University of Southampton

The application of Geographic Information Systems to archaeology

with case studies from Neolithic Wessex

by David Wheatley MSc

Submitted for the degree of Doctor of Philosophy

September 1994

To my wife, Rosina.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ARTS

ARCHAEOLOGY

Doctor of Philosophy

The application of Geographic Information Systems to archaeology
with case studies from Neolithic Wessex

by David Wheatley

Geographic Information Systems (GIS) - computer systems for the manipulation and storage of spatial data - have commanded substantial attention in recent archaeological literature. Applications of GIS have been extensively discussed within some archaeological contexts in the United States and yet, although the interpretation of regional sequences of archaeological change still relies heavily on the distribution map, and on supposed spatial relationships deduced from these, GIS have made little impact on archaeological research in Britain.

This thesis contends that the slow uptake of GIS within British archaeology is a result of the unthinking importation of analytical methods which are sometimes inappropriate to British archaeology or to archaeological study generally. It is held that the application of GIS within British archaeology requires that analytical methods be devised in response to specific archaeological problems rather than, as at present, archaeological problems sought to fit the available methods.

The thesis begins by reviewing the historical development of GIS, and the application of GIS to archaeology to date. It then describes two case studies which explore in detail the potential applications of GIS to the Neolithic archaeology of the Wessex chalklands, firstly through the analysis of earlier Neolithic monument locations of the Avebury and Stonehenge areas, and secondly, through the analysis of systematic survey data from the Stonehenge area. In the process of exploring these sequences of change, new methods are devised to approach the specific questions which arise from this material.

These case studies result in some new methodological tools for archaeological regional analysis, notably the definition of a methodology for substantiating claims for intervisibility among archaeological monuments, and some recommendations concerning the portability of predictive models. The thesis also presents some specific archaeological conclusions, particularly that the causewayed enclosure sites of the study areas cannot now be regarded as 'central places'; that long barrow monuments around Stonehenge may have been deliberately situated for intervisibility; and that the densities of flint in the immediate area of Stonehenge are related to distance from the ceremonial monuments of the area.

Contents

List of figures

List of tables

Preface

This thesis stems from work carried out in the Department of Archaeology at the University of Southampton between October 1990 and October 1994, funded by the Science and Engineering Research Council (as was) under the Cooperative Awardes in Science and Engineering (CASE) studentship scheme. Additional funding for the project was provided by the Royal Commission for Historic Monuments (England) who acted as the CASE collaborator on the project.

My original interest in computer systems for the manipulation of spatial data stems from 1989 when, under the supervision of Sebastian Rahtz and as part of the Southampton York Archaeological Simulation System (SYASS) project, I wrote a graphically based excavation simulation program. This program has subsequently been successfully used in a variety of educational circumstances, and has been re-written (rather better than the original) for a variety of hardware and software platforms. However, some of the problems of representing geographic space in a computer system which were encountered in the design and implementation of this program were so intractable and fundamental to the context of archaeology that this led me to search for an information technology which would deal with the complex business of spatial representation. Soon after this, and rather coincidentally, Jim Milne of the Department of Geography at Southampton told be about Geographic Information Systems.

The link with the prehistory of Wessex can be traced to my studies as an undergraduate at the Archaeology Department of University College Cardiff (now the School of History and Archaeology, University College at Cardiff). While (very occasionally) studying for my degree, I was involved with the excavation and post excavation of a ploughed-out round barrow under the direction of G. Swanton and Dr (now Professor) John Evans in 1987. The surprising evidence which was encountered for the use and re-use of this site as a burial monument between the later Neolithic and the later Bronze Age, revealed how the place had retained its ritual and ceremonial importance over many generations and this forced me to a consider the structure of the ceremonial landscape in detail.

Acknowledgements

That this thesis reached a conclusion is a credit to my supervisor Dr Steven Shennan, whose enthusiastic support for the project was always rightly coupled with a healthy scepticism. Many ideas were formulated, tempered or discarded (whether they know it or not) following conversations with Dr Shennan, Dr Timothy Champion and Dr Arthur ApSimon, and I have gained much from their willingness to share their knowledge and experience. Roger Leach, Bruce Eagles and Mike Dyne from the RCHM(E) were constantly supportive of the project.

During the last four years, many friends and colleagues, particularly Kris Lockyear, Peter Durham and Anthony Firth have provoked me to explore particular directions and ideas. Additionally, I particularly thank Roy Canham for access to the Wiltshire SMR and to digital data from the Stonehenge Environs Project, and Julian Richards for collecting the lithic data in the first place, and for comments on some of this work in progress.

Finally, this thesis would not have been possible without the intellectual and personal support I have received over the last seven years from Rosina Mount. Her understanding of the Kennet Valley has been a source of inspiration, and her wholly unreasoning faith that I had something interesting to say has sustained my efforts whenever I felt inclined to despair.

Chapter 1
Introduction

1.1. Introduction

The central aim of this thesis is to investigate the extent to which Geographic Information Systems (GIS) can be useful in the context of British archaeological analysis, and to develop or extend GIS-based methods for investigating archaeological material. These primary aims will be undertaken within the context of the archaeological case studies of chapters 5 and 6. However, they will also be undertaken within a particular theoretical and philosophical framework.

This chapter will therefore provide some historical background on the adoption of GIS within archaeology (detailed discussion of GIS itself will be postponed until chapter 2) in order to place the adoption of GIS technology within British archaeology in context, and to develop the theoretical and methodological aims of the thesis in more detail.

1.2. GIS in Archaeology

Savage (1990) has characterised the three main areas in which GIS has found application within archaeology as:

> *(1) site location models developed primarily for cultural resource management purposes; (2) GIS procedure related studies and (3) studies that address larger theoretical concerns related to landscape archaeology through GIS methods.* (Savage 1990 p 22)

Setting aside those in Savage's class (2), which have dealt primarily with the development of GIS as an end in itself and have used archaeology simply as an example, there are two broad classes of applications: *management applications,* which are primarily concerned with methods for managing cultural resources, and *research applications,* which are concerned primarily with using GIS functionality to improve interpretations of archaeological remains. By far the most numerous examples of the application of GIS to archaeological material have been management applications, usually predictive models.

