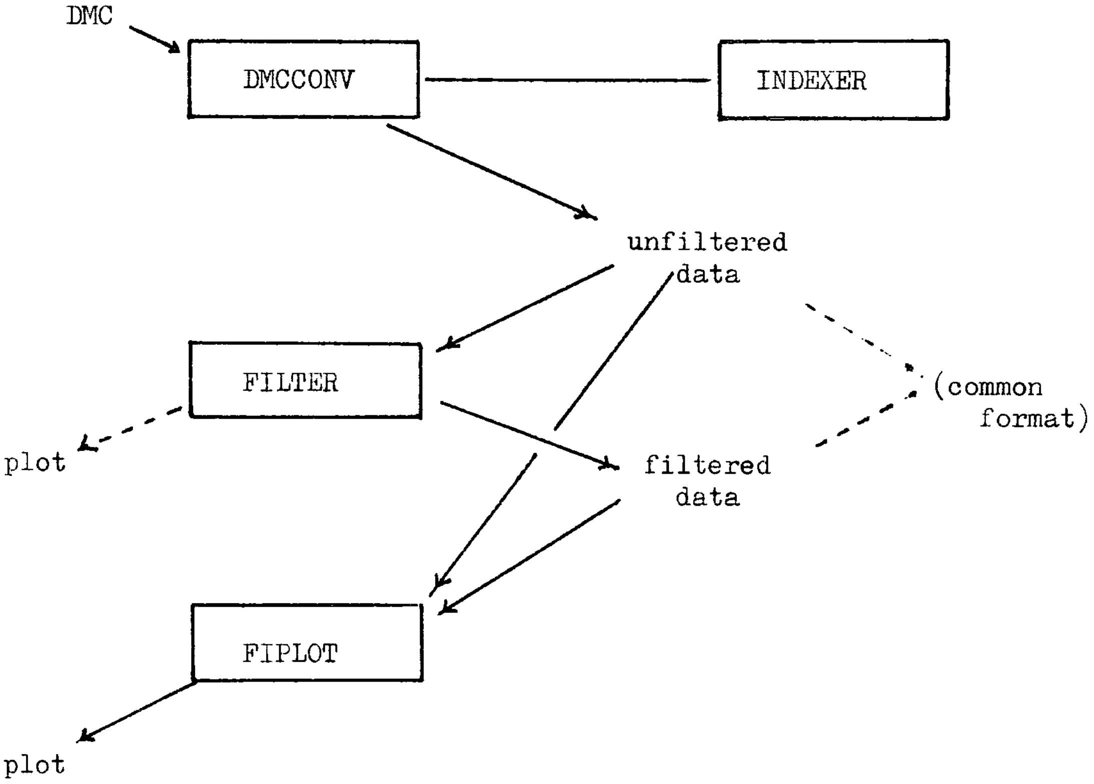
Two sets of data digitised from 1:50,000 scale maps were available. The first included hydrological and some cultural data for an area of South Devon (1:50,000 scale map sheet number 202) which had been digitised experimentally by Ordnance Survey in 1976/7 for evaluation by the Water Data Unit. These data are incomplete in some respects and the hydrological information and cultural information were digitised using slightly different techniques.

The second set of data was created for the Scottish Development Department's experimental Rural Land Use Information System (RLUIS). These data were digitised by Ordnance Survey to a high level of geometric accuracy and consistency of coding. It was therefore decided to use them for the experiments and a copy was obtained from the Program Library Unit, University of Edinburgh.

The RLUIS data set represents coastlines, rivers, roads, railways, powerlines and administrative boundaries for an area of Fife. It is coded in the Ordnance Survey customer format, DMC, and a simple conversion system was written to allow efficient use of the NUMAC system filestore. The system is outlined in Figure 5.1. The data were organised in spatial units of of 10 x 10 kilometre grid squares and it was decided to retain this organisation and concentrate upon one or two of these units for experimentation.

The initial processing of the data for each grid square resulted in:

- a file containing the co-ordinates of each point relative to the south west corner of the 10km square in which it lies, to a resolution of 1 metre.
- an index file to the above for each feature, giving the feature code and start and end co-ordinates.



DMCCONV - converted the DMC format data to a filestore file in the required format.

INDEXER - created an index to the converted data

FILTER - filtered data using interactively input parameters and produced optional plot (at source scale).

FIPLOT - used both unfiltered and filtered data to produce plots at source scale and 1:250,000 scale.

Figure 5.1 Outline of processing system for 1:50,000 scale map data in DMC format

5.3 Initial analysis of the RLUIS data.

As the data produced for the topographic information in the RLUIS dataset were in the same format as the map data derived from Ordnance Survey large scale maps, an analysis could be performed on them using programs written by Adams for his study of the likely characteristics of the national topographic databank (Adams 1979, Adams and Rhind 1981). The object of analysis was ta derive statistics about the source data by spatial unit and feature type. The full results of this are given for each of the spatial divisions of the data at Appendix B; a summary of the statistics for the whole dataset and selected feature types is presented in table 5.1.

The data are sorted into twelve contiguous spatial units of which seven include some area of sea and all contain only data pertinent to the RLUIS investigations. The total land area for which there is representation is approximately 500 square kilometres. The content of the data in qualitative terms was dictated by the requirements of the RLUIS study and by the cost of data capture; the data represents coastlines, administrative boundaries, communications and water features, together with some minor cultural features. For a topographical dataset the major ommissions were contours and habitation data, which were excluded on the grounds of cost or because they were to be supplied by one of the agencies participating in the RLUIS project. (Ordnance Survey unpublished paper, 1980)

The results of the preliminary analysis were used to indicate ways in which experimentation could proceed, and also provided an indication of the characteristics of the line data that can be used to represent a 1:50,000 scale map. However the results of this analysis have to be considered with regard to the primary use of the data and the methods by which it was captured and processed, together with the original purpose of Adams' investigations, which included estimates of line length and data quantity. No attempt to predict these quantities on a national scale has been made in the current work; the sample of the data cannot be considered represen - tative.

	1.	2.	3.	4.	5.	6.	7.	8.
Type of feature	Number of features	Number of segments	Total line length (KM)	Segment length (m)			Segments/ feature	mean feature length
				Mean	Max	Min		
Boundaries	145	17640	431.7	24	1472	1	121.6	3.0
Railways	370	4037	221.6	55	770	1	109.1	0.6
Electricity trans- mission lines	59	343	165.6	583	2346	12	5.8	2.8
River:								
Banks	269	14209	193.4	13	408	1	52.8	0.7
Single Streams	663	30141	543.4	18	1005	1	45.5	0.8
Centre Lines	14	1830	41.8	22	306	1	130.7	3.0
High Water Mark	40	5233	87.1	17	27	1	130.8	2.2
Roads:								
Dual Carriageways								
Motorways	27	235	14.4	61	513	1	8. 7	0.5
Trunk/main	14	112	6.7	60	610	1	8	0.5
Other	20	75	6.1	81	272	4	3.8	0.3
Single carriageway								
Major	579	4135	213.1	51	924	1	7.1	0.4
Secondary	363	2482	133.8	54	538	1	6.8	0.4
Minor	2558	22039	835.4	38	1079	1	8.6	0.3
	5121	102511	2894.1	2 8.2	-	-	20.0	0.6

Table 5.1 Summary of the characteristics of data representing some topographic features in western Fife

The methods of capture of data have affected the statistics in a number of ways. The minimum segment length for example,in the case of co-ordinates captured by digitising in a continuous stream, may be a functionof the rate of sampling of the digitising tablet, and the speed of movement of the cursor, the resolution of the tablet and the scale of the source document. The length of a linear feature may be restricted either by a convention(for example, to break linear features at particular points such as intersections) or by the concentration and physical ability of the draughtsman. The conclusions to be drawn from these statistics can only be general; it would be erroneous to conclude, for example, that motorway intersections occur every 500 metres. The convention used required the draughtsman to digitise motorways so that all cartographic intersections were shown by terminating a feature rather than the real situation. The minimum segment length is a function of the digitising tablet resolution (in this case 0.001 inches or 1.2 metres on the ground) and the post-processing spline routine that generates points along curves according to the algorithm designed by McConologue (1971) as implemented by the Ordnance Survey.

The statistics can, however, be used to confirm hypotheses about the data. From Table 5.1 a broad classification of feature types can be made on the basis of segment lengths into:

1. highly engineered features - electricity transmission lines
2. engineered features - dual carriageways, railways, major and secondary roads
3. non-engineered features - minor roads, hydrography boundaries.

This classification has important implications in cartographic generalisation because it represents lines of different 'character', and, as was stressed in an earlier chapter, preservation of line 'character' in generalisation is important for an acceptable cartographic result.

The value for some of these statistics is a function of the geographical character of the area itself. For example, the maximum segment length of a single stream is almost certainly the result of diversion either for agricultural purposes or because of mining subsidence.

The analysis using Adams' program was therefore of interest in exploration of the data set. However it relies wholly on the distance between points and gives only an overview of the problems likely to be experienced in generalisation.

5.4 Derivation of a line simplification method

Having carried out the exploratory analysis of the data, the possible methods for generalisation of the line data were considered. Because of the quantity of data involved, the primary consideration was to reduce the number of co-ordinates that had to be processed. A reduction by feature type was trivial provided the coding method used during digitising was accepted, and subsets of the data on this basis could be easily created.

Reduction of the number of co-ordinated points in each subset was more complicated. Each of the techniques proposed by Rhind (1973) was considered. These included:

- simple selection of points
- complex selection of points
- averaging
- arc substitution
- tolerancing
- frequency component filtering

The arbitrary selection of points either by the selection (or removal) of every nth point is a technique that is designed for gross reduction of data volume. It is conceivable that this would be of use in the case of data in which there were a low standard deviation and low mean value for the segment length. Such data might be the result of stream digitising on a short time or distance sampling basis. This technique, together with the more complex version proposed by Boyle (1970) were therefore rejected in this case.

The averaging techniques, such as those proposed by Koeman and van der Weiden (1972) or Gottschalke (1973) are also designed either for very large data sets or in cases where retention of geometric accuracy is not important, because, like arbitrary selection, there is no consideration of significant points.

Substitution of arc or polynomial segments for curves is an attractive concept, and can achieve great reduction in the number of points stored, but it is essentially a method suitable for archival purposes because for some data manipulations and almost all plotting the co-ordinates have to be regenerated, which may be a considerable overhead on processing. Breward (1972) estimated that storage requirements for data representing contour lines could be reduced by 80%. The substitution of mathematical functions for rectilinear features is impractical and whereas contour lines are usually smooth curves, the RLUIS data represents a combination of smoothly curved and rectilinear features in almost all feature types.

To test the possibility of mathematical representation, the program given in Baxter (1976) was adapted to accept data of the RLUIS area, but the test features were not successfully recreated.

Frequency component filtering was considered for generalising the semi-engineered and non-engineered curves. Rhind (1973) says that there are attractions in 'rigourously reducing the amplitude and sinuosities.....in relation to the scale at which they are to be produced'. For the data available there were a number of difficulties in this approach. The first problem comes in defining the frequencies, the second in defining the extent of the lines to be processed. Further conceptual problems occur in applying frequency analysis to the complex multivalued curves that constitute cartographic lines.

For some types of feature, within the dataset there are not enough line segments available within the average feature to allow any frequency analysis to take place without chaining consecutive features together. These include roads which are relatively easy to enchain, but the effect on the position of junctions must be considered.

The parameters of the frequencies involved, i.e. the 'amplitude' and 'wave length' are less easy to determine, simply because of the irregularities of the curves and the length of each that can be considered.

The analysis of frequencies was therefore carried over into tolerancing. In order to filter data representing lines by tolerance methods a number of parameters may be used, but they can be broken down into components of the angle subtended at a point by consecutive line segments meeting there and the distance between adjacent points.

The algorithm used for investigating line generalisation by tolerancing was based upon a review carried out by Ordnance Survey in 1979 (Ordnance Survey unpublished paper 1979) in which it was decided that the most effective algorithm for large scale map data should use the deviation angle and the distance from a baseline between fixed points. The advantage of this approach is that points greater than a certain distance from a based line (or trend line) between fixed points are retained, but in addition, points which have a deviation angle of greater than tolerance are also retained. By adjusting these two criteria, it was hoped to fulfil Rhind's requirement to select by 'frequency'.

5.5 Implémentation and experiments with a line filtering algorithm

The algorithm initially implemented was very similar to that described in the OS paper. It works on the basis of 'anchor points' which are fixed and output and 'floater' points which will be fixed in due course. All other points are tested either until they are found to exceed either of the preset tolerances, in which case they become floaters or they are rejected. As both ends of a feature must be fixed before filtering starts (to give the initial 'anchor' and 'floater'), the filter has to be carried out after digitising is completed.

The input to the algorithm is a vector containing a known number of co-ordinate pairs, and tolerances for the minimum offset distance and deviation angle for retention of a point. The meaning of these terms is illustrated in Figure 5.2

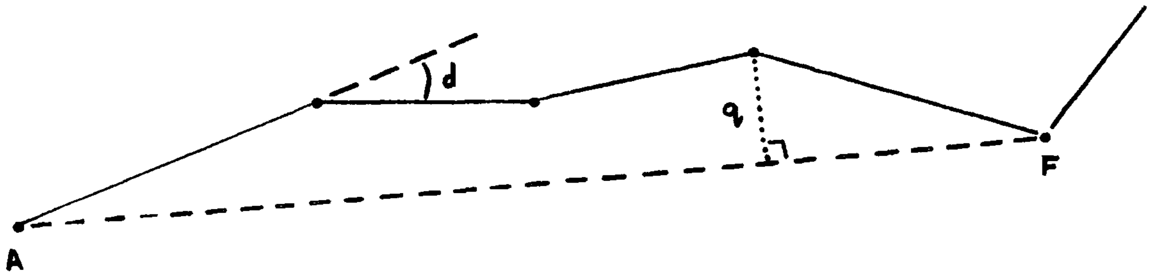


Fig. 5.2 Illustration of offset distance (q), deviation angle (d) and 'floater' and 'anchor' points.

The method of filtering data by this algorithm is described below. A stack is required to hold the index to points to be retained (but not yet output), and registers for the maximum offset value and position of the point at which this occurs. The feature has n co-ordinate pairs.

1. Assign 'first' point as anchor, nth as 'floater'.

2. a. Clear maximum value of offset distance, push 'floater' onto stack.
   If 'anchor' and 'floater' are adjacent then go on to 2c otherwide:
   b. From ('floater' - 1) to ('anchor' + 1): test each point for the offset distance from the current baseline (i.e. 'anchor' to 'floater'). While no point exceeds the tolerance, test the deviation angle at every point. If this exceeds the tolerance, then assign the current point as 'floater' and proceed from 2a.
   If the offset distance is greater than tolerance, and of the current maximum offset, record this value and the number of the current point.
   c. if the maximum offset is not zero proceed from 2d. Otherwise:
   output current 'anchor' point
   pop stack into anchor
   if the stack is empty then output the anchor and exit
   Otherwide:
   pop stack into 'floater' and proceed from 2a
   d. assign the 'furthest out' point as floater and proceed from 2a

The use of this algorithm poses two problems:

1. the definition of the vector to be a known number of co-ordinates
2. the derivation of the tolerances to be used

The first problem is data and machine-dependant. In the machines used, available memory was such that arrays could be dimensioned large enough to accommodate any likely feature. The features themselves were chained where appropriate - i.e. where two, and only two, features of the same type joined. The converse, of breaking features to filter, then recombining and subsequently refiltering was considered but not required. The program written could also deal with closed features.

The second problem is more complex and was approached in two ways. In the first, an analysis of the requirements of the data for the specific task of generating data for 1:250,000 scale mapping was considered. In the second, an empirical study of the results of using the filter was undertaken to determine how this compared with the analytical study.

The analysis took as a premise the (simplified) Ordnance Survey standard requirement for the accuracy of a traced line, or one scribed from a key to lie such that the centre of the new line is never more than one half of the width of the line from the centre of the key. Thus a 0.2mm line should never deviate by more than 0.1mm from the position of the original. This requirement can be directly compared with the offset tolerance of the filter algorithm by applying the line width for a particular feature type and the scale factor. For example, a 0.5mm wide line (e.g. a minor road on the 1:250,000 scale Ordnance Survey map) the offset tolerance could be set at 0.0005 x 250,000 = 62.5 metres.

This calculation is trivial and was easily implemented. An appropriate value for the deviation angle tolerance was more difficult to deduce.