It is also the case that until very recently (see below) there have been practically no applications of GIS technology to archaeology outside the United States. These US origins of GIS applications to archaeology are intimately connected with the dominance of management/predictive modelling applications and must be understood before the impact of GIS technology in the relatively new context of British archaeological research may be predicted.

1.2.1. The North America/Europe contrast

The earliest recognition of the potential of GIS technology within archaeology was during the early 1980s in North America (discussed in detail by Lock & Harris 1991), where there had been previous attempts to use automated cartography to undertake archaeological tasks (e.g. Upham (ed.) 1979). By the mid 1980s several papers covering the principles of GIS for an archaeological audience and the possible applications in regional analysis had been produced (e.g. Kvamme 1985b).

Most north American applications have, until recently, been undertaken within the context of the public bodies (such as the National Parks Service, the US Department of the Interior and state Museum services) responsible for cultural resource management, rather in Universities and other research institutions. These curatorial organisations are often responsible for quite extensive areas of landscape with cultural value, frequently not entirely surveyed. As a result of this context, applications of GIS have concentrated on its use for predictive modelling as an aid to the curation of such areas (see e.g. Thomas 1975, Carr 1985, Kohler & Parker 1986, Altschul 1988, 1990, Ebert 1988, Ebert & Kohler 1988, Kincaid 1988, Kohler 1988, Kvamme 1988, Rose & Altschul 1988, Sebastian & Judge 1988, Thoms 1988, Carmichael 1990, Hasenstab & Resnick 1990, Kvamme 1990a, Warren 1990b). The trend towards GIS applications within north American archaeology culminated in the recent volume *Interpreting Space: GIS and Archaeology* (Allen *et al* 1990), in which a number of papers are presented detailing both methodology and case studies within archaeological research and management contexts.

Similar recognition of the potential of GIS in Europe has been considerably slower. Harris (1986) noted that GIS was having little immediate impact (although with exceptions such as Wansleeben 1988 for example), and even by 1990 Harris and Lock (in Allen *et al* 1990) were able to comment of the situation in the United Kingdom: '*The take up of GIS technology in the UK at the present time is somewhat limited.*' (Harris & Lock 1990). The reason for such a differential take up is hard to assess, but seems likely to be related to two factors. Firstly, archaeological management in the United States has a different character to that in much of Europe. European archaeological managers rarely have responsibility for areas of land of the scale with which the US National Parks Service deal, and where large coherent areas of land do exist in Europe there has often been a longer history of study leading to the availability of more detailed survey records. Thus the characteristic problem of predictive modelling - how to extend knowledge from surveyed areas into unsurveyed areas - has not been as pressing in the European context as some north American situations.

It is possible to identify other reasons for the recalcitrance of British archaeology to develop GIS applications. In a field as small as the application of GIS to archaeology, individual researchers may have a disproportionate influence on the development of a discipline. While a number of key individuals may be identified within US archaeology, each of whom contributed to the dissemination and adoption of the technology among colleagues and institutions, there have been fewer of these within British archaeology. It may be more than an historical aside that Dr Trevor Harris, who as early as 1986 was advocating the adoption of GIS in British archaeology (Harris 1986), left the UK to take up a post in the US soon after this date. Of course there have been, and continue to be, innovators within the field in Britain and Europe but it may be that such single events are highly influential in the developmental trajectories of technologies.

It is also worth noting, on the other hand, that the slow take up of GIS technology does not seem to be restricted to Archaeology. Instead it may be a more general phenomenon, at least partly related to the availability of the `raw material' of information technology: the information itself. It is worth observing that another facilitating technology, the Internet, was begun by the US government (more specifically by the US armed forces to prevent a nuclear war disrupting computer communications), and remains more widely used by US agencies than by those of other areas of the world. Coupled with the United States freedom of information act, this has ensured that government agencies devote substantial time and effort to the provision of the information which they collect. For example, the contrast between the provision of topographic data by the United States Geological Survey (who provide global access to most terrain data for the entire continent either free via the Internet or on tape and disk for the cost of the supply media) with provision in the United Kingdom is highly revealing. In the UK, the Ordnance Survey are the responsible agency, and topographic data is not treated as public information: the Ordnance Survey therefore charge customers for the data on an annual licence basis. Even when paid for, the data is not available via networked computer as in the United States. It is no exaggeration to say that for most academic institutions in the UK, it is easier and cheaper to obtain detailed topographic data for the United States than for the UK. It seems distinctly possible that the result of these differences in political policy have had a direct effect on the development of technologies such as GIS, which rely heavily on the availability of suitable data.

Despite this, however, considerable interest in GIS has been shown recently amongst European archaeologists and a number of general papers advocating the adoption of GIS as either a research or management tool have appeared (e.g. Wansleeben 1988, Harris & Lock 1990, Lock & Harris 1991, Wheatley 1993, van-Leusen 1993, Ruggles *et al* 1993, Chartrand *et al* 1993). Although published studies are still rather few in number, there does seem to be evidence that the application of GIS technology in the European context will have a substantially different character to that of north America. Castleford (1992), for example, discussed the way in which GIS may develop into a temporal as well as spatial analysis tool, time being a subject which should be (but according to him rarely is) of particular concern to archaeologists. Castleford recognised that:

> *`... GIS technology is limited and one of the most severe limitations is the lack of a temporal dimension. Because GIS are atemporal, the data with which they deal are merely tenseless snapshots of a single state of the data at a single instant of time.' (Castleford 1992 p99)*

Although he provides no solutions, save for suggesting that the *z*-axis in a 3D GIS may be used to denote time, this identifies three major problems associated with the temporal dimension in GIS: temporal modelling, temporal databases and GIS architecture. He also recommends that archaeologists should be discussing the temporal dimension in order to make current an archaeological perspective on time which will then impact upon those actively designing temporal GIS (TGIS) software. Ruggles has also recently presented a technical paper on GIS in archaeology (Ruggles 1992), a discussion of data structures for GIS applications in Archaeology. He claimed that geographical entities within GIS can be defined as Abstract Data Types (ADTs) with spatial attributes. As a result of this Ruggles claimed that GIS can be regarded not merely as systems for the manipulation and display of geographical data but also as support tools for conceptual modelling.

As well as these technical papers, several European archaeologists have recently undertaken studies involving the use of GIS. Most notable amongst these is the Hvar Study of Gaffney and Stancic (Gaffney and Stancic 1991, 1992). This study utilised the raster-based GIS system GRASS to study the prehistoric and Roman archaeology on the Dalmatian island of Hvar. Gaffney and Stancic created environmental overlays representing surface topography, soils, lithology and micro-climate and used the existing survey information about the island to generate archaeological overlays by period (Neolithic, Bronze Age, Iron Age, Greek and Roman periods) and by site type. Although other GIS applications have been recently discussed, and are in progress, the Hvar study certainly represents the most thorough investigation of the potential of GIS to have been undertaken by European archaeologists to date.

With the uptake of GIS in European archaeology, one main contrast is beginning to emerge. Whereas US applications remain dominated by the techniques of prediction and cultural resource management (e.g. Warren 1990a, Zubrow 1990b, Allen 1990), the different nature of European resource management has led to far less emphasis on predictive models and far more interest in (1) the use of GIS as a replacement for the traditional sites and monuments database and (2) the research potential of GIS.

1.2.2. Sites and monuments records

The upkeep of sites and monuments databases (SMRs) is one of the primary responsibilities of archaeological managers because the computerised SMR is one of the primary tools of archaeological resource management. In the UK, for example, most county records are now held on computer and reference to sites and monuments databases are a regular feature of the management of archaeology within the planning process. At a national level also the Royal Commission for Historic Monuments of England (RCHM(E)) maintains a computerised archaeological record (the National Archaeological Record). However it has recently been appreciated (e.g. Harris & Lock 1991, Lang 1993) that existing SMR databases could greatly benefit from implementation as spatial databases within a GIS or from being linked to GIS. The functional advantages which may be obtained can be grouped in terms of: improvements in *storage* of archaeological data; *integration* of archaeological data with other databases and in the *retrieval* of archaeological data.

Traditional text-based databases do not adequately store the geographical features of data objects. In archaeological terms this has meant that large linear monuments, such as long earthworks and roman roads are frequently represented without adequate geographic references. It is not uncommon to find that such features have been disassembled into many artificial small monuments to make storage of their condition and other attributes possible. An alternative tactic has been to arbitrarily divide the monument into a series of records and record a geographic location at regular intervals of, for example, 1km in order to allow for database searches based on location. Were the SMR to be implemented as a spatial database, then such monuments could be represented as line entities and the problem of search by location would be the responsibility of the GIS. A similar advantage may be obtained in the maintenance of archaeological or other landscape areas (SSSIs, World Heritage Sites, Conservation areas etc.). These can be represented as choropleths (rasters) or polygons (vectors) within the GIS, with the major advantage that the areas and perimeters of the areas are correctly coded.

One of the major advantages of GIS for archaeological databases is the capability to integrate archaeological data within the same data structure as other types of data about the environment. Aerial photography has been consistently used within archaeology but has, until now, required an entirely different structure to maintain and search. It is also straightforward within a GIS to store and manipulate soil, geological and topographic information which is currently used only in paper form within SMRs. Farley, Limp and Lockhart discuss the integration of remote sensing and traditional database management in more detail (Farley Limp & Lockheart 1990).

The most significant advantages of the additional storage and integration capabilities of GIS in the context of SMRs, however, will only be realised at the retrieval stage. GIS implementation of archaeological records will allow comprehensive retrieval of sites on spatial criteria. This should allow straightforward and intuitive retrieval of all recorded sites within threatened areas. Moreover, the existing data may be given 'added value' through the integration of existing records with other environmental data. For example, if the GIS contains appropriate data themes, it is a simple operation to add fields to the existing SMR database to record the topographic height of monuments (from digital terrain data), the land use of monuments (from remote sensed data) or the soil and geological substrate on which they occur (from digitised map data). Without GIS, each of these attributes would have had to be recorded individually for each monument within the SMR.

1.2.3. GIS in archaeological research

Although considerable interest has been expressed in the potential of GIS within archaeology, the actual use of GIS for research has concentrated on the extension and improvement of existing techniques. Most notable of these has been the use of cost surface calculations to improve site catchment analysis. Site Catchment analysis (Vita-Finzi & Higgs 1970) is a technique for analysing the locations of archaeological sites with respect to the resources which are available to them. The basic tenet of the method is that the farther from the base site resources are, the greater the economic cost of exploiting them. Eventually there is a point at which the cost of exploitation outstrips the return and an economic boundary can be defined at this point to define the *exploitation territory* of a site.

The exploitation territory or *catchment* of any given site can then be approximated as a circular area centred on the site in question - the proportions of given resources within this area can then be analysed and compared with the resources within similar territories from other sites. Having obtained the characteristics of each of the catchments to be compared, the different character of the sites should then become apparent. For example, if some sites are settlements used for cereal growing while others are bases for hunting activities, then the catchments of the settlements associated with agriculture might be expected to exhibit higher proportions of lighter, fertile soils. Using the facilities, transformations and operators which have been described so far it becomes easy to implement Site Catchment Analysis as follows:

1. Create point files for each site
2. Create choropleth overlays for each resource
3. Generate buffers of appropriate distance from a site
4. Mask the resource map with the buffer map
5. Perform an area operation the result
6. Repeat steps 3 to 5 for each site, then compare the relative areas.