Figure 5.3 illustrates the problem. A B C are fixed points, I is any point under consideration. The deviation angle is the angle between the extension of AB and BI. The maximum offset, q , is also shown. From table 5.1 it can be seen that the mean segment length for all lines was 28 metres and the detailed analysis at Appendix B shows that the segment lengths were positively skewed, indicating that the majority of lines are far shorter than an appropriate offset distance tolerance. This means that in most cases the deviation angle will be expected to control the points retained and points such as I and J in figure 5.3 will be retained because they have large deviation angles while points such as K and L will not.

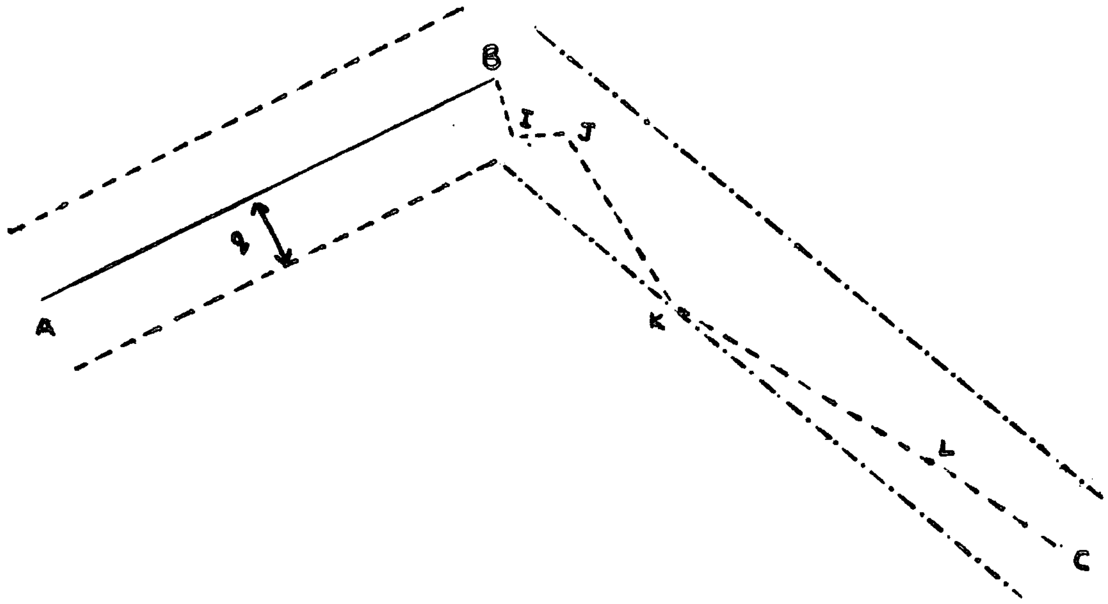


Figure 5.3

The problem posed by points such as J is relatively common in data derived from stream digitising or from some raster to vector conversion algorithms where the distances between points are of the order of the resolution either of the tablet, or the resulting transformed co-ordinates.

In Figure 5.4, the line AB is represented by points $P_I \ldots P_{I+n}$ and it is apparent that very few of these actually lie on the required line AB, and the resulting drawn line is irregular. When δ is of the order of 0.0005 inches and more (o.1mm approx) the effect can be very noticeable, even on modest plotters.

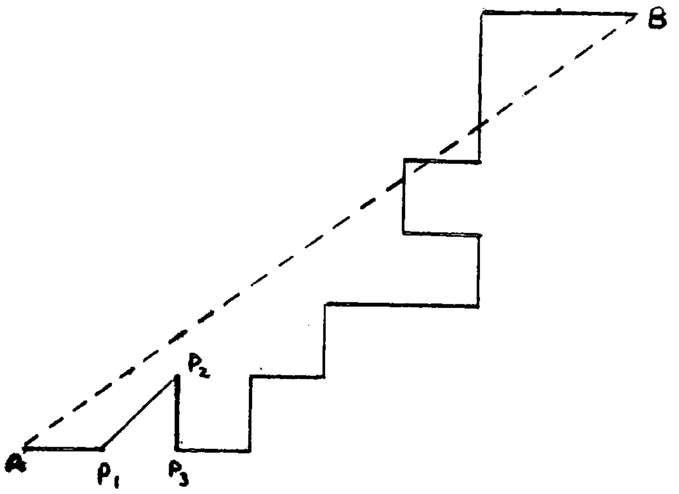


Figure 5.4 Example of the effect of co-ordinate resolution on a non-axial stream digitised line where δ is the increment of resolution.

More important for the present discussion, however, is the effect on the filter routine in which the angles subtended at stream digitised points may be consistently of the order of 90^{o}. This may be a constraint on the use of a filter based on angle tolerance, unless overcome by means which include the following:

1. by accepting the situation and either averaging over the whole length of a line, or smoothing prior to filtering.
2. by removing points which subtend angles of 90^{o} and treating them as a special case, using an average value where the interpoint distances are equal and small integer multiples of Cases where the interpoint distances are not equal but are small integer multiples of might also be treated in this way.
3. by passing the data through a tolerance filter as described, using a high angle tolerance (90^{o}) and a low offset tolerance distance (of the order

2δ). This will eliminate all points in the situation shown, but at the risk of straightening slightly curved lines.

By adopting course 1, problems could occur in a situation where a straight stretch joins a curve, the averaging perhaps displacing the line considerably, as shown in Fig. 5.5

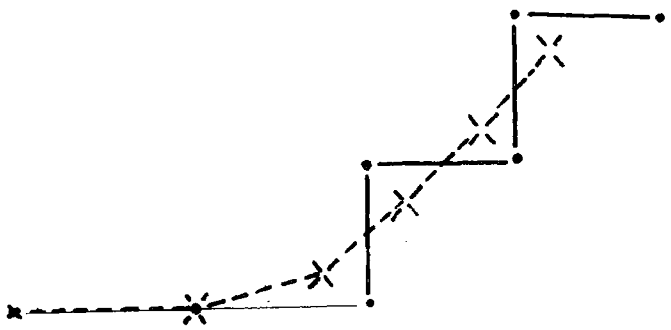


Figure 5.5 Example of curve distortion caused by using a running mean of three points

Courses (2) and (3) are related to the specific problem, and either could be used prior to filtering, but there is a penalty in processing time and a risk of losing acceptable data.

In testing, the data used was found to be free of this problem because it was generated by the McConalogue spline, and such techniques were not required.

The analytical solution to determine the value for the tolerance angle was therefore inconclusive in the case of these experimental data, because the mean segment length is less than the required offset distance, thus suggesting that a distance filter, of the type proposed by Douglas and Peucker (1973) might be all that is required.

In parallel with this analysis a series of empirical tests were undertaken using a subset of the data for one grid square, representing the 7315 co-ordinate pairs for 468 features. These data were filtered for 56 different combinations of tolerance. The results for these trials are tabulated in table 5.2, the percentage of retained points are displayed in figure 5.6. In table 5.2 the parameter N is the ratio between the number of retained points and the minimum possible number of retained points - i.e. twice the number of features. This index is a rough measure of sinuoisity, but of limited value because it is dependent on the digitising method, the draughtsman's technique and the spline routine as well as the sinuosity.

The graphs in figure 5.6 demonstrate the action of the filter routine. Taking into account the geometric progression (approximate) of the offset distances, the curves are asymptotic for each value of the deviation angle. Those for the angles of 45° and 90° are approaching the minimum possible percentage (12%) that could be retained.

Two further points should be considered in connection with these curves: the realistic precision of the source data and of the 'intelligence ' of the processing, i.e. the efficiency of the filtering in retaining significant points. The resolution of the souce data can be taken as that of the digitising tablet - 0.001 inch, or 1.27 metres at 1:50,000 scale (assuming that the source was not enlarged). The realistic precision of the data is likely to be of the order of the minimum linewidth of the source, i.e. 5 metres. This effectively means that offset distances of less than 5 metres are filters of noise rather than useable information.

The significance of points to be retained is less easy to determine, but as the points outside the noise level must be assumed to be significant for some purpose, the minimum deviation angle should be used to retain subtle changes in direction that form many curves.

For reduction in scale the noise level will be higher. To calculate the appropriate filter parameters, the precision can be calculated, or be preset according to the requirement for the data.

Table 5.2 - Example of the effect of various combinations of deviation angle and offset distance tolerance upon a data set consisting of 7315 points in 493 features of different types.

offset distance	2	5	10	20	40	80	160	320
0°	7.1 96%	7.1 96%	7.1 96%	7.1 96%	7.1 96%	7.1 96%	7.1 96%	7.1 96%
5°	5.1 69%	4.8 65%	4.7 64%	4.7 64%	4.7 64%	4.7 64%	4.7 64%	4.7 64%
10°	4.2 57%	3.6 49%	3.3 45%	3.1 42%	3.0 41%	3.0 41%	3.0 41%	3.0 41%
15°	4.0 54%	3.1 42%	2.7 36%	2.4 32%	2.3 31%	2.2 30%	2.1 28%	2.1 28%
22°	3.9 53%	3.0 41%	2.4 32%	2.1 28%	1.9 26%	1.7 23%	1.7 23%	1.6 22%
45°	3.8 51%	2.9 39%	2.3 31%	1.9 26%	1.6 22%	1.4 19%	1.3 18%	1.2 16%
90°	3.8 51%	2.9 39%	2.3 31%	1.9 26%	1.5 20%	1.3 18%	1.2 16%	1.1 15%

Top value - N index, where N is the ratio between the points retained and the minimum possible number of points retained. In this case:

$$7.4 > N > 1.0$$

The minimum possible number of points retained is twice the number of features - in this example, 986

Lower value - percentage of points retained.

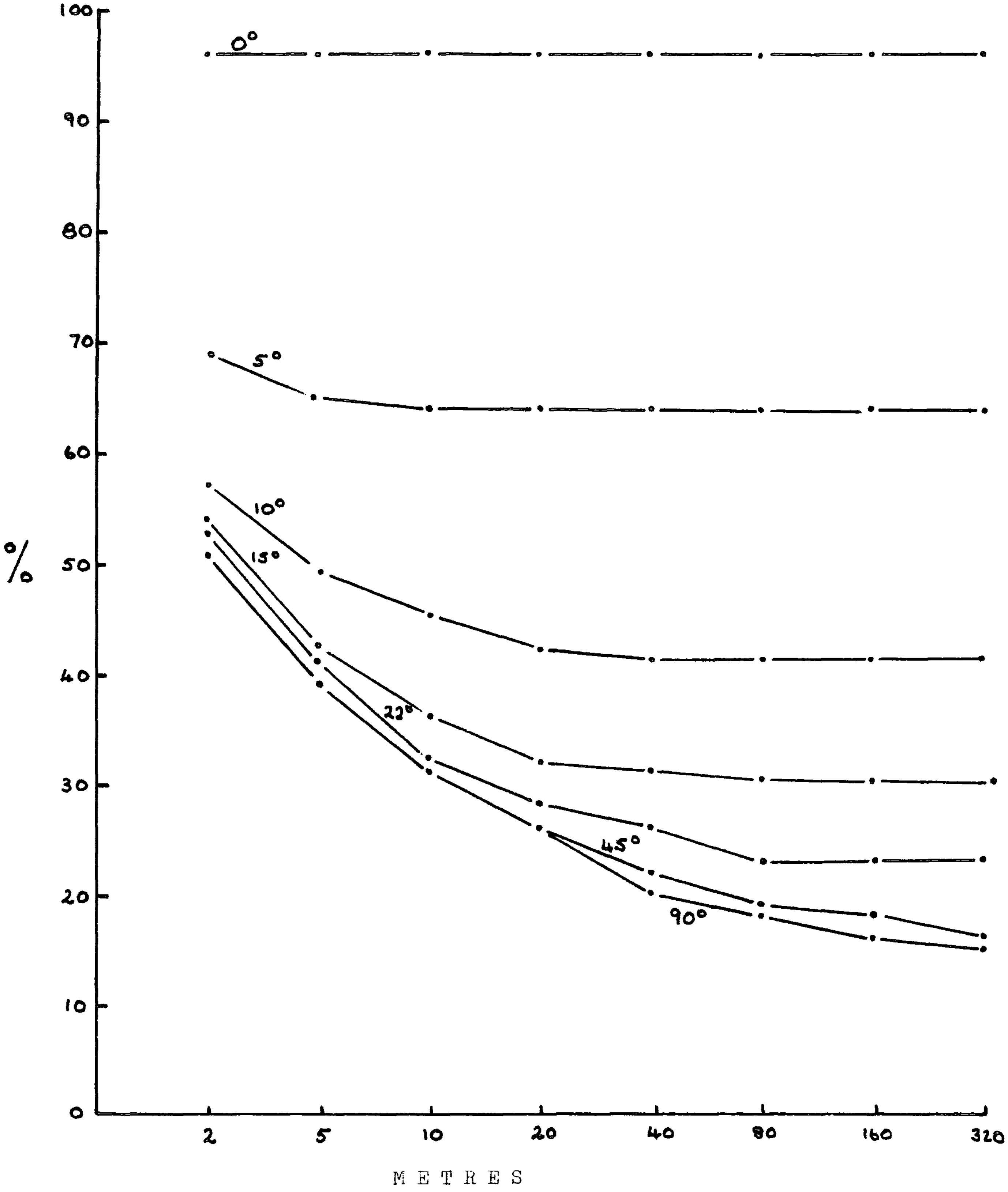


Figure 5.6 Graphs of percentage of points retained for given deviation angles for a series of offset distances.

N.b. The horizontal scale is non-linear; the offset distances form a near geometric progression.

and the appropriate value for the deviation angle can be estimated from the curve which is asymptotic from that point. Thus for a required maximum offset of 25 metres the most efficient deviation angle would be 10°.

In practice, empirical derivation of the parameters could not be efficient because of the processing involved to produce the curves, and analysis of the data, or a sample of it, is more likely to be used. Although for many cartographic processes the deviation angle is important in determining the appearance of the final map, for non-cartographic purposes the known maximum positional uncertainty of the offset distance will be of use. The ability to filter data and know that the result is not going to cause gross geometric distortion is vital for all forms of automated cartometry. However, the algorithm used in this experiment was biased towards the cartographic filter, because of the way in which the deviation angle tripped the recursion, and other filters will have to be devised for other purposes.

5.6 Extensions and modifications to the line filter algorithm

The Ordnance Survey/Rand filter algorithm as implemented has a number of weaknesses. It does not consider the sign of the deviation angle and hence cannot detect inflexions which may be highly significant. Some noise cannot be distinguished from valid data except, perhaps, by undertaking multiple passes. The tolerances are static - they are fixed at the beginning of the run in the current program for all features, but could be set differently for separate types of features or an iterative method could be used.

A number of modifications and extensions to the basic algorithm were considered. To investigate the detection of inflexions and noise, a few small linear features were broken down into a 'traverse array' in which the co-ordinate pairs were replaced by the deviation angle and interpoint distance. An example of such an array is shown at table 5.3, together with comments about the function of each point. Using this table and a large scale plot (Figure 5.7 a) an attempt was made to analyse the geometry of each feature and from this determine how generalisation might be approached.