An improvement on the basic site catchment technique presented above was used by Gaffney & Stancic in their study of the Island of Hvar (Gaffney & Stancic 1991), and more specifically presented by Hunt (1992). First the study used the GIS to generate site catchments for the Iron Age Hillforts of Hvar, improving on the existing notion of a *sphere of influence* by utilising a cost-surface analysis. This generated theoretical territories for the hillforts based not on linear distance but on the calculated energy cost of travel to any point in the landscape from a given starting point (in this case a hillfort). These were `calibrated' by walking in a straight line from the hillfort and measuring how far could be walked in a known time. These new catchments were then used to examine and measure the proportions of different soil types within each hillfort territory, and to suggest that the hillforts were in fact sited for different reasons: mostly agricultural (having strong associations with good soils), but at least one (Lompic with an anomalously low association with good soil) could demonstrably not have been sited for this reason. Next, the study used GIS to investigate the association between stone cairns and tumuli on the one hand and agricultural land on the other. This showed a clear association between cairns and the best agricultural land while there was a strong negative association between the cairns and the areas of dolomite (which are generally mountainous). This supported the contention that the cairns have a functional explanation, in terms of clearance of fields prior to cultivation rather than a symbolic one, although the acceptance of a functional explanation does not of course preclude that they may have also had symbolic value.

The third area of interest to Gaffney and Stancic was the location of Roman villas on the island. Like the cairns, the study demonstrated a clear association between the villa sites and the best agricultural soils on the island, and further demonstrated that the villas had a similar association with areas of `better' climate (better for agriculture that is). As in the case of the hillfort studies, one of the most useful aspects of this study was the way in which it identified `anomalous' sites, in this case a villa which seemed to be sited on more marginal land. Gaffney and Stancic suggest that the explanation of this site may be that it is of a later date and thus have been sited in more marginal land due to the intensive use of nearby better locations.

1.3. Theoretical and methodological themes

Although the applications of GIS to date have been interesting and beneficial to archaeology, the position that is adopted here is that they represent only a small fraction of the potential of applying GIS technology to European, and more particularly to British archaeology. It is also held that the impact of GIS on archaeology is not yet fully understood within either management or research contexts in British archaeology, and that some consideration of the theoretical impact of a widespread adoption of GIS is not only desirable but essential to fullfil the potential of the technology. To this end, a number of themes may be identified within which the impact of GIS may be considered.

1.3.1. GIS and theoretical neutrality

It may be useful to consider the relationship between technologies (such as GIS) and theory. It may be held that tools and technologies are merely the product of theories, that the tools are *theory neutral*. This is a position which has been repeated often enough to have become accepted as a truism. This is not the position which is held here; rather, it is an assumption of this work that the use and widespread adoption of tools such as GIS have identifiable effects on the methodology of a discipline, and on the theoretical basis of understanding within that discipline. The notion that GIS is a 'theoretically neutral' tool can and has been recently criticised (e.g. Wheatley 1993), and in other contexts it has been argued that tools and technologies in their role as components of the milieu in which the practice of archaeology takes place are never theoretically neutral (Richards 1986). Zubrow (1990b) puts this position clearly when he states that, when one begins to use a tool such as GIS '*one rapidly discovers that it is not equivalent to a mechanic changing wrenches. The changes are more profound*' (p68). Zubrow justifies this claim in the context of GIS by considering the differences in thinking which lie behind a raster and vector approach to GIS. A raster approach, he observes, locates the 'meaning' of the data equally within all the units of the raster, essentially allowing the entire landscape to be treated in the same way. A vector approach, on the other hand, places all the meaning (in the form of attribute data) in the lines and nodes of the data structure (see chapter 2 for technical details). The implication of the vector approach is that some element of the meaning of the structure of the landscape can be induced from study of the boundaries and interfaces betwen areas of the landscape rather than the land which is enclosed. Unlike vectors, which fundementally depend on the creation of boundaries and interfaces, raster systems cannot truly represent a boundary because, whatever the selected resolution, the raster must represent a line as a series of small areas. It seems clear that Zubrow has observed something interesting in this, and that the choice of methodology will inevitably effect the type of archaeology which is undertaken. The requirement in vector systems to define a boundary around whatever is interesting provokes classification, demarkation and site definition, while raster systems are better at modelling surfaces and continuous variation. Vector approaches to modelling cultural landscapes assume that there are clear distinctions to make between 'inside' and 'outside' places, while raster approaches almost expect to find a gradation which can be modelled as a process.

It is not only the selection of the type of GIS representation strategy which will have an effect on the type of archaeology which is likely to be done, and hence the theoretical development of the discipline, but whether or not GIS is adopted at all. Zubrow (1990b) is right to conclude that

> *'GIS will change the way archaeology is done and, as the work changes to make use of these systems, it will both expand and limit the types of archaeological research which are possible.'* (p72)

The truth is that GIS force the user to think in a particular way by representing a world which has particular characteristics. In GIS, as in maps, the landscape is generally presented, top down, orthogonally projected so that all places are equally distant or equally close. This is a position from which no contemporary human could have viewed the world and from which, in reality, no archaeologist can ever view the landscape. It is a privileged position, a position which emphasises the role of the archaeologist as 'all-seeing' and omnipresent. As Ingold (1993) has observed, it presents archaeology as

> *'a single picture independent of any point of observation ... as it could be directly apprehended only by a consciousness capable of being everywhere at once and nowhere in particular.'*

This does not (instinctively at least) encourage the archaeologist to get close to archaeological materials and to try to experience the material remains of the past as they would be lived and experienced. Although this charactaristic is not exclusive to GIS, and certainly is not new - it is a charactaristic which GIS share with plans, distribution maps and even with much of western representational art - it must be considered and accepted as the technology is used.

It has also been argued that the ease with which GIS manipulates environmental and economic data has led to GIS-based archaeologies consistently avoiding those situations where economic or settlement evidence is scarce, and to a concentration almost exclusively on areas and periods for which the majority of the evidence is economic, particularly on hunter-gatherer communities (e.g. Kvamme 1985b, Kvamme & Jochim 1989, Carmichael 1990). This general observation should not be taken to imply criticism of many of these studies by archaeologists who are justly interested in the economic aspects of prehistory and in hunter-gatherer communities. Nonetheless, it should be accepted that

> *"The use of GIS modules may lead to the unwitting exposition of an environmentally deterministic viewpoint of a type which has largely been rejected by most archaeologists".* (Gaffney and Stancic, forthcoming)

and that the over representation of examples of this type of archaeology should be of concern, and may encourage a return to the situation which caused Richard Bradley to remark of British archaeological accounts that *"Successful farmers have social relations with one another, while hunter-gatherers have ecological relationships with hazelnuts"* (Bradley 1984b).

Partly because of this avoidance of data with symbolic rather than economic meaning, there is a real danger that deterministic, generalising analyses may come to dominate the application of GIS to archaeology. Given the rejection of the theoretical position of which many of these procedures are characteristic, this would clearly be undesirable. As an example, it is necessary to reflect that the use of Thiessen's polygons in archaeology became unfashionable not because they were methodologically inadequate, but because they were used to infer 'political control' of sites over spatial areas - a naive theoretical approach which fails to recognise the nature of political control: that it is a social mechanism concerning social relationships between people. Thus it is social theory and not spatial science which will lead to a genuine understanding of political control. To read the recent GIS literature, one may be forgiven the belief that such techniques were rejected simply because they were methodologically limited.

None of this should be taken to mean that work concerning the improvement of techniques of spatial analysis is unnecessary or retrogressive. Clearly it can be extremely valuable to explore pattern in spatial organisation within human communities. However, archaeologists using GIS in this way should perhaps think more carefully about the application of these methodologies they are to avoid the trap of producing ever more complicated, but never more profound archaeologies.

1.3.2. Theoretical orientation

If GIS is not theory-neutral, it may be helpful to outline the position of this thesis at the start, rather than allow the theoretical orientation of the analyses to take the path of least resistance and be guided by the methodology. The archaelogical case studies undertaken here, and the interpretations of the results, are an attempt to integrate existing GIS methods into, and develop new ones within, a framework for archaeological explanation whose aim is a generalising one: to observe and understand long term patterns of change through the Neolithic period. It is recognised, however, that archaeological evidence is generally not of long term processes, but of individual actions and individual events and that, consequently, in order to understand long term cultural sequences, it is necessary to understand the human agency which created the material remains.

In attempting to understand the human individuals of the Neolithic, it is also recognised that it is the remains of their activities which archaeologists are forced to interpret rather than their beliefs, psychology or social conditions, and that it is consequently necessary to write what Thomas (1991) has termed a *'history less of people than of practices'*, and that this is, superficially, in opposition to the desire to understand human agency. However, it is held here that to write of the practices through which individuals of the Neolithic communities of the chalklands defined themselves, negotiated and re-negotiated their relationships to social institutions and reproduced and transformed social structures is not what Barrett (1994) terms:

> *'burying the individual, a move in archaeological writing which creates an unbridgeable distance between our own images of the past and the subjective and local intimacies of people's own lives as they were once lived'*

Rather, it is a means of positioning the individual and collective events, which are evidenced by the archaeological record, within a broader social context by explaining the consequences of those individual actions over a long period of time, and their relationship with each other over a wide spatial area. These consequences of actions and practices, which archaeological material provides evidence for, might be intentional or they might be unintentional in that the consequences might be foreseen by the individuals or they might be entirely unforeseen, but nevertheless direct results of those practices.

One of the consequences of this particular nature of archaeological investigation is that archaeology is not a subset of geography or anthropology, although archaeological perspectives may be of interest in either of these disciplines, but that these particular characteristics of archaeological investigations require that methods are developed in response to archaeological questions. To date, GIS methods have been almost exclusively imported into archaeology from other subject domains (geography, ecology, marketing) and these are, therefore, frequently inappropriate.

Most of all, however, it is the position of this thesis that, in order to interpret archaeological material in a satisfactory way, elements of both 'processual' and 'post-processual' archaeological theories are necessary. The methodological rigour of 'processual' themes such as behavioural and evolutionary approaches are entirely appropriate to an understanding of those patterns of change which occur over long periods of time as a result of the unforeseen consequences of social practices, while the analysis and understanding of the foreseen and intentional consequences of the practices of social actors requires that the critique of processual archaeology - most of all that it fails to account for individual, knowing actions - is taken seriously.

1.3.3. Scale

The nature of GIS analyses is such that the technology is applicable at a wide variety of spatial scales, varying from global scale to a few square metres. However, while the same technology can be applied to each of these scales of reference, it is clear that different methodologies will be appropriate at different scales, and even where the same method may be applicable at a variety of scales, different theories will be required to interpret the results. In addition, different data sources will be required to study, for example, a continent as opposed to a town and different theoretical perspectives on space at different scales will require that different questions are asked of different scales of GIS.

Under many circumstances, the scale of a GIS application would be determined by the requirements of a user. In non-research-oriented GIS systems the scope of the system will normally be defined by current political or practical boundaries: a County records office, for example, would naturally define the scale of study as the relevant County even though the county boundaries can have no relevance to prehistoric archaeology. This is what Firth (forthcoming 1994) refers to as an *institutional effect*: something which derives from outside the institutional context of archaeology (in this case from local government) but nonetheless has impact on the way archaeologists work. In this case the questions asked of such a system would be driven by the planning and resource management concerns of that particular application.

Researchers, however, are far less constrained by pragmatic concerns and will ask questions of a GIS system which relate to the period of study. This in turn must lead to a far more considered definition of the scope of the GIS study area. Two options are available to the researcher who wishes to formalise the selection of an area; the first may be characterised as a positivistic approach, in that it seeks to define the area in terms of the empirical relationships between those things observable in the present, while the second can be seen as a more normative approach, because it requires that the limits of the study area coincide to some extent with a distinction which was made in the past.