Table 5.3 - traverse array for statistical filter

Point number	Bearing	Deviation angle	Interpoint distance	Remarks
1		-		start point
	142		31	
2		+56		turn
	198		183	
3		0		redundant
	198		3	
4		-7		noise?
	191		5	
5		0		redundant
	191		5	
6		0		redundant
	191		5	
7		+7		noise?
	198		3	
8		-18		noise?
	180		2	
9		+18		noise?
	198		6	
10		-47		turn
	161		6	
11		-43		
	124		11	
12		-29		
	95		12	
13		-5		
	90		2	
14		+11		noise? turn?
	101		5	
15		-16		turn?
	85		12	
16		-40		turn? noise?
	45		1	
17		-18		noise?
	27		2	
18		-7		
	34		4	
19		0		redundant
	34		4	
20		+3		noise?
	37		5	
21		0		redundant?
	37		5	
22		-18		
	19		147	
23		+60		turn
	79		121	
24		-		end point

The above example is of a feature representing a dock railway i.e. a fairly simple rectilinear object. The long parallel sides are between points 2 - 10 and ?18- 23. One end is open

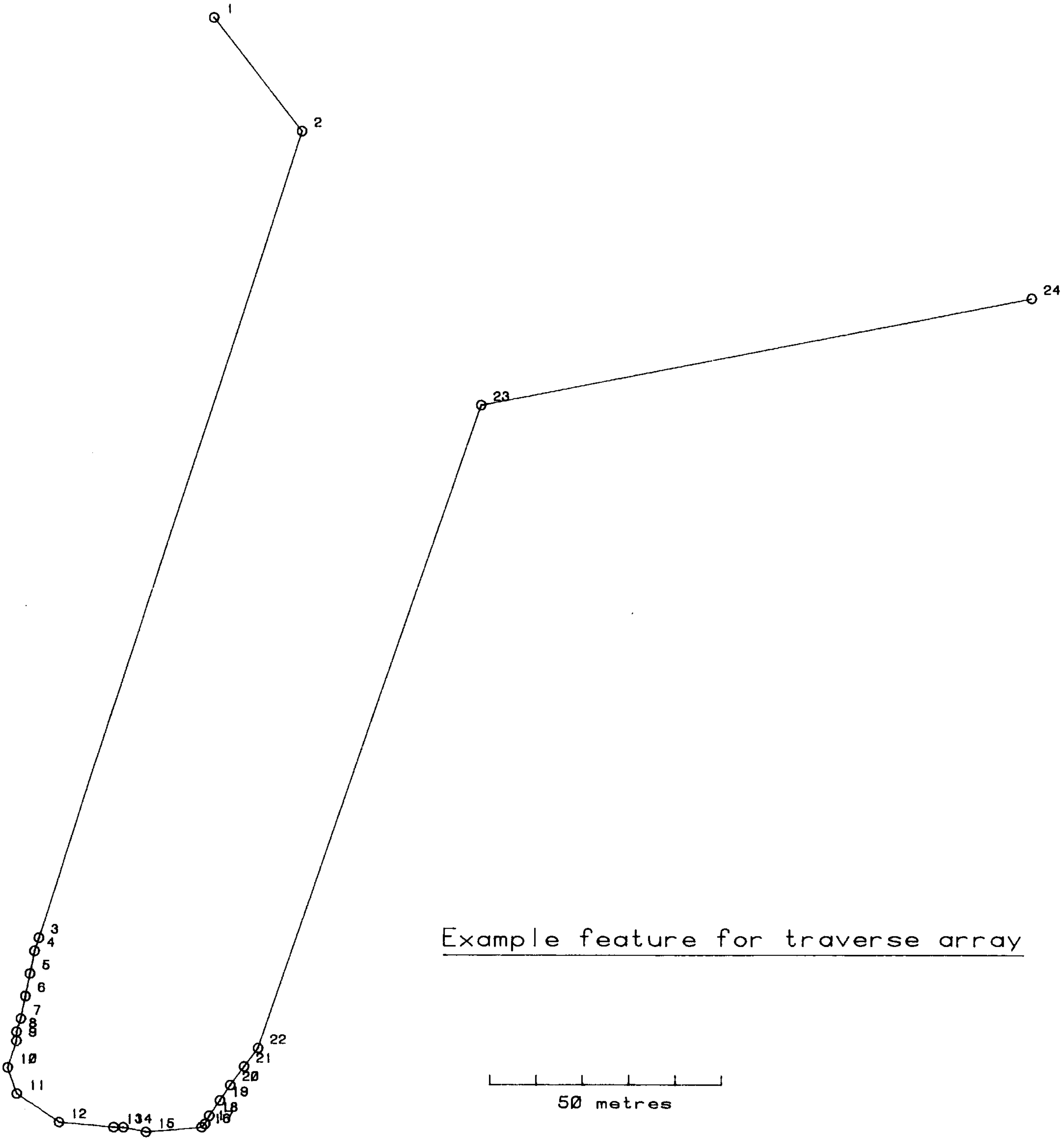
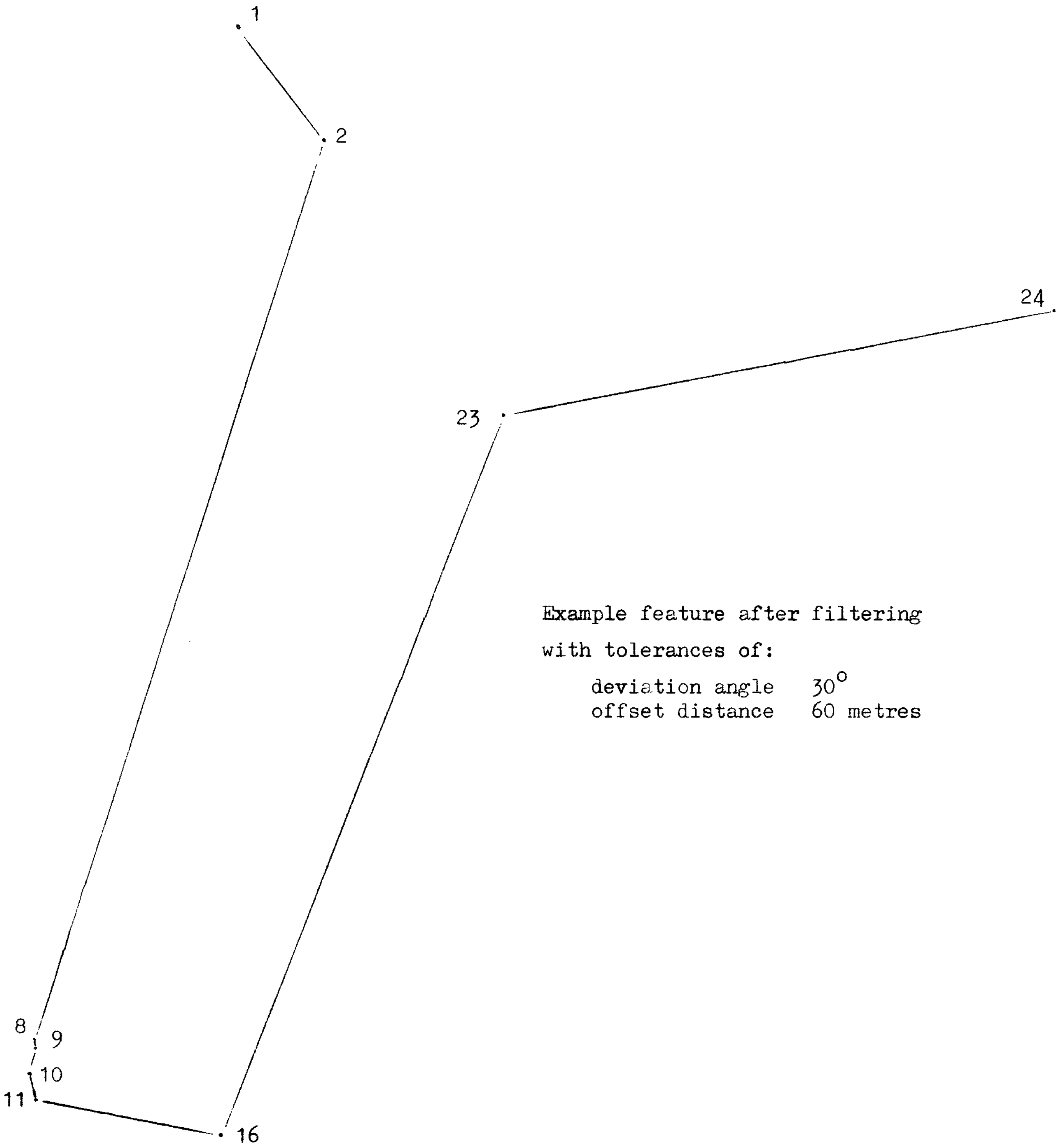

Example feature for traverse array
50 metres
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Figure 5.7 (a)



Example feature after filtering
with tolerances of:

deviation angle 30°
offset distance 60 metres

Figure 5.7 (b)

The first evidence of the example derived from the feature code, is that it is a railway-suggesting that curves should be smooth, with deviation angles generally less than 90°, a fact confirmed by the table. If the feature has been digitised using a Mconologue spline, then a series of near equal interpoint distances is to be expected, typically in a group with low variance for each arc of the generated curve. This is evident between points 4, 9 and 11 and inconclusively so between points 17 and 22.

If a filter routine of the type implemented is applied to this feature, the results are as seen in Figure 5.7, with tolerances of 60 metres and 30°. Clearly the shape is preserved, but the filter has retained points that may well be noise - in preference to points that have more claim to be turning points.

The first stage of these extensions to the filter was to add a condition to the test for deviation angle to the effect that the distance to the next point should be greater than a preset tolerance. The value of this tolerance is, however, extremely critical. In the case of the feature described above, a tolerance of greater than 12 metres results in a triangle, but one of two metres results in the removal of only one extra point.

For the case of generalisation from 1:50,000 to 1:25,000 these arguments are probably unneccessary for this particular feature because a satisfactory result would be obtained by determining the co-ordinates of the intercept between all lines greater than 10 metres, which could be used as a means of reducing the effect of 'noise'.

This examination of features opens up a very wide range of possibilities for generalisation and filtering, but each feature exhibits its own problems. The most important conclusion is that, before analysing data for generalisation purposes, it is essential to know the history of the data - the precision, methods and processing used in its capture. The analysis of data for the example feature might have been totally different had the facts not been known that a McConalogue spline been available, and that the alternative was 'point made' digitising.

The attempt to detect inflexions was based on the sign of the deviation angle after filtering of the insignificant shorter segments. Having detected an inflexion the question remains as to what to do about it. A simple example is a phase of a sinusoidal curve, as in figure 5.8

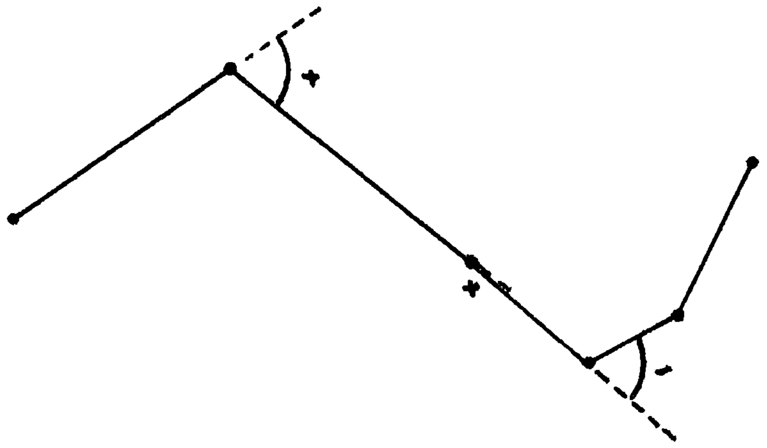


Figure 5.8

Using a clockwise positive convention for deviation angles, the feature can be logically represented as + + - - (working from either end), and an inflexion must exist. One of the points, or a new, generated point generated between the points at the change of sign can be designated a 'floater' and the filter run from there. However, in the situation illustrated, the curves would not be represented if the offset distances and deviation angles are less than tolerance. The answer would appear to be to retain, regardless of tolerance, the point with maximum offset on either side of an inflexion, together with a point representing the inflexion. The justification for this lies in that the slight curve produced will satisfy the requirement for maintaining the character of a line, and allow the line to be distinguished from a straight one.

The detection of unwanted or moved line intersections is outside the scope of the algorithm under consideration and has been discussed in an earlier chapter. The importance of this was, however, seen in a number of cases where gross geometric distortion occurred. Examples can be seen on the plots in Appendix C.

5.7 Conclusions

The results of the experiments carried out, and the extensions proposed are in confirmation that a relatively simple algorithm using basic properties of the geometry of lines can be used to simplify the geometry of those lines in such a way that the quantity of data is reduced and the majority of 'significant' points are retained. How many significant points were lost is difficult to estimate, as the definition of a point's significance is obscure, dependant upon the feature type, its method of digitising, the number of points, feature chaining and so on, so no 'efficiency' index for this particular algorithm can be determined. The experiment's major use was in the handling of spatial data and the insemination of ideas as to possible structures and content of data sets designed for map generalisation.

The data used partially satisfied the requirements for data identified in chapter 4. There was partial rational coding of the data, but only a simple attribute coding system; indirect representation of administrative areas only and no information about logical connection or separation of features. The only text was related to administrative areas and it did act as a modifier to the boundary to which it referred.

CHAPTER 6

Conclusions

The study for this dissertation was divided between the theoretical considerations described in Chapters 2 and 4 and the assessment of production procedures in Chapter 3 which were undertaken concurrently with the experimental work described in Chapter 5. All three main tasks were closely interwoven, the practical work outlining inadequacies of the present methods of data presentation, the theoretical work and examination of current maps and procedures providing the 'ideal' solution in cartographic terms.

It is believed that automated cartographic generalisation and generalisation for information systems use are related, but that the former requires considerably more processing effort and many of the processes are outside the capability of current data processing systems. Manufacturers of cartographic data processing systems using vector methods (for example Computervision, Kongsberg, Ferranti, Wild, among others) claim to aid generalisation but the emphasis is on aiding interactive generalisation, transferring the need to draw on paper or plastic to 'drawing' on a cathode ray tube. Even these systems, however are designed to deal with relatively small quantities of data at one time. None of the systems are yet capable of handling the quantities of data that are envisaged in a data set derived from 1:50,000 scale maps in a time frame that would allow rapid automated compilation of maps covering the whole of the United Kingdom at one time. The data quantities are not excessive by current standards at 800 million bytes (Haywood, forthcoming) but sequential processing is a highly inefficient method of treating some aspects of two dimensional data (Adams, forthcoming) and 'two dimensional' computing of the type used in array processors may allow new techniques of generalisation to be devised to allow genuine automated compilation.

For small data-sets, sequential processing is viable and the experiments carried out in this study confirmed that a line filtering algorithm can be used on network data, but with greater information encoded into the source data, more 'intelligent' data reduction could have been undertaken. Against this, the extra information would add considerably to the data set size which might in turn, slow down processing.

For the future, the widely predicted fall in relative costs of storage and processing of digital data will undoubtedly bring closer the possibility of complete automation of spatial generalisation, but it is thought that for many years to come the real map, as displayed on paper or tube will retain at least some human input.

The work for this dissertation proved to be very useful to the author in giving the opportunity to think about map generalisation both in conceptual and practical terms, ranging from the theoretical aspects of the various processes and how they might be implemented to the effects of the method of data capture upon one aspect of line generalisation.

REFERENCES

ADAMS T.A. (1979) Characteristics of a national topographic digital database. Unpublished MSc dissertation, University of Durham

ADAMS T.A. (forthcoming) Doctoral theseis on the use of raster encoding methods for manipulating Ordnance Survey digital map data

ADAMS T.A. and RHIND D.W. (1981) The projected characteristics of a national topographic digital databank. Geog. Journ. Vol 147 No 1

BAUGH I.D.M. (1975) Accuracy in geographical measurements Orb - Aberdeen University Geog. Soc. Vol VI No 1

BAUGH I.D.M. and BOREHAM J.R. (1976) Measuring the coastline from maps: a study of the Scottish Mainland. Carto Journ Vol 13

BAXTER R.S. (1976) Computer and statistical techniques for planners Methuen, London

BAXTER R.S. (1977) Data structure problems in interactive manipulations and graphical display of nested zone thematic maps. In DUTTON (1977) ed.

BOYLE R. (1970) The quantised line. Carto. Journ. Vol 7 No.2

BRASSEL K. (1977) A topological data structure for multi-element map processing. In DUTTON (1977) ed.