In the first case, the researcher can define an area based either on some form of statistical sampling theory or, pragmatically, on the grounds of data availability, much as the non research-orientated user might do. In this case, she or he must accept that this form of selection restrains and directs the questions which can be asked of the study. It would be possible, for example, to derive theoretical population densities for this type of area, allowing some investigation of economic and ecological themes, but it would not always then be possible to move from this stage of explanation to an exploration of the social and political organisation implied by such statements about an economic base.

Alternatively, a researcher can claim, in a rather more normative way, that the area of study was recognised as distinct by its occupants in the past. This is clearly a far better situation from the point of view of understanding social and political organisation in the past, and removes to a large extent the restrictions of the first method. Two drawbacks, however, are apparent if this course is adopted. Firstly, the case for identifying an area as 'real', and claiming that it was recognized as a discrete area in the past must be good. It is obviously impossible to categorically identify such a characteristic in the archaeological record, and this leads to the distinct possibility that, should the area in question subsequently be dismissed as a 'real' region, the underlying assumption of the entire study would collapse. This may not be a problem in certain circumstances, such as the Hvar study of Gaffney and Stanicic (Gaffney and Stancic, 1991), where the researchers selected an island (Hvar in Croatia) as the area of study, and could reasonably argue that this was a geographical area which was recognised in the past. This is far more of a problem for studies like those described below, where the regions are not so clearly geographically separated. Secondly, this normative mode of area selection may lead to a rather monolithic concept of political regions in prehistory, and place undue weight on one level of political organisation, when in fact a variety of levels of political organisation are equally valid. Recent discussion of the nature of cultural identity has highlighted the essential fluidity and dynamic nature of the concept of identity, and pointed to the possibility that cultural identity, and hence political organisation, may be latent within human groups until being defined or re-defined at times of conflict, in response to those particular circumstances.

1.3.4. Perception and landscape

The case can now be made that GIS technology *can* be effectively used in new archaeological contexts, and within theoretical frameworks other than those in which it has been most comfortably adopted. This requires, however, that archaeologists who wish to apply GIS techniques to archaeology undertake a reconsideration of the link between archaeological theory and practice so that the privileged viewpoint of GIS does not dominate the interpretation of archaeological material.

Most particularly, a number of workers have recently identified the *act of perception* to be a missing element in the approach of many studies. This absence effectively de-contextualises archaeological case studies presented with GIS, by tacitly presenting a theoretical model of the culture-environment interaction which takes no account of the cultural preconceptions and consequent constrained interpretations that human social actors bring to their physical environment before they interact with it. This perception prior to engagement, I would maintain, is what separates the environment from the landscape and suggests that GIS offers an ideal tool for the approach which has become known as landscape archaeology.

Landscape archaeology and landscape itself, however, are equally difficult to define precisely. A narrow definition is given by Crumley and Marquardt, who define landscape as *'the spatial manifestation of the relations between humans and their environment'* (1990 p73). This though may be an altogether too narrow definition unless environment is taken to include both the natural environment of soils, geology, and vegetation and the humanly modified environments of fields, buildings and monuments. Ingold (1993) suggests that landscape encompasses '*the world as it is known to those who dwell therein, who inhabit its places and journey along the paths connecting them*' which clearly incorporates perception of the landscape.

But although landscape and environment as intellectual ideas may be separated by the act of perception, in reality there is no clear distinction between them. Instead they are intimately related. Entirely 'natural' features such as hilltops or springs may have symbolic importance, while supposedly cultural features such as the ditches surrounding monuments may provide new food resources. What is required is a methodology which allows the entire landscape, cultural and natural to be stored and manipulated within a common framework.

The flexibility of the GIS here sets it apart from traditional thematic maps and distribution maps. Using the ability of GIS to manipulate complex surfaces, and to represent geographic data in new ways it should be possible to move away from the position of omnipotence, by generating alternative representations of the landscape which are not derived from abstract mathematical rules, but geared towards the experience of people. Through the ability to manipulate huge quantities of spatial information in complex ways, GIS holds the promise that it may allow the archaeologist to retain the best qualities of the map - a holistic summary of spatial variation - while escaping from the artificial position which this has always presented.

Cost surface analysis, for example, allows far more than a the modification of site catchment analysis described above, or the weighting of Thiessen's polygons, to which it has been applied so far. It may be used to generate alternative cognitive maps of landscape. Usually these are transformed according to the energy expenditure of traversing the terrain, such as a map of Britain coded according to the journey time from London, but these functional and economic parameters do not define the only possible transformations. Some years ago, for example, Rodney White was generating transformed maps of Britain according to the expressed preferences of school children in different parts of the country (White 1969) and these 'mental maps', which are as revealing about the regional variation in cultural attitudes as they are about migration patterns in contemorary Britain, are actually a simple manual version of cost-surface analysis using cultural parameters. While it is not possible to interview Neolithic people to elucidate their cognitive maps of Wessex, it may be possible to gain an insight into the Neolithic landscape using cost surface analyses in other ways. Similarly, GIS offers the opportunity to study relationships of visibility within the landscape. Without discarding the ability of maps to present summary information about entire groups of monuments, it may be possible to develop techniques for analysing and presenting the relationships of visibility in terms of visibility surfaces.

1.4. Stucture of the thesis

The rest of this thesis will attempt to substantiate some of these ideas, and to develop some new methods and approaches to archaeological remains. Before this, however, chapter 2 will present a more thorough account of the historical development of GIS, describe the special character of spatial databases and spatial analysis and describe the analyitical capabilities of GIS. Chapter 3 will then provide the archaeological background to the case studies, describing the material remains which characterise the Neolithic period of the area and reviewing some of the archaeological interpretations which have been placed on it. Chapter 4 will then describe the construction of the GIS databases, the sources of data and some of the procedures which were followed. The case studies themeselves are included in chapters 5 and 6.