BREWARD R.W. (1972) A mathematical approach to the storage of digitised contours Carto. Journ. Vol 10 No 2

COX N.J. and RHIND D.W. (1978) Geographical networks and automated data processing. University of Durham Dept. of Geography Occasional Paper No. 11

DOUGLAS and PEUCKER T.K. (1973) Algorithms for the reduction of the number of points required to represent a digitised line or its caricature. Canadian Cartog. Vol 10 No 2

DREWETT B. (1973) The changing face of the map user in Britain. Carto. Journ. Vol 10 No 2

DUTTON G.H. (1977) ed. Proceedings, First International Advanced Symposium on Topological Data Structures for Geographical Information Systems (7 Volumes) Addison-Wesley, New York

GOTTSCHALKE H-J. (1972) Derivation of a measure for the diminished content of information of cartographic lines by means of a gliding arithmetic mean Nach. aus dem Karto. und Vermess. Series II Vol 30

GREGGOR K.N. (1975) Computer aided map compilation. Cartography Vol 9 No 1

HAKE G. (1975) Zum Begriffsystem der Generalisierung. Commemorative volume, Nach. aus dem Karto. und Vermess. (cited in LICHTNER (1979))

HARLEY J.B. (1975) Ordnance Survey maps - a descriptive manual H.M.S.O. London

HAYWOOD P.E.H.. (1981) The size of a 1:50,00 scale national topographic database and some considerations thereon. Unpublished dissertation for the degree of M App Sc, University of Glasgow

H.M.S.O. (1979) Report of the Ordnance Survey Review Committee Her Majesty's Stationery Office, London

JOHANNSEN T. (1973) A program for editing and for some generalisation operations (for derivation of a small scale map from digitised data in 1:50000) Nach. aus dem Karto. und Vermess. Series II Vol 30

KADMON N (1972) Automated selection of settlements in map generalisation Carto Journ. Vol 9 No 2

KEATES J.C. (1972) Symbols and meaning in topographic maps Int. Yearbook of Carto. 1972 p 168-182

KEIR K.M. and MOTT P.G. (1976) A digital mapping system designed for the commercial air survey market. Proc. International Symposium on Photogrammetry, Helsinki

KISHIMOTO M (1967) Cartometric measurements Geog. Instit. der Univ. Zurich (reviewd in MALING (1968))

KOEMAN C. and Van der WEIDEN F.L. (1970) The application of computation and automatic drawing instruments to structural generalisation. Carto Journ. Vol 7 No 2

LANG T. (1969) Rules for robot draughtsmen Geog. Mag Vol XLII No 1 Oct. 1969

LICHTNER W. (1979) Computer assisted processes of cartographic generalisation in topographic maps Geoprocessing Vol 1 No 2

MALING D.H. (1968) How long is a piece of string? Carto Journ. Vol 5 No 2.

MANDELBROT B. (1967) How long is the coast of Britain? Statistical Self similarity and fractional dimension. Science 155 p 63 p 636-638

MANDELBROT B.B. (1977) Fractals: Form, chance and dimension W.H. Freeman, San Francisco

MARK D.M. (1979) Phenomenon based data structuring and digital terrain modelling Geoprocessing Vol 1 No 1

MOELLERING H. (1980) Strategies of real time cartography
Carto Journ. Vol 17 No 1

MORRISON A. (1971) Experimental maps of road travel
Carto Journ. Vol 8 No 2

ORDNANCE SURVEY - Undated, unpublished draughtsmen's guides
to the specification of the 1:50,000 and 1:250000 scale maps

ORDNANCE SURVEY - Undated, unpublished report on the filtering
of cartographic lines in digital form ?1979

ORDNANCE SURVEY - Annual Reports, 1975 - 1980, published by
Her Majesty's Stationery Office

ORDNANCE SURVEY - Reply to article criticising removal of
woodland symbols in OS News Jan. 1980

ORDNANCE SURVEY (1980) Unpublished internal report on the OS
input to the RLUIS project.

PAGE G. (1978) Filtering of digitized line features.
Proc. Digital Terrain Models Symposium, Am. Soc. Photogram.
St Louis, May 1978

PEUQUET D. (1979) Raster processing: an alternative appraoch to
automated cartographic data handling. American Cartog.
Vol 6 No 2

RHIND D.W. (1973) Generalisation and realism within automated
cartographic systems Canadian Cartog. Vol 10 No 1

RHIND D.W. (1976) Towards universal, intelligent and useable
automated cartographic systems. ITC Journal 1976 No 3.

RHIND D.W. and HUDSON R. (1981) Land Use. Methuen London

ROBINSON A.M. and SALE R.D. (1969 3rd ed., 1978 4th. ed.) Elements
of Cartography. John Wiley and Sons, New York

ROYAL SOCIETY (1978) Evidence submitted to the Ordnance Survey
Review Committee

SHARPLEY W.K., LIESERSON J.F. and SCHMIDT W. (1978) A unified
appraoch to mapping, charting and geodesy (M,C and G) database
structure design. US Army Engineer Topographic Laboratories
Report No ETL - 0144

SHIRYAYEV Y.Y. (1977) A possible way of computer generalisation of linear objects using a river network as an example. Geodesy, Mapping and Photogrammetry Vol 18 No 2

STENHOUSE H. (1979) Selection of towns on derived maps Carto. Journal. Vol 16 No 1

STEWARD H.J. (1974) Cartographic generalisation: some concepts and explanation. Cartographica Monograph No. 10

THOMPSON C.N. (1978) Digital mapping in the Ordnance Survey, 1968-1978. Paper presented at the International Soc. of Photogramm. Ottawa

TOPFER F. and PILLEWIZER W. (1966) The principles of selection Carto. Journ. Vol 3 No 1

TOPFER F. (1974) Kartographische Generalisierung. VEB Herman Haack Geographische-Kartographische Anstalt. Gotha/Leipzig

UITERMARK (1980) Generalisation within automated cartographic systems. Report (in Dutch) to Netherlands Soil Institute and National Physical Planning Agency.

U.S.G.S. (1980) Unpublished report on the data structures of the Digital Line Graph system.

VANICEK P. and WOOLNOUGH D.F. (1975) Reduction of linear cartographic data based on a generation of pseudo hyperbolae. Carto Journ. Vol 12 No 2

APPENDIX A. Occurrence of Features shown on 1:50,000 and 1:250,000 scale Ordnance Survey maps by type

P = Point
L = Line
A = Area
N = Named

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
<u>Antiquities</u>									
Roman	✓	✓	✓	✓		✓		✓	
Non-Roman	✓	✓	✓	✓					
Other	✓			✓	✓			✓	
Battlefield	✓			✓	✓			✓	
Milecastle: Turret				✓					
Cairn	✓			✓					
<u>Boundaries</u>									
National		✓				✓			
County		✓				✓			
District		✓							
Parish		✓							Outline only
LB		✓							
Forest Park		✓							
National Park		✓							
NT (always open)		✓							
NT (opening restricted)		✓							
New Forest		✓							
Country Park				✓					
<u>Communications</u> <u>Air</u>									
Airport, Aerodrome		✓		✓	✓	✓			
Heliport	✓				✓	✓			
Landing Strip		✓		✓					

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
<u>Communications (contd)</u>									
<u>Railways</u>									
Standard Gauge multiple / single		✓				✓			
Narrow gauge		✓							
Freight/sidings/tramway		✓							
Private		✓		✓					
Principal Station	✓	✓			✓			✓	
other	✓				✓				
closed	✓								
Tunnel		✓		✓		✓			
Viaduct		✓							
Level crossing	✓			✓	✓				
Railway Bridge		✓							
<u>Roads</u>									
Bus Station	✓								
Danger Areas			✓	✓					No bounding lines
Ford		✓		✓					
Gradient 1:5	✓				✓				
do. 1:7	✓				✓				
Road numbers				✓				✓	
Rights of way		✓							
Bridleway		✓							
Footpath		✓							
Road		✓							
Road bridge		✓		✓		✓		✓	
Tunnel		✓		✓		✓			
Footbridge		✓							
Subway		✓							
Toll	✓			✓		✓			

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
Communications (contd)									
Roads									
Motorway dual CW		✓				✓			
junction		✓				✓			
single CW		✓							
Elevated roads		✓							
Trunk/main roads									
dual CW		✓				✓			
single CW		✓				✓			
narrow		✓							coloured only
Unfenced roads		✓							
Secondary roads									
dual CW		✓				✓			
single CW		✓				✓			
narrow		✓				✓			coloured only
Minor roads tarred		✓							
dual CW		✓				✓			
narrow		✓							
untarred etc		✓							coloured only
Paths (non ROW)		✓							
Single CW roads under const.		✓							
Dual CW do.		✓							
Motorways do.		✓							
Motorway Service Areas	✓				✓				
Road viaduct		✓							
Roundabouts	✓				✓				
do. (large)		✓				✓			
Water									
Beacon - coastal	✓								
Canal full		✓				✓			
dry		✓							
ship		✓				✓			
Docks				✓					

FEATURE	1:50,000 P	L	A	N	1:250,000 P	L	A	N
Communications contd								
Water								
Ferry vehicle		✓		✓		✓		
passenger		✓		✓		✓		
continental		✓		✓		✓		
Lighthouse used	✓				✓			
disused	✓							
Locks	✓							
Rivers > 8m wide		✓				✓		
8m wide		✓						
Weir	✓			✓				
Culture								
Abbey		✓		✓				✓
Aquaduct		✓						
Artillery range				✓				
Barracks/camp				✓				
Aerial ropeway		✓		✓				
Beerhouse/public	✓	✓		✓				
house	✓	✓		✓				
Hotel	✓			✓				
Motel	✓			✓				
Inn	✓			✓				
Breakwater		✓				✓		
Building groups			✓				✓	
Buildings isolated	✓				✓			
public		✓		✓				
Cable way		✓						
Cairns	✓			✓				
Cattle grid	✓			✓				
Cemetaries/ crematoria				✓				
Churches spired	✓							
towered	✓							
no spire/ tower	✓							
ruined	✓			✓				

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
Culture contd.									
Chimney	√			√					
CGCr station	/√			√					
Clay pit				√					
College				√					
Conduit		√							
do. open		√							
Cricket ground				√					
Cutting/embankment		√							
Dam		√		√					
Deer Park				√					
Dump			√	√					
Elec. Gen Stat.				√					
" sub stat.									
ELT		√							
Field study centre				√					
Factory/works				√					
Fences		√							around woods etc
Firebreaks		√							
Fire tower				√					
Football ground				√					
Forest Park									
Garage				√					
Gasworks				√					
Gate	√								
Glasshouses			√						
Gravel pits				√					
Golf course/links	√								
Govt. office				√					
Groynes		√							
Hill figures	√			√					
Hospital				√					
Hostel	√								
Law Court				√					
Obelisk	√			√					

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
Culture contd.									
Tower	✓			✓					
Lifeboat Station	✓			✓					
Inshore Rescue boat	✓			✓					
Lodge				✓					
Milepost	✓			✓					
Mill				✓					
Mine				✓					
Mineshaft	✓			✓					
Moat		✓		✓					
Monument	✓			✓					
Mountain Rescuepost	✓			✓					
Municipal Office		✓		✓					
Museum		✓		✓					
Nature Reserves				✓					
Nuclear power Station					✓			✓	
Observatory		✓		✓					
Oil Refinery	✓	✓		✓				✓	
Open pit				✓					
Parks & ornamental grounds			✓						Minimum 8h at 1:50,000
Pier		✓							
Pipeline (above ground)		✓							
Post Office	✓								
Prison, Borstal		✓		✓					
Public convenience				✓					
Public House				✓					
Pumping Station									
Quarry			✓						
" disused			✓						
Quay		✓							
Race course		✓		✓					
Recreation ground				✓					
Refuse, spoil, dump			✓						
Rifle range		✓		✓					

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
Culture contd.									
Royal Palace				✓					
Ruin				✓					
School		✓		✓					
Sewage works				✓					
Shielings	✓			✓					
Shooting ride									as paths
Sports ground				✓					
Ski/chair/lift/tow	✓	✓		✓					
Sluice		✓		✓					
Tanks	✓								
Telephone Call Box	✓				✓				AA&RAC blue GPO Black at 1:50,000
Television mast	✓				✓				
Town Hall		✓		✓					
Tunnel		✓		✓					
University		✓		✓					
Water works				✓					
Windmill - with sails	✓				✓				
" less "	✓								
Wind pump	✓								
Works				✓					
Zoological garden				✓					

Horizontal Control

Graticule intersections									1:50,000 5'intervals at
"									30' " 1:250,000
Grid									1km at 1:50,000
									10 km at 1:250,000
Triangulation pillar	✓								

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
Hydrography									
Drains		✓							
Dykes		✓							
High Water mark		✓				✓			
Lakes, ponds		✓	✓	✓		✓	✓	✓	
reservoirs		✓	✓	✓		✓	✓	✓	
covered do.				✓					
Low Water mark		✓				✓			edge of foreshore
Rivers/streams		✓		✓		✓		✓	
Saltings marshes			✓						
Waterfall	✓			✓	✓			✓	
Wells & springs	✓			✓					
water areas			✓						
Topography									
Beacons	✓			✓					
Caves	✓			✓					
Cliffs			✓						
Coastline		✓							
Dunes			✓						
Flat rock			✓						
Pot holes	✓								
Sand, mud, shingle			✓						
Slopes			✓						
Rock outcrops			✓						
Vegetation									
Orchard			✓						
Woods & forests			✓						
Firebreaks		✓							
Osiers			✓						

FEATURE	1:50,000				1:250,000				
	P	L	A	N	P	L	A	N	
Vertical control									
Bathymetric contours		✓							VI 50m
Contours		✓							VI 10m/200feet
Hill shading							✓		
Vertical layers									
Spot heights	✓			✓	✓			✓	
Tourist Information									
Beauty spot				✓					
Place of historic interest				✓					
Historic House				✓					
Ancient Monument				✓					
Camp site	✓								
Caravan site	✓								
Information centre	✓								
Picnic site	✓								
Parking area	✓								in rural areas only
Viewpoint	✓								

Source: Ordnance Survey Draughtsman's Manuals

Appendix B DETAILED STATISTICAL ANALYSIS OF SOURCE DATA

The analyses whose results are given in Appendix B were undertaken using a program written by T.A.Adams. Each table represents the results for one data set which includes major road, railway and water information for a 10 kilometre square in Fife. The results for dataset number 5 have not been included; there were only 6 co-ordinate pairs in this data set.

FEATURE CODE DESCRIPTION

FEATURE CODE	FEATURE TYPE	DESCRIPTION
7	LINE	BOUNDARY – PARISH OR COMMUNITY
8	LINE	BOUNDARY · DISTRICT
9	LINE	BOUNDARY – COUNTY OR REGION
15	LINE	RAILWAY – STANDARD GUAGE
19	LINE	RAILWAY – DISMANTLED CENTRE LINE
47	LINE	ELECTRICITY TRANSMISSION LINE
59	LINE	BANK OF DOUBLE RIVER / STREAM
62	LINE	BANK OF LAKE / POND
64	LINE	SINGLE STREAM
66	LINE	CENTRE LINE OF DOUBLE WATER FEATURE
71	LINE	MEAN HIGH WATER (MHWS)
89	LINE	CL TRUNK/MAIN DUAL CARRIAGEWAY
90	LINE	CL TRUNK/MAIN SINGLE CARRIAGEWAY
92	LINE	CL SECONDARY DUAL CARRIAGEWAY
93	LINE	CL SECONDARY SINGLE CARRIAGEWAY
96	LINE	CL MINOR SINGLE CARRIAGEWAY MORE 4 M
97	LINE	CL MINOR SINGLE CARRIAGEWAY LESS 4 M
98	LINE	CL MINOR OTHER ROADS
102	LINE	CL RAILWAY – SIDINGS
141	LINE	PATH / TOWPATH
155	POINT	STANDARD ROUNDABOUT
158	LINE	ORNAMENTAL PARK BOUNDARY (NON – ROAD)
159	LINE	ORNAMENTAL PARK BOUNDARY (ROAD)
193	LINE	LIMIT OF SAND / MUD / SHINGLE

```
                                                                                                                                   1.
         FEATURES INCLUDED
         =================

                                                                        GROUND                                     GROUND
                     NO.      NO.      TOTAL LINE      GROUND      MEAN   STANDARD
F/CODE    FREQ.      PTS .   LINES    LENGTH (MM)     DIST (M)     DIST   DEVIATION      SKEW     KURTOSIS     MAXIMUM     MINIMUM

    7      11.       948.     938.        910.0      45500.75     48.5     69.135      2.5612       7.579      464.82        1.00
    8       5.        0.
    9       5.        0.
   15      14.       198.     184.        131.4       6570.74     35.7     33.253      3.6488      16.980      255.93        1.00
   19      11.       279.     268.        209.7      10484.54     39.1     41.846      3.2265      11.918      296.41        1.00
   47       4.        12.       8.        214.5      10725.52   1340.7    357.342     -0.3004      -1.643     1736.61      799.60
   59      14.      1150.    1136.        513.8      25691.70     22.6     25.389      2.5861       7.919      198.86        1.00
   62       9.       331.     322.         75.4       3770.14     11.7     20.490      4.4413      22.920      176.50        1.00
   64      52.      1414.    1362.        928.6      46428.57     34.1     48.274      3.2872      14.065      410.57        1.00
   66       8.       555.     547.        260.9      13043.84     23.8     28.822      3.6399      20.838      306.15        0.94
   71       3.       151.     148.         80.1       4002.98     27.0     31.462      3.0949      12.008      214.21        1.00
   90      53.       339.     286.        424.5      21226.34     74.2    107.710      2.5819       8.238      695.97        1.00
   93      27.       121.      94.        170.5       8524.26     90.7     98.398      1.3332       1.296      438.55        1.00
   96      54.       351.     297.        453.8      22692.04     76.4     89.053      1.6064       1.852      385.75        1.00
   97       7.        42.      35.         54.7       2733.75     78.1     78.538      1.4766       2.142      348.53        2.24
   99     188.      1148.     960.       1396.0      69801.19     72.7     83.757      1.8792       3.980      529.14        1.00
  102       9.        60.      51.         41.1       2052.80     40.3     37.261      1.1238       0.260      137.93        1.00
  141      30.       216.     186.        326.6      16328.22     87.8    102.399      1.8656       3.897      545.37        2.24
  155       1.        1.
  193       2.        0.


A TOTAL NUMBER OF      402.  CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0.  GND.  METRES

TOTAL INKED IN LINES  =        6822.

TOTAL INVISIBLE LINES =           0.

TOTAL LINES GENERATED =        6822.

TOTAL DISTANCE GENERATED BY LINES =       309399.88 METRES AT GROUND SCALE
                                  =          618.800 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =          11063.

NUMBER OF POINTS IN THE FILE  =           7316.

NUMBER OF IGNORED CODES (0 OR >271)    =         1.

NUMBER OF IGNORED RECORDS (0 OR >271)  =        24.
```

```
FEATURES INCLUDED
=================

                                                                        ------GROUND------                                       ------GROUND--
                      NO.        NO.      TOTAL LINE        GROUND      MEAN    STANDARD
F/CODE     FREQ.      PTS       LINES     LENGTH (MM)       DIST (M)    DIST    DEVIATION       SKEW      KURTOSIS      MAXIMUM       MINIMUM
------     -----    -------    -------    -----------     ----------   ------   ---------
   7        10.      579.       569.          492.6        24632.22     43.3     63.097        2.6716        8.723       426.30          1.00
   8         7.        0.
   9         7.        0.
  15        17.      240.       223.          211.1        10555.74     47.3     77.954        5.4858       39.088       770.88          1.00
  19        28.      365.       337.          318.4        15918.82     47.2     60.512        2.8017        8.225       379.38          2.00
  47         5.       46.        41.          306.9        15344.95    374.3    193.054       -0.1543       -1.111       775.82         27.23
  59         4.      670.       666.          213.5        10675.00     16.0     18.245        2.6119        7.666       121.49          1.00
  62         9.      332.       323.           35.8         1787.85      5.5      7.397        4.9392       30.602        71.37          1.00
  64         6.      285.       279.          160.6         8030.19     28.8     55.491        3.9386       16.862       368.44          1.00
  66         1.      402.       401.          106.6         5327.91     13.3     14.944        3.1083       12.151       104.09          1.00
  71         3.      473.       470.          219.1        10955.56     23.3     34.849        4.3622       24.861       335.88          0.94
  90        46.      291.       245.          354.5        17726.89     72.4    112.312        2.4638        6.514       632.19          1.00
  93        26.      161.       135.          184.3         9213.68     68.2     81.721        1.6060        1.501       332.44          3.61
  96        14.      162.       148.          131.0         6548.37     44.2     71.813        2.8251        8.231       409.68          1.00
  98       110.      740.       630.          745.1        37253.93     59.1     93.247        2.6007        7.246       560.31          1.00
 102        23.      286.       263.          115.9         5793.75     22.0     27.933        3.4509       13.588       210.24          1.00
 141        26.      430.       404.          295.3        14766.02     36.5     54.919        3.2532       13.311       432.21          1.00
 158         9.      162.       153.           57.9         2893.36     18.9     31.393        3.0152        9.385       174.16          1.00
 159         9.      116.       107.           69.9         3495.09     32.7     46.351        2.2528        4.643       229.21          1.41
 193         7.        0.


A TOTAL NUMBER OF      528. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES   =       5394.

TOTAL INVISIBLE LINES =           0.

TOTAL LINES GENERATED =        5394.

TOTAL DISTANCE GENERATED BY LINES =        200814.63 METRES AT GROUND SCALE
                                  =           401.629 CMS AT MAP SCALE


NUMBER OF RECORDS IN THE FILE =            9477.

NUMBER OF POINTS IN THE FILE   =           5740.

NUMBER OF IGNORED CODES (0 OR >271)     =           0.

NUMBER OF IGNORED RECORDS (0 OR >271)   =           0.
```

```
                    FEATURES INCLUDED
                    =================

                                                                   -----GROUND-----                                      -----GROUND-----
                     NO.      NO.      TOTAL LINE     GROUND      MEAN   STANDARD
F/CODE    FREQ.      PTS     LINES    LENGTH (MM)    DIST (M)     DIST   DEVIATION     SKEW      KURTOSIS      MAXIMUM      MINIMUM
------    -----    ------   ------    -----------    --------    ------  ---------    ------    ---------     --------     -------
     7       7.     904.     897.          701.6     35081.56     39.1     72.629    5.0407        36.268      533.16        1.00
     8       4.       0.
     9       2.       0.
    15      10.      64.      54.           95.5      4776.63     88.5     86.767    1.3501         0.123      297.40        3.16
    19       9.      81.      72.          102.7      5135.93     71.3     73.187    1.5614         1.573      335.07        1.41
    47       2.      10.       8.          115.1      5753.65    719.2    763.209    1.0576         0.259     2346.00       61.14
    59      15.     633.     618.          440.8     22039.01     35.7     48.698    4.8064        35.565      491.25        1.00
    62      25.     921.     896.          191.7      9582.70     10.7     20.147    5.2952        36.834      226.11        0.94
    64      71.    1600.    1529.          898.1     44903.34     29.4     37.690    2.8941        11.134      336.03        1.00
    66       3.     262.     259.          201.0     10048.59     38.8     37.417    1.2674         1.041      195.14        1.00
    89      11.      48.      37.           56.4      2819.42     76.2     68.122    1.4120         0.719      258.16       20.00
    90      33.     202.     169.          219.9     10994.54     65.1     78.691    2.1007         4.533      444.97        1.00
    93      12.      61.      49.           98.9      4945.12    100.9     83.821    0.5962         1.125      270.17        4.12
    96      65.     482.     417.          547.6     27379.30     65.7     72.551    1.7936         2.841      397.18        1.00
    97       4.      53.      49.           59.0      2952.45     60.3     69.009    1.4613         1.027      274.29        4.24
    98     108.     731.     623.          914.3     45716.56     73.4     84.976    2.0414         4.453      481.04        1.00
   102       5.      84.      79.           56.7      2837.31     35.9     40.030    2.8442         9.549      245.36        1.00
   141       6.      83.      77.           99.3      4966.85     64.5    145.403    3.4966        12.966      849.30        1.41
   155       6.       6.
   158       4.      59.      55.           53.6      2679.35     48.7     32.205    0.8927         0.292      149.93        4.47
   159       5.      90.      85.           71.4      3568.13     42.0     38.401    1.1707         0.189      153.71        1.00

A TOTAL NUMBER OF      243. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES   =       5973.

TOTAL INVISIBLE LINES =           0.

TOTAL LINES GENERATED =        5973.

TOTAL DISTANCE GENERATED BY LINES =        246031.75 METRES AT GROUND SCALE
                                  =          492.063 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =            9492.

NUMBER OF POINTS IN THE FILE  =            6374.

NUMBER OF IGNORED CODES (0 OR >271)    =           0.

NUMBER OF IGNORED RECORDS (0 OR >271)  =           0.

FEATURE CODE DESCRIPTION
------------------------
```

```
                 FEATURES INCLUDED
                 =================

                                                                           GROUND                                       GROUND
                    NO.      NO.      TOTAL LINE       GROUND      MEAN  STANDARD
F/CODE   FREQ.      PTS     LINES    LENGTH (MM)      DIST (M)     DIST  DEVIATION     SKEW     KURTOSIS     MAXIMUM      MINIMUM

     7     16.     1246.    1230.        1132.2      56611.68      46.0     77.644    3.6763      18.691      798.96         1.00
     8      7.       15.      14.          55.7       2782.63     198.8     88.613    0.2073       1.211      340.01        45.28
     9      5.        0.
    15     35.      215.     180.         356.2      17811.35      99.0    108.348    2.3745       6.005      609.66         3.16
    19     23.      233.     215.         237.6      11880.58      55.3     54.324    1.9847       3.398      272.66         2.24
    47     14.       82.      68.         982.8      49139.75     722.6    465.239    0.6451      -0.336     1923.79        84.46
    59      6.     1194.    1188.         491.2      24560.58      20.7     23.389    2.8658       9.659      182.43         1.00
    62     25.      636.     611.         229.0      11451.99      18.7     21.520    2.9641      12.104      192.29         1.00
    64     62.     1674.    1612.        1422.1      71107.06      44.1     65.670    4.7854      41.639     1005.11         1.00
    66      1.      545.     544.         245.5      12275.01      22.6     25.740    2.5861       7.253      162.52         1.00
    71      4.      103.      99.          85.4       4271.25      43.1     45.402    1.7411       2.351      188.62         1.41
    89      4.       16.      12.          32.8       1639.22     136.6    184.815    1.4188       0.841      610.61        14.42
    90     84.      476.     392.         580.2      29010.74      74.0     77.077    1.6673       2.957      464.71         1.41
    92      6.       30.      24.          44.2       2209.19      92.0     82.809    1.1325       0.243      272.42        10.20
    93     68.      418.     350.         547.9      27396.15      78.3     88.023    2.1470       5.840      538.70         1.00
    96    116.      642.     526.         936.9      46845.00      89.1     97.397    2.1842       7.336      681.61         1.00
    97     13.       88.      75.          68.5       3424.26      45.7     97.765    4.8780      27.540      709.89         1.46
    98    311.     2160.    1849.        2344.0     117200.56      63.4     87.008    3.1470      17.282     1079.18         1.00
   102     70.      593.     523.         573.3      28667.23      54.8     74.128    3.0363      12.008      638.49         1.00
   141     37.      318.     281.         472.2      23612.34      84.0    123.365    2.9610       9.859      750.95         1.00
   155      3.        3.
   158     10.      183.     173.         103.2       5161.71      29.8     35.244    2.2961       5.457      137.42         1.41
   159      9.      157.     148.          90.5       4523.88      30.6     37.615    2.4904       7.412      240.84         1.00
   193      3.        0.


A TOTAL NUMBER OF       774. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES  =       10114.

TOTAL INVISIBLE LINES =           0.

TOTAL LINES GENERATED =       10114.

TOTAL DISTANCE GENERATED BY LINES =        551310.94 METRES AT GROUND SCALE
                                  =          1102.622 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =          18048.

NUMBER OF POINTS IN THE FILE  =          11032.

NUMBER OF IGNORED CODES (0 OR >271)   =          0.

NUMBER OF IGNORED RECORDS (0 OR >271) =          0.
```

```
                        FEATURES INCLUDED
                        =================

                                                                           GROUND                                              GROUND
                   NO.        NO.       TOTAL LINE       GROUND        MEAN    STANDARD
F/CODE    FREQ.    PTS        LINES     LENGTH (MM)      DIST (M)      DIST    DEVIATION       SKEW       KURTOSIS      MAXIMUM       MINIMUM

     7      16.    1246.      1230.        1132.2        56611.68      46.0      77.644       3.6763       18.691       798.96         1.00
     8       7.      15.        14.          55.7         2782.63     198.8      88.613       0.2073        1.211       340.01        45.28
     9       5.       0.
    15      35.     215.       180.         356.2        17811.35      99.0     108.348       2.3745        6.005       609.66         3.16
    19      23.     238.       215.         237.6        11880.58      55.3      54.324       1.9847        3.398       272.66         2.24
    47      14.      82.        68.         982.8        49139.75     722.6     465.239       0.6451       -0.336      1923.79        84.46
    59       6.    1194.      1188.         491.2        24560.58      20.7      23.389       2.8658        9.659       182.43         1.00
    62      25.     636.       611.         229.0        11451.99      18.7      21.520       2.9641       12.104       192.29         1.00
    64      62.    1674.      1612.        1422.1        71107.06      44.1      65.670       4.7854       41.639      1005.11         1.00
    66       1.     545.       544.         245.5        12275.01      22.6      25.740       2.5861        7.253       162.52         1.00
    71       4.     103.        99.          85.4         4271.25      43.1      45.402       1.7411        2.351       188.62         1.41
    89       4.      16.        12.          32.8         1639.22     136.6     184.815       1.4188        0.841       610.61        14.42
    90      84.     476.       392.         580.2        29010.74      74.0      77.077       1.6673        2.957       464.71         1.41
    92       6.      30.        24.          44.2         2209.19      92.0      82.809       1.1325        0.243       272.42        10.20
    93      68.     418.       350.         547.9        27396.15      78.3      88.023       2.1470        5.840       538.70         1.00
    96     116.     642.       526.         936.9        46845.00      89.1      97.397       2.1842        7.336       681.61         1.00
    97      13.      88.        75.          68.5         3424.26      45.7      97.765       4.8780       27.540       709.89         1.46
    98     311.    2160.      1849.        2344.0       117200.56      63.4      87.008       3.1470       17.282      1079.18         1.00
   102      70.     593.       523.         573.3        28667.23      54.8      74.128       3.0363       12.008       638.49         1.00
   141      37.     318.       281.         472.2        23612.34      84.0     123.365       2.9610        9.859       750.95         1.00
   155       3.       3.
   158      10.     183.       173.         103.2         5161.71      29.8      35.244       2.2961        5.457       187.42         1.41
   159       9.     157.       148.          90.5         4523.88      30.6      37.615       2.4904        7.412       240.84         1.00
   193       3.       0.


A TOTAL NUMBER OF       774. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES  =       10114.

TOTAL INVISIBLE LINES =           0.

TOTAL LINES GENERATED =       10114.

TOTAL DISTANCE GENERATED BY LINES =        551310.94 METRES AT GROUND SCALE
                                  =         1102.622 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =          18048.

NUMBER OF POINTS IN THE FILE  =          11032.

NUMBER OF IGNORED CODES (0 OR >271)    =          0.

NUMBER OF IGNORED RECORDS (0 OR >271)  =          0.
```