Chapter 5 uses GIS methods in the interpretation of the locations of the ceremonial monuments of the period, particularly the situations of the first major monuments to be built in Wessex. Cost surface analysis is applied to the monuments of the Avebury and Stonehenge areas, and new approaches are developed from this. The locations of the long barrows are also approached with GIS methods in chapter 5, although the different nature of these monuments required a different approach to the analysis of their locations. This resulted in the analysis of the visibility and intervisibility of the long barrow monuments of the two areas, initially with existing methodology, but eventually leading to the development of a new method for the analysis of intervisibility among groups of locations.

Chapter 6 explores the analysis of systematic survey data with GIS, specifically the distribution of lithics recovered around Stonehenge by the Stonehenge Environs Project. This is approached first through the analysis of the spatial patterning within fieldwalking data itself and then through the analysis of the relationships between the lithic densities, the environment of the area and the other cultural features which constitute the Neolithic landscape of the area. Chapter 7 then attempts to draw together the separate strands of the thesis which were developed in chapters 5 and 6 and to reach some conclusions as to the role of GIS within archaeology.

Chapter 2
Geographic Information Systems

2.1. Introduction

The purpose of this chapter is firstly to define what is meant by a Geographic Information System (GIS) and provide background on methods and theory which relates to GIS, particularly to those attributes which distinguish it from other forms of information technology. The origins of GIS will be discussed, prior to an outline of the method and theory which underpins the creation of spatial databases. A section on the storage and products of digital elevation data will be included at this stage because elevation data represents the most widely used data theme in most GIS applications, this section will also serve to link the discussion of spatial databases and the analytical capabilities of GIS which will then be described.

2.2. Definitions

Geographic Information Systems (GIS) have proved difficult to define in precise terminology. This has not prevented large numbers of workers from trying to do so, and debate has ensued both in public and in private about what is and what is not a GIS. It is first perhaps worth citing the accepted `standard text' on GIS (Burrough 1986). According to Burrough a GIS can be defined as:

> *"... a powerful set of tools for collecting, storing, retrieving at will, transforming, and displaying spatial data from the real world for a particular set of purposes." (Burrough 1986)*

Some of the other possible definitions range from the minimal:

> *"A system for capture, storage, retrieval, analysis and display of spatial data" (Chorley 1987)*

> *"Automated systems for the capture, storage, retrieval, analysis and display of spatial data" (Clarke, 1990)*

to the more thorough:

> *"An information system that is designed to work with data referenced by spatial or geographic coordinates. In other words a GIS is both a database system with specific capabilities for spatially-referenced data as well as a set of operations for working (analysis) with the data" (Star and Estes 1990)*

An alternative tactic is to attempt a definition of GIS in terms of components. According to Marble (1987) a system is a GIS if it contains each of four component subsystems as follows:

1. *Data Entry subsystem which handles all problems with the translation of raw or partially processed spatial data (either in analogue or digital form) into an input*

stream of known and carefully controlled characteristics (the most common form of input today involves the manual digitising of maps);

2. *Data Storage and retrieval subsystem which accepts the input stream of spatial data and structures the database for efficient retrieval by the users of the GIS;*
3. *Data Manipulation and analysis subsystem which takes care of all data transformations initiated by the user and either carries out spatial analysis functions internally or provides a two-way interface between the GIS and specialised spatial modelling systems; and*
4. *Data Visualisation and reporting subsystem which returns the results of queries and analyses to the user in the form of maps and other graphics as well as in the textual form. Regretfully, most GIS designers have yet to realise the visualisation potential opened to them by the separation of data storage and data visualisation in a computer environment and an amusing goal of many GIS operations is to create map output that cannot be distinguished from traditional manual forms*

(Marble 1987)

Each of the above definitions has merit, although characteristically the definition by Burrough is both comprehensive and to the point. GIS, in reality, comprise a wide variety of hardware and software components sometimes solely connected because their main purpose is the manipulation of data which is referenced by geographic position.

For example, GIS software now operates on an extremely wide range of hardware platforms, from mainframes and large distributed systems through to desktop computers such as those based on the Intel 80x86 and Motorola 6800 families of processors. GIS software has also been developed on and ported to a wide variety of operating systems - the most popular being the main proprietary UNIX versions (e.g. ESRI's Arc/Info, the US Corps of Engineers' GRASS and several implementations of Tydac Technology's SPANS) but also MS-DOS (e.g. Clark University's IDRISI, PC Arc/Info and an assortment of smaller products), MS-Windows (e.g. MapInfo), IBM OS/2 (Tydac Technology's SPANS and SPANS MAP products) and the Macintosh (e.g. PAX technology's Geo-Navigator). As a direct result of, or perhaps because of, this wide variety of platforms, the systems now vary in functionality and purpose so greatly that attempts at further definition seem doomed.

2.3. Origins of GIS

Although GIS are a relatively new technology, the component elements of GIS have been available for some considerable time in other sub-disciplines of the Information Sciences. GIS is actually a combination of many information technologies including Computer Aided Cartography (CAC) which is sometimes called Computer Aided Mapping (CAM), Computer Aided Design (CAD), Database Systems and Computer Graphics. In addition the field draws on elements deriving from the manipulation of remotely sensed and photographic data (Image Processing) and from Spatial Analysis.

2.3.1. Computer aided mapping

The development of GIS can therefore be claimed in any of these disciplines, each of which have received extensive development for particular reasons during the last two to three decades. Computer Aided Mapping (CAM or Computer Aided Cartography: CAC), for example, was developed during the 1960s and 1970s to automate the production of paper maps.

These systems could reproduce (generally crude) spatial information on a computer screen, plotter or line printer. One of the major characteristics of such programs was that it became possible to produce new types of map such as `3-d' maps from contour or choropleth data. Early CAC programs included SYMAP (and SYMVU), GRID, IMGRID, GEOMAP and CPS-1 and were capable of various presentation tasks such as producing contour plots, 3d net plots and interpolating between known points. None had significant capability for analysis.