5

```
                FEATURES INCLUDED
                =================

                                                                      ----GROUND----                                         ------GROUND--
                     NO.       NO.      TOTAL LINE      GROUND      MEAN    STANDARD
F/CODE    FREQ.      PTS      LINES    LENGTH (MM)     DIST (M)     DIST   DEVIATION     SKEW     KURTOSIS     MAXIMUM      MINIMUM
------    -----   --------   -------  -------------  -----------   ------  ----------   -------  ----------   ----------   ---------
    7       18.     1846.      1828.       1064.1     53204.73     29.1      49.692     3.2681     12.987       416.77        1.00
    8        9.       32.        30.         78.0      3900.64    130.0     226.926     2.3134      4.679       937.25        5.00
    9        5.        0.
   15       14.      115.       101.        143.2      7159.17     70.9      77.145     2.3564      5.492       398.18        7.08
   19       47.      522.       475.        612.7     30635.23     64.5      81.466     3.2486     14.310       698.37        1.00
   47        6.       27.        21.        219.1     10954.48    521.6     358.071     0.8378      0.017      1440.81       36.69
   59        3.      111.       108.         56.6      2827.96     26.2      23.239     1.8608      3.855       132.85        1.00
   62       50.     3250.      3200.        549.4     27470.95      8.6      16.094     6.2851     60.882       291.31        0.94
   64      123.     2247.      2124.       1701.3     85065.00     40.0      51.153     2.3660      6.345       367.92        1.00
   85       10.      102.        92.        112.4      5617.52     61.1      66.161     2.5969      7.027       381.93        2.24
   86        4.       20.        16.         30.8      1538.29     96.1      77.203     0.4441     -1.510       237.78        5.10
   90       35.      178.       143.        281.6     14080.51     98.5     129.731     2.6994     11.087       924.73        1.00
   93       59.      310.       251.        461.4     23072.38     91.9      97.890     1.6136      2.609       527.09        1.00
   96       36.      273.       242.        378.1     18905.88     78.1     100.742     1.7032      2.489       497.09        1.00
   98      195.     1453.      1258.       1551.9     77594.94     61.7      80.434     2.4277      7.235       590.76        1.00
  102       14.      103.        89.        101.3      5063.11     56.9      73.112     2.3874      5.499       351.65        4.00
  141       25.      194.       169.        318.2     15910.40     94.1      72.885     1.2565      2.561       422.69        3.16
  158        1.        8.         7.         14.1       706.30    100.9      62.849     0.4466     -1.530       205.01       33.42
  159        1.       10.         9.         12.0       601.14     66.8      85.578     0.6097     -1.767       195.58        7.21


A TOTAL NUMBER OF       774. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES   =      10163.

TOTAL INVISIBLE LINES =          0.

TOTAL LINES GENERATED =       10163.

TOTAL DISTANCE GENERATED BY LINES =        384037.88 METRES AT GROUND SCALE
                                  =          768.075 CMS AT MAP SCALE

.NUMBER OF RECORDS IN THE FILE =           15766.

NUMBER OF POINTS IN THE FILE  =           10806.

NUMBER OF IGNORED CODES (0 OR >271)    =            0.

NUMBER OF IGNORED RECORDS (0 OR >271)  =            0.

FEATURE CODE DESCRIPTION
------------------------

 FEATURE    FEATURE
```

```
                 FEATURES INCLUDED
                 =================

                                                                      GROUND                                          GROUND
                          NO.      NO.     TOTAL LINE        GROUND     MEAN  STANDARD
F/CODE    FREQ.           PTS     LINES    LENGTH (MM)       DIST (M)   DIST  DEVIATION       SKEW      KURTOSIS      MAXIMUM      MINIMUM

   7        7.          1228.     1221.         623.3       31165.44    25.5    46.721      3.4310       13.597       386.09         1.00
   8        2.             0.
   9        2.             0.
  19       20.           292.      272.         234.9       11747.14    43.2    69.474      5.1844       32.988       656.88         1.41
  47        1.             4.        3.          45.7        2287.11   762.4     0.0        0.0          0.0         890.89       636.78
  62       28.          1108.     1080.         184.8        9239.29     8.6    12.849      4.4096       25.932       132.82         1.00
  64      135.          8558.     8423.        1873.0       93652.44    11.1    21.079      7.3962       90.495       486.95         0.94
  90       21.           419.      398.         242.6       12131.78    30.5    55.514      3.9133       17.939       428.07         1.00
  93       42.           333.      291.         286.5       14323.80    49.2    71.939      2.5476        6.830       428.69         1.00
  96       42.           498.      456.         280.6       14029.29    30.8    59.896      4.3861       24.664       575.95         1.00
  97        5.            84.       79.          51.2        2559.76    32.4    48.456      2.0865        3.364       201.90         3.61
  98      158.          2475.     2317.        1152.9       57646.27    24.9    48.147      4.2021       22.173       499.19         1.00
 102       15.           208.      193.          78.3        3912.64    20.3    26.108      2.9559        9.405       152.70         0.94
 141       18.           406.      388.         303.0       15152.27    39.1    57.188      2.7648        8.832       403.11         1.00


A TOTAL NUMBER OF      225. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES   =      15121.

TOTAL INVISIBLE LINES  =          0.

TOTAL LINES GENERATED  =      15121.

TOTAL DISTANCE GENERATED BY LINES =        267557.56 METRES AT GROUND SCALE
                                  =           535.115 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =          19235.

NUMBER OF POINTS IN THE FILE  =          15613.

NUMBER OF IGNORED CODES (0 OR >271)     =           0.

NUMBER OF IGNORED RECORDS (0 OR >271)   =           0.

FEATURE CODE DESCRIPTION
------------------------

 FEATURE   FEATURE
  CODE      TYPE                     DESCRIPTION
 -------   -------                   -----------

   7        LINE              BOUNDARY - PARISH OR COMMUNITY
```

```
           FEATURES INCLUDED
           =================

                                                                       GROUND                                        GROUND
                   NO.      NO.     TOTAL LINE      GROUND     MEAN   STANDARD
F/CODE   FREQ.     PTS     LINES    LENGTH (MM)    DIST (M)    DIST   DEVIATION     SKEW     KURTOSIS     MAXIMUM     MINIMUM

    7      5.     1544.    1539.        381.6      19082.02    12.4     32.021     6.1536     43.876      329.96        1.00
    8      3.        0.
    9      3.        0.
   15      6.       17.      11.         43.2       2158.22   196.2    156.896     0.9563      0.367      545.05       47.18
   47      1.        6.       5.         33.5       1672.53   334.5    202.381     0.7002      1.410      656.62      175.54
   62      3.      196.     193.         22.6       1129.17     5.9      7.664     4.1111     18.450       50.25        1.00
   64     18.     2973.    2955.        446.5      22323.96     7.6     12.086     6.3185     60.485      214.84        0.94
   90     21.      162.     141.        152.7       7633.24    54.1     75.188     2.7308      8.730      461.47        3.61
   93      6.       63.      57.         41.0       2048.96    35.9     48.347     1.9432      2.728      198.98        5.39
   96      6.       59.      53.         61.0       3048.90    57.5    102.009     3.0715     11.143      577.97        2.29
   97      3.       76.      73.         21.1       1054.81    14.4     23.545     4.0554     18.975      155.90        1.00
   98     64.      799.     735.        491.3      24565.48    33.4     59.333     3.1675     11.370      403.44        1.00
  141      2.       39.      37.         25.5       1274.99    34.5     88.319     4.5707     21.949      520.79        2.00
  155      1.        1.


A TOTAL NUMBER OF      210. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES   =       5799.

TOTAL INVISIBLE LINES =           0.

TOTAL LINES GENERATED =        5799.

TOTAL DISTANCE GENERATED BY LINES =         85941.13 METRES AT GROUND SCALE
                                  =           171.882 CMS AT MAP SCALE


NUMBER OF RECORDS IN THE FILE =            7005.

NUMBER OF POINTS IN THE FILE  =            5935.

NUMBER OF IGNORED CODES (0 OR >271)     =          0.

NUMBER OF IGNORED RECORDS (0 OR >271)   =          0.

FEATURE CODE DESCRIPTION
------------------------

 FEATURE   FEATURE
  CODE      TYPE                      DESCRIPTION
-------   -------                    -----------

   7        LINE                 BOUNDARY - PARISH OR COMMUNITY
```

8

```
                          FEATURES INCLUDED
                          =================

                                                                          ------GROUND------                                                  ------GROUND------
                      NO.      NO.       TOTAL LINE      GROUND       MEAN   STANDARD
F/CODE     FREQ.      PTS      LINES     LENGTH (MM)     DIST (M)     DIST   DEVIATION       SKEW       KURTOSIS       MAXIMUM       MINIMUM
------     -----    ------    ------     -----------    ----------   ------  ---------     -------     ----------     ---------     ---------
   7        12.     1073.     1061.          514.3       25717.01     24.2     37.641      3.1443        11.921         295.15          1.00
   8         8.        0.
   9         8.        0.
  15        17.      221.      204.          275.9       13795.35     67.6     80.614      2.3737         5.920         465.54          3.16
  19         1.        3.        2.           17.3         867.46    433.7      0.0        0.0            0.0           547.81        319.65
  47        19.       69.       50.          644.6       32229.50    644.6    512.103      1.2286         0.983        2255.56         17.74
  62        19.     1239.     1220.          358.8       17940.09     14.7     27.616      6.4477        56.792         359.34          0.94
  64        29.     1194.     1165.          519.4       25972.03     22.3     28.760      3.7957        24.521         358.08          1.00
  71         3.      203.      200.          218.7       10935.68     54.7     62.506      1.6486         1.895         283.21          2.00
  90        25.      103.       78.          202.1       10104.27    129.5    138.234      1.1084         0.288         516.98          1.94
  93        24.      232.      208.          198.3        9915.00     47.7     54.913      2.5023         8.694         372.76          1.00
  96        11.       69.       58.          102.6        5129.92     88.4    103.157      1.6370         2.489         457.23          5.00
  97         9.       79.       70.           90.9        4547.13     65.0     60.763      1.2421         1.045         253.21          1.00
  98       152.     1365.     1213.         1139.0       59447.91     49.0     65.997      3.2751        18.044         753.11          1.00
 102        14.      191.      177.          127.5        6373.81     36.0     47.589      2.9856         9.382         275.63          1.00
 141        36.      376.      340.          304.6       15228.25     44.8     62.582      2.6235         7.415         392.37          1.00
 155         1.        1.
 193         7.        0.


A TOTAL NUMBER OF      537. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES    =        6046.

TOTAL INVISIBLE LINES =            0.

TOTAL LINES GENERATED =          6046.

TOTAL DISTANCE GENERATED BY LINES =        238080.25 METRES AT GROUND SCALE
                                  =          476.160 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =            9445.

NUMBER OF POINTS IN THE FILE  =            6418.

NUMBER OF IGNORED CODES (0 OR >271)   =            0.

NUMBER OF IGNORED PECORDS (0 OR >271)  =           0.

FEATURE CODE DESCRIPTION
------------------------

 FEATURE   FEATURE
  CODE       TYPE                         DESCRIPTION
```

```
                    FEATURES INCLUDED
                    =================

                                                                  ----GROUND----                                     ----GROUND----
                       NO.        NO.      TOTAL LINE      GROUND      MEAN    STANDARD
F/CODE    FREQ.        PTS       LINES     LENGTH (MM)    DIST (M)     DIST    DEVIATION      SKEW      KURTOSIS       MAXIMUM       MINIMUM
------    -----     -------    -------    -----------    ---------    -----    ---------     ------
     7      13.       2400.      2387.          861.4     43070.77     18.0       41.997    11.4402     204.849       989.08          0.94
     8       5.          0.
     9       5.          0.
    15      44.        601.       557.          551.5     27575.14     49.5       68.073     2.9302       8.887       399.41          1.00
    19       7.        111.       104.           64.1      3202.79     30.8       28.390     3.1519      12.022       171.41          4.47
    47       2.         42.        40.          414.6     20727.54    518.2      210.967    -0.3096       0.536       927.98        115.14
    59       3.        184.       181.           42.6      2127.80     11.8       14.969     3.7484      16.349       104.14          1.41
    62      15.        579.       564.          192.4      9120.00     16.2       41.072     5.9147      40.638       408.95          1.00
    64      57.       4883.      4826.         1124.8     56237.60     11.7       14.749     4.5529      28.856       175.98          0.94
    66       1.         80.        79.           21.0      1048.55     13.3       15.388     3.2005      10.282        84.39          1.41
    71       5.        846.       841.          343.8     17189.80     20.4       28.461     3.4351      14.306       227.18          1.00
    89       2.         15.        13.           10.5       525.49     40.4       30.296     1.1442       0.968       120.02          9.00
    90     117.        966.       849.          716.4     35821.96     42.2       66.014     3.0036      10.179       139.13          1.00
    93      26.        336.       310.          153.1      7652.70     24.7       38.161     3.8780      20.566       348.57          1.41
    96      67.        887.       820.          588.0     29400.29     35.9       62.710     3.3483      12.990       466.49          1.00
    97      16.        308.       292.          150.3      7514.07     25.7       50.511     3.6753      14.271       345.76          1.00
    98     222.       2650.      2428.         1564.5     78222.50     32.2       57.506     3.6646      17.778       663.07          1.00
   102      16.        165.       149.           98.3      4914.93     33.0       64.540     4.0320      18.528       431.70          1.41
   141      21.        555.       534.          272.9     13643.21     25.5       40.587     4.5533      27.508       419.89          1.00
   155       2.          2.
   158       9.        423.       414.          201.5     10072.52     24.3       48.415     4.3800      22.363       398.22          1.00
   159       8.        181.       173.           89.5      4475.73     25.9       47.641     5.6923      38.139       402.25          1.00
   193       9.        120.       115.           47.7      2386.36     20.8       37.671     3.5325      12.391       221.94          1.00


A TOTAL NUMBER OF       492. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METRES

TOTAL INKED IN LINES  =       15676.

TOTAL INVISIBLE LINES =            0.

TOTAL LINES GENERATED =       15676.

TOTAL DISTANCE GENERATED BY LINES =        374625.50 METRES AT GROUND SCALE
                                  =          749.251 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =           21352.

NUMBER OF POINTS IN THE FILE  =           16334.

NUMBER OF IGNORED CODES (0 OR >271)   =          0.

NUMBER OF IGNORED RECORDS (0 OR >271)  =          0.
```

```
                 FEATURES INCLUDED
                 =================

                                                              ----GROUND----                                    -----GROUND-  -
                    NO.        NO.     TOTAL LINE     GROUND      MEAN   STANDARD
F/CODE   FREQ.      PTS  .   LINES    LENGTH (MM)    DIST (M)     DIST   DEVIATION      SKEW        KURTOSIS      MAXIMUM       MINIMUM
------   -----   --------   -------   -----------   ----------   ------  ---------    ---

     7     28.     3843.     3815.        1201.5     60074.90     15.7     36.795    18.8036       659.337      1472.57         0.94
     8     21.        5.        4.          12.5       626.88    156.7     93.605     0.3343        -1.843       293.69        58.31
     9     19.        0.
    15     54.      635.      581.         600.1     30003.18     51.6     67.311     3.5741        16.234       529.76         1.00
    19      3.       28.       25.          12.6       627.60     25.1     14.441     1.4926         3.653        76.11         6.74
    47      4.       39.       35.         291.2     14560.54    416.0    192.647    -0.0933        -0.630       790.32        12.29
    62     27.     1269.     1242.         184.4      9220.55      7.4     12.239     6.2605        55.013       179.04         1.00
    64     83.     4448.     4365.        1414.4     70719.69     16.2     27.451     5.7628        46.196       392.40         0.94
    71     10.     1879.     1869.         460.4     23019.92     12.3     24.606    10.8412       179.139       524.39         1.00
    85     17.      160.      143.         175.7      8783.53     61.4     68.547     3.2440        14.505       513.52         1.00
    86     14.      177.      163.         102.6      5129.45     31.5     33.914     2.7560        11.403       251.77         1.00
    89      8.       95.       87.          91.1      4557.11     52.4     58.521     2.7316         8.485       352.22         2.24
    90     88.      866.      778.         611.3     30556.21     39.3     64.571     3.2645        12.188       522.78         1.00
    92      3.       17.       14.          22.4      1119.88     80.0    104.008     0.7912        -1.336       263.70         4.12
    93     59.      667.      608.         446.4     22320.14     36.7     62.242     3.2044        11.469       429.07         1.00
    95      1.        2.        1.           5.8       291.61    291.6      0.0       0.0           0.0        291.61       291.61
    96     62.      775.      713.         569.7     28484.46     40.0     67.381     3.7066        17.710       609.93         1.00
    97     25.      401.      376.         203.0     10152.49     27.0     46.804     3.4250        14.942       372.21         1.00
    98    233.     3341.     3108.        1720.2     86009.44     27.7     52.295     4.4080        24.947       575.31         1.00
   102     26.      289.      263.         152.2      7608.48     28.9     48.151     3.4000        13.879       372.14         1.00
   141     29.      572.      543.         297.2     14858.30     27.4     44.857     3.9145        20.175       422.86         1.00
   155      7.        7.
   193     22.       47.       43.          65.7      3285.63     76.4    123.565     2.4684         6.179       586.96         1.00


A TOTAL NUMBER OF     1569. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =        0. GND. METPES

TOTAL INKED IN LINES  =      18776.

TOTAL INVISIBLE LINES =          0.

TOTAL LINES GENERATED =      18776.

TOTAL DISTANCE GENERATED BY LINES =        431523.25 METRES AT GROUND SCALE
                                  =          863.046 CMS AT MAP SCALE


NUMBER OF RECORDS IN THE FILE =          25872.

NUMBER OF POINTS IN THE FILE  =          19562.

NUMBER OF IGNORED CODES (0 OR >271)    =          0.

NUMBER OF IGNORED RECORDS (0 OR >271)  =          0.
```

```
                    FEATURES INCLUDED
                    =================

                                                                      -----GROUND-----                                       ------GROUND------
           NO.      NO.     TOTAL LINE       GROUND      MEAN   STANDARD
F/CODE  FREQ.  PTS      LINES   LENGTH (MM)      DIST (M)     DIST   DEVIATION     SKEW     KURTOSIS    MAXIMUM    MINIMUM
------  -----  -------  ------  -----------      --------     -----  ---------     ----     --------    -------    -------
    7    18.   2174.    2156.         752.8      37637.98      17.5     32.983    5.7976      49.528     479.88       0.94
    8    15.      0.
    9    12.      0.
   15    10.    182.     172.         221.8      11090.09      64.5     66.147    2.4628       6.140     352.63       4.12
   47     1.      6.       5.          46.3       2314.73     462.9    264.620    0.1544      -1.790     833.43     151.14
   62    17.    675.     658.          98.2       4909.73       7.5     11.575    4.7044      26.133     104.04       1.00
   64    27.    865.     838.         378.7      18932.74      22.6     42.822    3.9786      19.431     373.49       1.00
   71    12.   1618.    1606.         334.4      16718.49      10.4     18.149    5.6301      44.642     264.17       0.94
   90    56.    712.     656.         476.5      23827.25      36.3     51.232    3.3365      14.216     438.87       1.00
   93    14.    143.     129.          88.1       4403.60      34.1     42.373    2.6821       7.398     238.11       2.83
   97    33.    528.     495.         310.4      15517.89      31.3     47.747    3.4065      13.566     357.30       1.00
   98   129.   1873.    1744.         847.5      42376.38      24.3     40.480    4.1044      21.917     428.24       1.00
  102     3.    123.     120.          73.4       3668.97      30.6     40.971    3.4403      13.816     269.22       1.00
  141    27.    614.     587.         318.7      15933.55      27.1     43.080    3.8596      18.681     334.03       1.00
  155     1.      1.
  193    16.     91.      87.           6.0        298.79       3.4      2.605    4.0952      23.226      20.97       0.94


A TOTAL NUMBER OF     1031. CHARACTERS EXIST

TOTAL DISTANCE OF -15 CODED LINES =         0. GND. METRES

TOTAL INKED IN LINES  =       9253.

TOTAL INVISIBLE LINES =          0.

TOTAL LINES GENERATED =       9253.

TOTAL DISTANCE GENERATED BY LINES =        197426.63 METRES AT GROUND SCALE
                                  =          394.853 CMS AT MAP SCALE

NUMBER OF RECORDS IN THE FILE =           12524.

NUMBER OF POINTS IN THE FILE  =            9605.

NUMBER OF IGNORED CODES (0 OR >271)    =           0.

NUMBER OF IGNORED RECORDS (0 OR >271)  =           0.

FEATURE CODE DESCRIPTION
------------------------

 FEATURE   FEATURE
  CODE      TYPE                          DESCRIPTION
 -------   -------                        -----------
```

Appendix C Example plots made using filtered data

The plots in this appendix were made using the University of Durham CIL incremental plotter. While the quality of this plotter is insufficient to meet the requirements of a commercial or government mapping agency, it is representative of a number of plotters that are likely to be owned by users of topographical data in digital form.

Appendix D Source code listing of Fortran Program to filter vector encoded map data.

```
C
C
C ******************************************************************
C
C
C     FILTER PROGRAM TO INPUT DATA IN THE FORMAT OUTPUT BY PROGRAM
C     DMCD AND FILTER IT BY A METHOD BASED ON THAT DEVISED BY ORDNANCE
C     SURVEY.  PLOT ROUTINES FROM *GHOST ARE ALSO UTILISED TO PLOT
C     UNFILTERED AND FILTERED DATA TOGETHER ON ONE PLOT FOR COMPARISON
C     PURPOSES.
C
C
C     THE MAIN CALLING PROGRAM COMES LAST IN THE SOURCE LISTING
C
C
C               N S SMITH  JULY 1980
C
C
C*******************************************************************
C
C
C
        SUBROUTINE FILTA
C
C     SUBROUTINE TO FILTER ORDNANCE SURVEY DMC FORMAT DATA BY A OFFSET
C     DISTANCE AND DEVIATION ANGLE METHOD BASED ON THAT
C     DEVISED BY OS AND M RAND
C
        COMMON/Z/IEAST(500),NORTH(500),NCO,ISN,IFC,NFCO,T,R,ICONT
         DIMENSION IX(200),IY(200)
C
        INTEGER S,ANCH,FLOT,STK(200),DE,DN,X1,Y1,DX,DY,CURPT
C
C
C
C     FUNCTION TO CALCULATE HYPOTENUSE
C
          HYPOT(I,J)=SQRT(FLOAT(I*I+J*J))
C
C     ELIMINATE CONSECUTIVE IDENTICAL POINTS
C
C     AND READ CO-ORDS OF FEATURE INTO TEMPORARY ARRAYS
C
C
C
        PI=3.1415927
        IF(NCO.GT.2)GO TO 3
        NFCO=NCO
        RETURN
 3      CONTINUE
         N=NCO-1
        J=1
        DO10 I=1,N
        IF(IEAST(I)-IEAST(I+1))5,4,5
 4      IF(NORTH(I)-NORTH(I+1)) 5,10,5
 5      IX(J)=IEAST(I)
        IY(J)=NORTH(I)
        J=J+1
 10     CONTINUE
        IX(J)=IEAST(NCO)
```

```
      IY(J)=NORTH(NCO)
      NCO=J   2
      DO 12 I=1,NCO
      IEAST(I)=0
      NORTH(I)=0
   12 CONTINUE
C
C  INITIALISE PRIOR TO RECURSIVE ROUTINE
C  ANCH - ANCHOR POINT LOCATION
C  FLOT - FLOATER POINT LOCATION
C  S - POSITION OF 'STACK' POINTER
C  NFCO - NUMBER OF RETAINED CO-ORDS
C  AMAXOF - MAXIMUM OFFSET DISTANCE (OUTSIDE TOLERANCE)
C
C
C
      S=1
      ANCH=1
      FLOT=NCO
      NFCO=1
C
C  START ROUTINE
C
20    STK(S)=FLOT
30    S=S+1
      AMAXOF=0.0
C
C  TEST FOR ADJACENT ANCHOR AND FLOATER POINTS
C
40    IF((FLOT-ANCH).LE.1) GO TO 400
C
C  CREATE CO-EFFICIENTS TO DESCRIBE PERPENDICULAR TO LINE BETWEEN ANCHOR
C  AND FLOATER POINTS.
C
C  THESE ARE THEN APPLIED TO EACH POINT IN TURN ALONG THE FEATURE
C  TO, DETERMINE THE DISTANCE FROM THE 'BASELINE'.
C
C
      DE=IX(ANCH)-IX(FLOT)
      DN=IY(ANCH)-IY(FLOT)
      IF(IABS(DE)-IABS(DN))50,55,55
50    A1=1.0
      IF(DN.NE.0)GO TO51
      A2=0
      GO TO 60
51    A2=-1.0*DE/DN
      GO TO 60
55    A2=1.0
      IF(DE.NE.0) GO TO 56
      A1=0
      GO TO 60
56    A1=-1.0*DN/DE
60    A3=IX(ANCH)*A1+IY(ANCH)*A2
       DENOM=SQRT(A1*A1+A2*A2)
C
C  CALCULATE INITIAL DISTANCE FOR DEVIATION TEST
C
      DE=IX(FLOT-1)-IX(FLOT)
      DN=IY(FLOT-1)-IY(FLOT)
       FDIST=HYPOT(DE,DN)
```

```
      IXBK=IX(FLOT)
      IYBK=IY(FLOT)
C
C  TEST EACH POINT BACKWARDS FROM FLOAT TO ANCHOR FOR OFFSET DISTANCE AN
C  DEVIATION ANGLE. RESTART ROUTINE IF GREATER THAN TOLERANCE
C
C
C  CALCULATE OFFSET DISTANCE
C
      K3=FLOT-ANCH-1
      DO 100  I=1,K3
      CURPT=FLOT-I
      X1=IX(CURPT)
      Y1=IY(CURPT)
      Q=(X1*A1+Y1*A2-A3)/DENOM
      Q=ABS(Q)
      IF(Q.LE.T) GO TO 80
      IF(Q.LE.AMAXOF)GO TO 80
      AMAXOF=Q
      K=CURPT
C
C
C  CALCULATE DEVIATION ANGLE USING COSINE FORMULA
C  ON THE TRIANGLE FORMED BY THE CURRENT, LAST, AND NEXT POINTS.
C
80    BDIST=FDIST
      X2=IX(CURPT-1)
      Y2=IY(CURPT-1)
      DE=X2-X1
      DN=Y2-Y1
      FDIST=HYPOT(DE,DN)
      DE=X2-IXBK
      DN=Y2-IYBK
      TRIG=HYPOT(DE,DN)
      IF((BDIST+FDIST).GT.TRIG) GO TO 89
      DEV=PI
      GOTO87
  88   FORMAT(3F20.4)
89    DENM=FDIST*BDIST*2
      TERM1=BDIST*BDIST/DENM
      TERM2=FDIST*FDIST/DENM
      TERM3=TRIG*TRIG/DENM
      DEV=ARCOS(TERM1+TERM2-TERM3)
      IXBK=X1
      IYBK=Y1
87      DEV=PI-DEV
        IF(DEV-R)100,100,110
C
  100   CONTINUE
        GO TO 400
101    FORMAT(1X,F20.3,F8.2)
C
C  IF DEVIATION IS GREATER THAN TOLERANCE, ASSIGN POINT AS NEW FLOATER
C  AND START AGAIN
C
  110   FLOT=CURPT
        GOTO20
C
C
C  IF MAXIMUM OFFSET = 0 THEN OUTPUT ANCHOR,POP STACK AND CONTINUE FROM
```

```
C
400   IF(AMAXOF.EQ.0.0) GO TO 500
      FLOT = K
      GOTO 20
500   IEAST(NFCO)=IX(ANCH)
      NORTH(NFCO)=IY(ANCH)
      NFCO=NFCO+1
      S=S-1
       ANCH=STK(S)
       IF(S.LT.2)GOTO 600
      S=S-1
      FLOT=STK(S)
       S=S+1
      AMAXOF=0.0
      GO TO 40
C
C  OUTPUT FINAL ANCHOR INTO MAIN ARRAYS AND RETURN
C
 600  IEAST(NFCO)=IX(ANCH)
      NORTH(NFCO)=IY(ANCH)
      RETURN
      END
C
C
C
C **************************************************************************
C
C
      SUBROUTINE PLOTGH(IQ)
C
C
C
C  THIS SUBROUTINE USES *GHOST ROUTINES TO PRODUCE A PDF CONTAINING
C  THE ORIGINAL DATA IN PECKED LINES AND THE FILTERED LINES IN SOLID
C
C
C
C
      COMMON/Z/IEAST(500),NORTH(500),NCO,ISN,IFC,NFCO,T,R,ICONT
      DIMENSIONX1(500),Y1(500)
C
C
C  VARIABLE ICONT SWITCHES THE SUBROUTINE TO TERMINATE THE PLOT IF
C  SET TO -1 IN SUBROUTINE COREAD
C
C
      IF(ICONT)999,5,5
C
C
C  VAIABLE IQ IS USED AS A SWITCH TO CONTROL PLOTGH, CAUSING
C  THE SUBROUTINE TO PRODUCE EITHER PLOT INITIALISATION, PECKED OR
C  SOLID LINES
C
C
5     IF(IQ)100,10,200
10    CALL PAPER(1)
      CALL PSPACE(0.2,0.95,0.3,0.95)
      CALL CSPACE(0.1,0.95,0.3,0.95)
      CALL MAP(0,10000.0,0,10000.0)
      CALL CTRSET(2)
```

```
      CALL CTRSIZ(10)
      CALL BORDER
      CALL BLUPEN
      CALL GRATSI(1000.0,1000.0)
      RETURN
C  PLOT UNFILTERED CO-ORDINATES WITH A SOLID LINE
C
C
C  CONVERT INTEGER ARRAY INTO TEMPORARY REAL
C
100   DO 110 I=1,NCO
      X1(I)=FLOAT(IEAST(I))
      Y1(I)=FLOAT(NORTH(I))
110   CONTINUE
      CALL BROKEN(5,5,5,5)
      CALL BLKPEN
      CALL POSITN(X1(1),Y1(1))
      DO 117 I=1,NCO
      CALL JOIN(X1(I),Y1(I))
117   CONTINUE
      RETURN
C
C  PLOT FILTERED LINES IN RED SOLID LINES
C
200   DO 210 I=1,NFCO
      X1(I)=FLOAT(IEAST(I))
      Y1(I)=FLOAT(NORTH(I))
210   CONTINUE
      CALL REDPEN
      CALL FULL
      CALL POSITN(X1(1),Y1(1))
      DO 217 I=1,NFCO
      CALL JOIN(X1(I),Y1(I))
217   CONTINUE
      RETURN
C  TERMINATE PLOT
C
C
999   CALL GREND
      RETURN
      END
C
C
C
C ************************************************************************
C
C
C
      SUBROUTINE CORITE
C
C
C  SUBROUTINE TO WRITE OUT CO-ORDINATES IN THE FORMAT OF
C  'DMCRD' FOR FURTHER PROCESSING IF REQUIRED.
C
C  THIS FORMAT IS
C
C   EASTINGS, NORTHINGS,(OF SW CORNER) GRID INTERVAL, SCALE (4I12)
C   SERIAL NO., FEATURE CODE, NO OF CO-ORDS  (3I2)
C   SHEET EASTINGS, SHEET NORTHING  (2I8)
C  ......
```

```
C
C   -1,   EASTINGS, NORTHINGS  (3I12) - END OF SHEET
C
C
      COMMON/Z/IEAST(500),NORTH(500),NCO,ISN,IFC,NFCO,T,R,ICONT
      COMMON/Z1/IESW,INSW,IGRID,ISCALE
      IF(ICONT)999,10,100
10    WRITE(3,1050)IESW,INSW,IGRID,ISCALE
C
C ZERO COUNTERS FOR NUMBER OF FEATURES, INPUT AND OUTPUT POINTS
C
      IFCOUN=0
      IPCOUN=0
      IFILCO=0
      RETURN
C
C
C  WRITE OUT FEATURE PARAMETERS TO CHANNEL 3 AND 6, PLUS CO-ORDINATES
C  TO CHANNEL 3
C
C
100   WRITE(3,1000)ISN,IFC,NFCO
      WRITE(6,1300)ISN,IFC,NCO,NFCO
      DO150I=1,NFCO
      WRITE(3,1100)IEAST(I),NORTH(I)
150   CONTINUE
      IFCOUN=IFCOUN+1
      IPCOUN=IPCOUN+NCO
      IFILCO=IFILCO+NFCO
      RETURN
C
C
C  AT THE END OF A SHEET WRITE OUT TOTALS ETC.
C
C
999   WRITE(3,1000)ISN,IESW,INSW
      WRITE(6,1300)IFCOUN,IPCOUN,IFILCO
      R=R*180/3.1419527
      IR=IFIX(R)
      WRITE(6,1400)T,IR
      WRITE(6,1500)IESW,INSW
      WRITE(6,1600)
1000  FORMAT(3I12)
1050  FORMAT(4I12)
1100  FORMAT(2I8)
1200  FORMAT('  FEATURES - ',I5,' INPUT POINTS - ',I6,' OUTPUT POINTS -
     1',I6)
1300  FORMAT(/,'  SERIAL - ',I5,' F.CODE - ',I5,' POINTS - INPUT',I5,' O
     1UTPUT',I5)
1400  FORMAT(//,'DISTANCE TOLERANCE - ',F4.1,'DEVIATION ANGLE TOLERANCE
     1- ',I4)
1500  FORMAT(//,' SHEET SOUTH WEST CORNER',/' EASTINGS - ',I10,' NORTHIN
     1GS - ',I10)
1600  FORMAT(///,'*****************************************************
     1*****')
      RETURN
      END
C
C
C
```

```
C ************************************************************************
C
C
C
        SUBROUTINE COREAD
C
C
C
C
C   SUBROUTINE TO READ IN CO-ORDINATES IN FORMAT OUTPUT BY 'DMCRD'
C   AND PUT THEM INTO ARRAYS IEAST AND NORTH.  THIS IS DONE
C   FEATURE BY FEATURE FROM CHANNEL 2.
C
C
        COMMON/Z/IEAST(500),NORTH(500),NCO,ISN,IFC,NFCO,T,R,ICONT
        COMMON/Z1/IESW,INSW,IGRID,ISCALE
        IF(ICONT)999,10,20
C   READ START OF SHEET PARAMETERS
C
   10   READ(2,1000)IESW,INSW,IGRID,ISCALE
 1000   FORMAT(4I12)
        RETURN
 20     READ(2,1100)ISN,IFC,NCO
 1100   FORMAT(3I12)
        IF(ISN)990,990,21
 21     DO25I=1,NCO
        READ(2,1200)IEAST(I),NORTH(I)
 25     CONTINUE
 1200   FORMAT(2I8)
        RETURN
 990    ICONT=-1
        RETURN
 999    STOP
        END
C
C
C
C***********************************************************************
C
C
C
        SUBROUTINE FILPAR
C
C
C
C   READS IN FILTER TOLERANCES FROM CHANNEL 5 (*SOURCE*)
C   AND CONVERTS THEM TO REAL FORMAT
        COMMON/Z/IEAST(500),NORTH(500),NCO,ISN,IFC,NFCO,T,R,ICONT
        WRITE(5,1000)
        READ(6,100)IR
        R=IR*3.1419527/180
 100    FORMAT(I4)
 1000   FORMAT('  DEVIATION ANGLE TOLERANCE(DEGREES) IN I4 FORMAT')
        WRITE(5,1001)
        READ(6,100)IT
        T=FLOAT(IT)
        WRITE(6,1002)IR,IT
 1001   FORMAT('  MAXIMUM OFFSET DISTANCE(METRES) IN I4 FORMAT')
 1002   FORMAT('  TOLERANCES:',//'   -  ',I4,' DEGREES',/'   -  ',I5,' MET
       1RES')
```

```
      RETURN
      END
C
C
C
C ************************************************************************
C
C
C
C THIS PROGRAM INPUTS DATA IN THE FORMAT OUTPUT BY 'DNCRD'
C AND FILTERS IT USING SUBROUTINE 'FILTA'
C A PLOT SUBROUTINE 'PLTGH' MAY ALSO BE CALLED TO PLOT
C BOTH UNFILTERED AND FILTERED DATA ON ONE PLOT
C
      COMMON/Z/IEAST(500),NORTH(500),NCO,ISN,IFC,NFCO,T,R,ICONT
C
C IEAST AND NORTH HOLD NCO CO-ORDINATES OF A FEATURE BEFORE
C FILTERING AND NFCO AFTERWARDS. THE DISTANCE TOLERANCE IS
C T METRES, ANGLE TOLERANCE R RADIANS. ICONT IS A CONTROL
C PARAMETER FOR ROUTING THE START AND END OF SHEET ROUTINES
C
C
      ICONT=0
      IQ=0
C INITIALISE PLOT
      CALL PLOTGH(IQ)
C
C INPUT FILTER PARAMETERS
C
      CALL FILPAR
C
C INPUT CO-ORDINATE PARAMETERS
C
      CALL COREAD
C WRITE OUT SHEET PARAMETERS
      CALL CORITE
      ICONT=1
C
C START READING CO-ORDINATES AND PLOTTING THEM.....
C
5     CALL COREAD
      IF (ICONT)100,10,10
10    IQ=-1
      CALL PLOTGH(IQ)
C
C
C AND FILTERING THEM......
C
      CALL FILTA
      IQ=1
C
C BEFORE PLOTTING THE FILTERED VERSION
C
      CALL PLOTGH(IQ)
      CALL CORITE
      GO TO 5
C
C IF AN END OF SHEET IS FLAGGED, TERMINATE PLOT
C
100   CALL PLOTGH(IQ)
```

```
              CALL CORITE
              WRITE(6,1000)
        1000  FORMAT('  END OF FILTER RUN')
              STOP
              END
) OF FILE
```

ROADS AND RAILWAYS

PLOT NO. 0

1:50,000

NO OF DATA POINTS

UNFILTERED 6773
FILTERED 6773

SOUTH WEST CORNER

EASTINGS 310000
NORTHINGS 680000

1:250,000

ROADS AND RAILWAYS

PLOT NO. 1

1:50,000

NO OF DATA POINTS

UNFILTERED	6773
FILTERED	4255

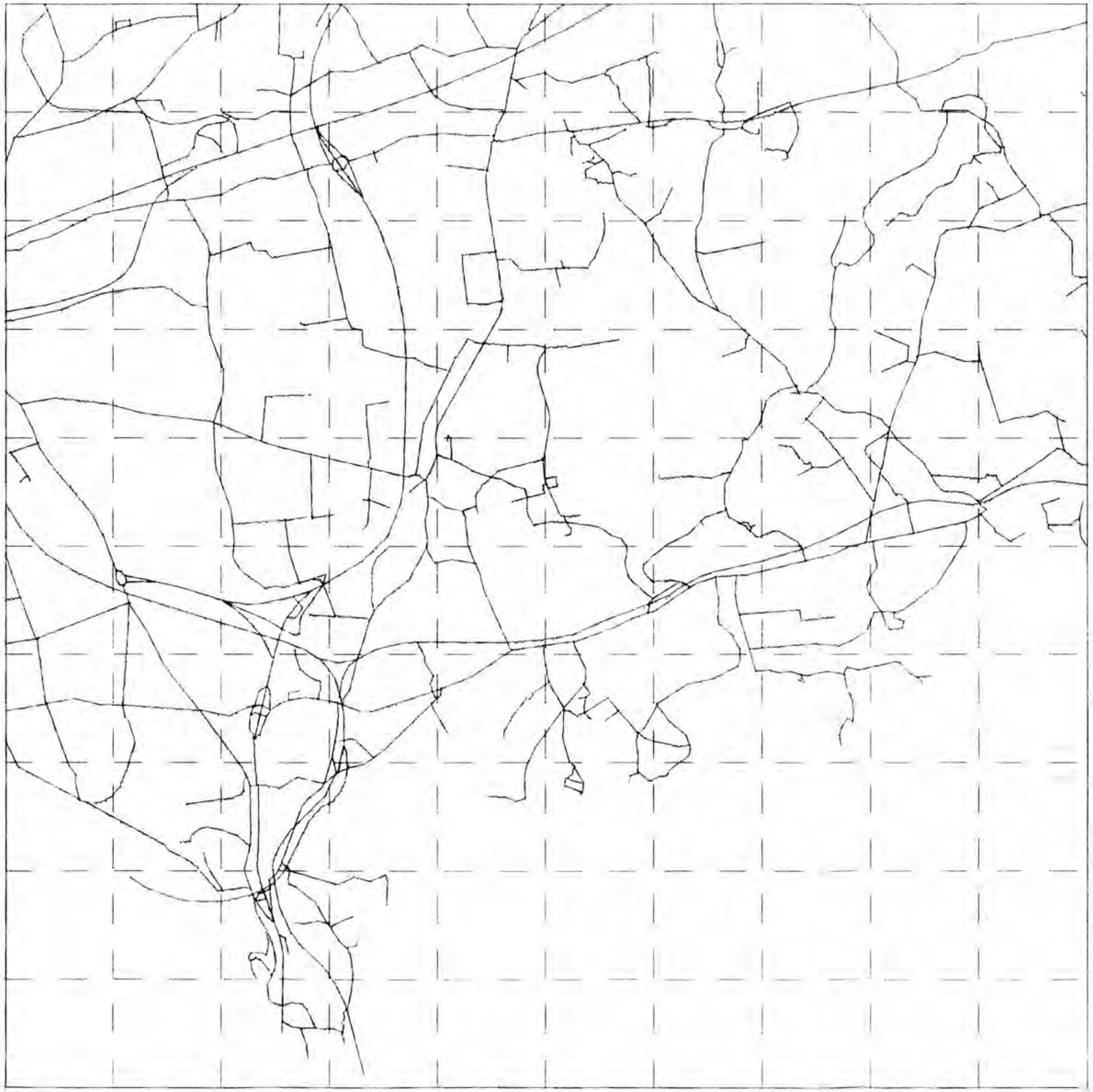
SOUTH WEST CORNER

EASTINGS	310000
NORTHINGS	680000

1:250,000

ROADS AND RAILWAYS

PLOT NO. 5



1:50,000

NO OF DATA POINTS

UNFILTERED 6773
FILTERED 1892

SOUTH WEST CORNER

EASTINGS 310000
NORTHINGS 680000

1:250,000

WATER AND ELEC. TRANS LINES

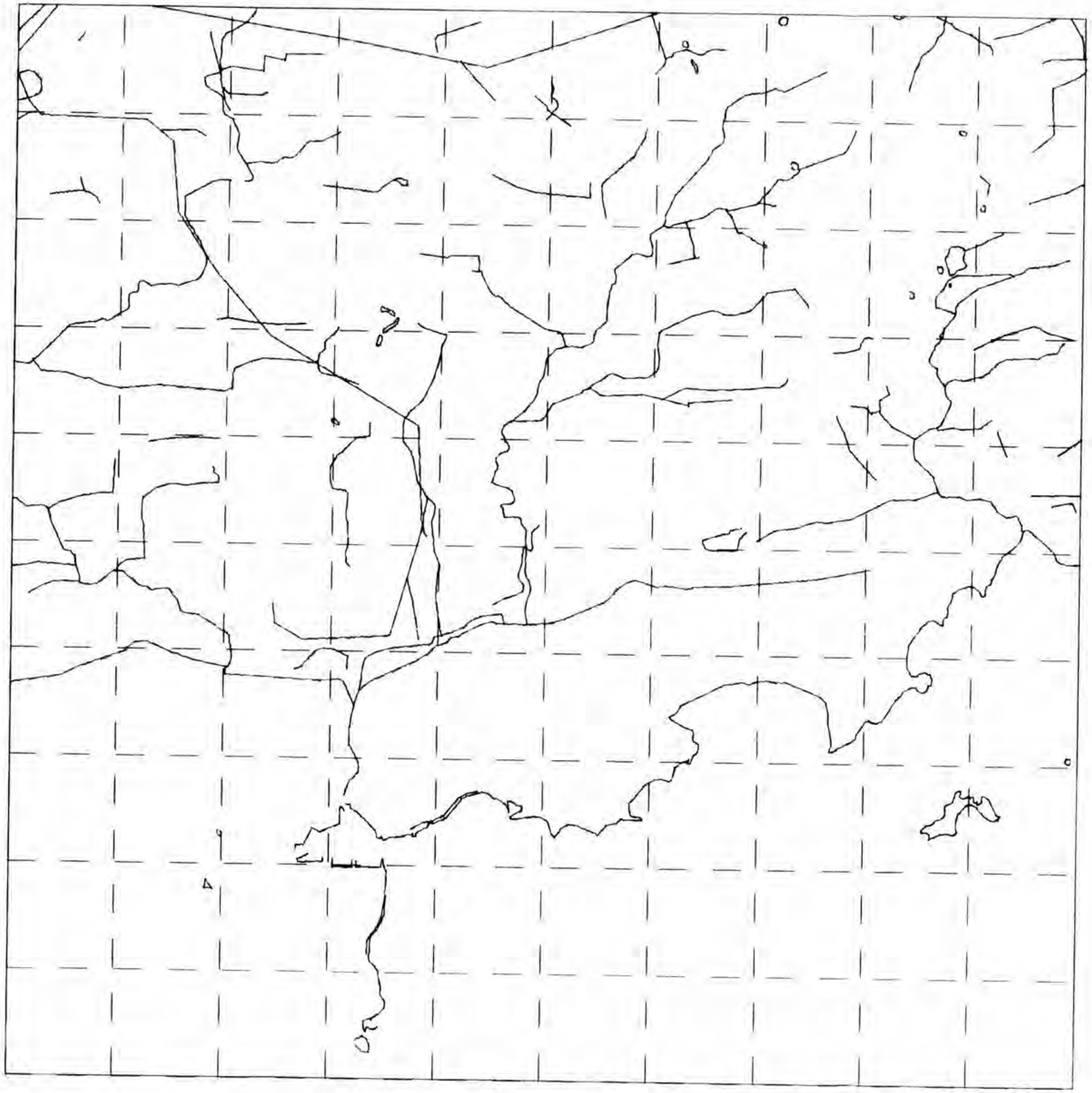
PLOT NO. 1

NO OF DATA POINTS

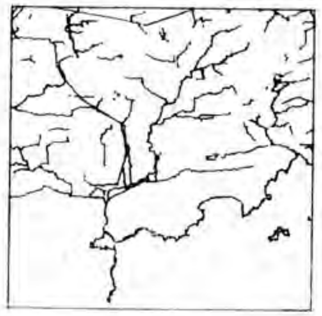
UNFILTERED 7074
FILTERED 3227

SOUTH WEST CORNER

EASTINGS 310000
NORTHINGS 680000



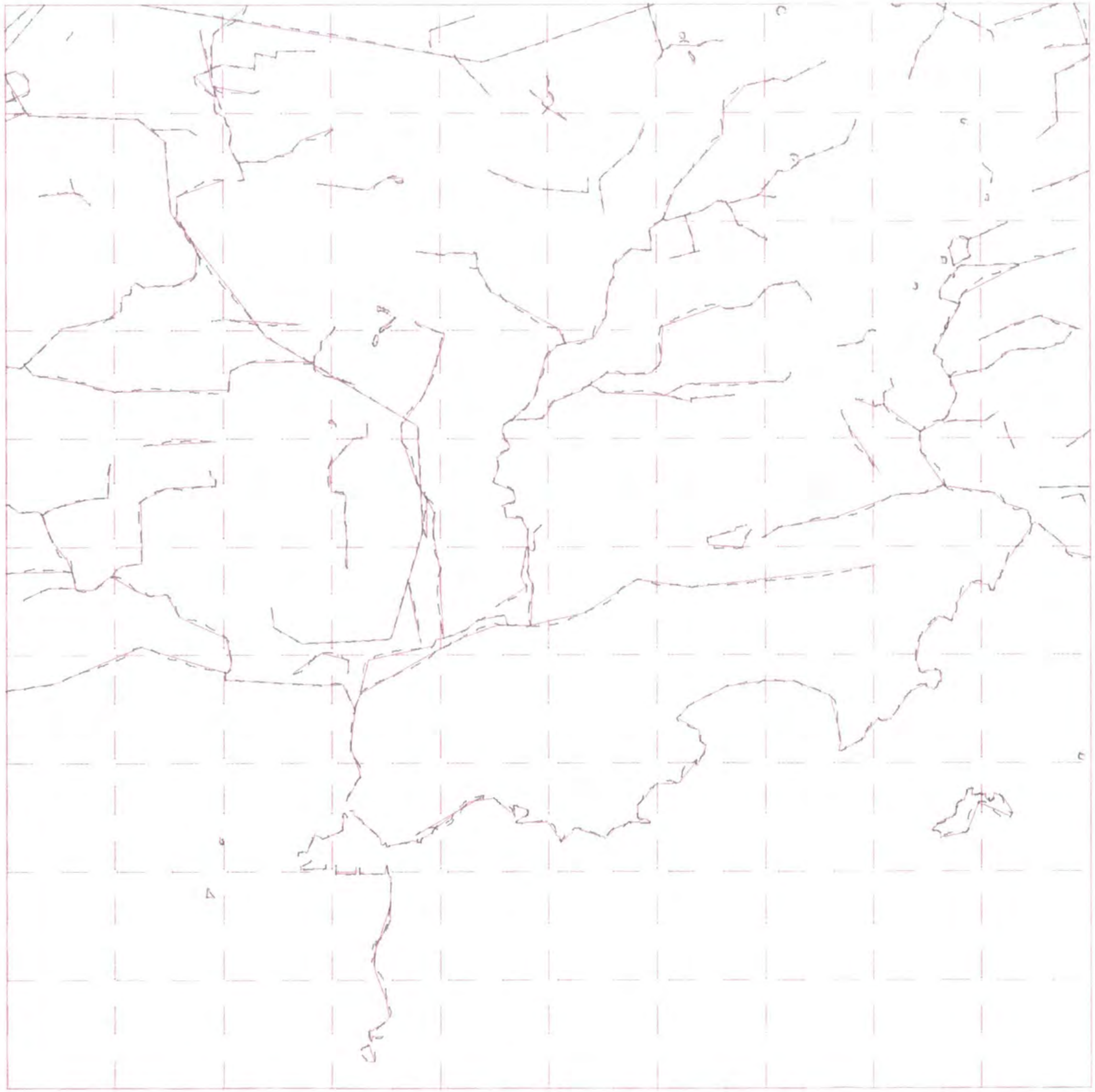
1:50,000



1:250,000

WATER AND ELEC. TRANS LINES

PLOT NO. 4



1:50,000

NO OF DATA POINTS

UNFILTERED	7074
FILTERED	1282

SOUTH WEST CORNER

EASTINGS	310000
NORTHINGS	680000

1:250,000