One system, SYMAP, was very popular in archaeology, particularly in the United States (see for example Arnold III 1979, Jermann & Dunnell 1979, Dean & Robinson 1979) and was typical of the systems available. SYMAP was developed by Harvard University's Laboratory for Computer Graphics in the early 1960s, and required firstly `a medium sized computer' and secondly a line printer. SYMAP was capable of accepting input from punch cards, tape or magnetic disk. As input, it needed a set of coordinates defining the boundary of the map area (OUTLINE), a set of coordinates relating to the actual data locations (DATA POINTS), a set of values which corresponded with these points (VALUES) and a set of cards specifying the options to be used. Output could be either isoline, choropleth or `proximal' map (a form of choropleth). Although the output was crude, and the resolution necessarily restricted by the cost of memory, SYMAP was capable of accepting ungridded point data as input and producing gridded (raster) data as output: in other words of obtaining a trend surface from the data.

By the early 1980s, CAC software had advanced in implementation so that better output was available from pen plotters, and electrostatic printers; input could be obtained from digitising tablets and scanners and computer technology allowed rapid viewing and updating of data. Computer aided mapping received substantial investments (particularly in the United States) where the motivation for creating these programs was listed by Rhind (1977). Slightly paraphrased, these were:

1. To make existing maps more quickly.
2. To make existing maps more cheaply.
3. To make maps for specific user needs.
4. To make map production possible without skilled staff.
5. To allow experimentation with different representations.
6. To facilitate map making and updating when data are already in digital form.

7. To facilitate analyses of data.
8. To minimise use of printed maps and thereby the effects of classification and generalisation.
9. To create maps that are difficult to make by hand e.g. 3d and stereo maps.
10. To create maps in which selection and generalisation procedures are explicitly defined and consistently executed.
11. To lead to a review of the whole map-making process, which can lead to savings and improvements.

By the late 1970s, CAC/CAM applications had been extensively developed, most particularly in North America. A parallel development to this is found in the development of CAD software. This was developed at the same time as CAC/CAM systems and shares many of the components of these. CAD, however, was primarily designed for the production of architectural and engineering drawings. In terms of the development of these types of software into GIS, two characteristics recur in these systems, and are of some importance to their eventual development into true GIS.

The Use of Overlays

In order to make map data more manageable within an information system, CAC/CAM and CAD systems generally adopt a conceptual model of spatial data which breaks down the whole database into a series of `overlays' or `layers' (the terminology presumably derives from the practice of drafting on several transparent layers of drawing film to group related elements of a complex drawing). In CAC/CAM and CAD systems, this refers to the practice of storing digital map information about different data themes in a series of different files. This led to the ability to store and combine (although only for presentation) information from any of a selected subset of the available data themes in order to generate different maps for particular purposes.

Representation by Vectors

Most CAC/CAM and all CAD software share the general property that they store pictorial information in a vector form, in other words the data files contain data which models the world in terms of points, lines and areas. Consequently these related technologies can be regarded as the origins of vector-oriented GIS systems. Each, however, lacks a data model which implicitly stores the connections between entities in geographic space.

2.3.2. Remote sensing

An alternative, and equally important, approach to the storage of spatial information derives from the expansion during the 1970s of remote sensing and the related discipline of image processing. The term `Remote Sensing' refers to the acquisition of data about the surface of the earth by means of sensors which detect (generally) electromagnetic radiation. Areas of the earth with different properties reflect different quantities of radiation in different areas of the spectrum. Image Processing refers to the manipulation of the images obtained in this way in order to extract valuable information from them (such as different vegetation types, or the location of roads). Although most commonly associated with satellite sensing, remote sensing can in fact also refer to sensors mounted on aircraft or simply to aerial photographs taken from aircraft, kites or hot air balloons. The most significant contributions of remote sensing, with regard to the development of GIS are probably the following:

Representation by Raster

Unlike digital drawings, remotely sensed images generally comprise a regular square grid of sampled locations, each of which corresponds to an area of the earth surface. This representation of reality is more intuitive to a worker familiar with photographic images, and can be contrasted with the vector approach deriving from workers who were more concerned with drawings. Approaches based on storing spatial information as both rasters and vectors are now extensively utilised within GIS systems.

Rectification and Georeferencing

In order to relate remotely sensed images to real locations on the earth's surface, software technology was developed to correct the optically skewed images produced by remote sensors into regular grids. This allowed every element of the image to be accurately referenced to a point within a known coordinate system. Although not exclusively derived from remote sensing, this concept of georeferenced data can be seen as a critical step towards the integration of spatial information within a unified information system.

2.3.3. Hardware advances

The software technologies discussed above were steadily refined during the 1970s and early 1980s in parallel with the increasing development of computer hardware. The development of integrated spatial information management and analysis systems from these existing technologies was only made possible by these advances in hardware. The most important of them is almost certainly the decrease in cost of computers which followed the introduction of the IBM PC in 1982, and the availability since that date of increasingly powerful desktop computers. The three most significant improvements which made GIS feasible on desktop computers, and therefore allowed the discipline to build up a `critical mass' of users since the end of the 1980s are probably as follows:

Disk Storage Capacity

Geographic data, particularly raster data, can occupy enormous quantities of storage space. At the start of the 1980s, mass storage on computers was essentially based on magnetic tape. By the mid 1980s, storage space on desktop computers was available in the form of hard disk units, storing in the order of 10-20 Mbytes (Million Bytes) of information. Most importantly this could be transferred to and from the electronic memory (RAM) of the machine at a speed which made feasible the processing of data files in the order of kilobytes (thousands of bytes). As an illustration, the IBM PC released in 1982 was supplied for its mass storage with one 360Kb floppy disk drive. This document was written on a comparably priced machine bought in 1991, which is supplied with a 100Mb (100,000Kb) hard disk with an access time several hundred times faster than the floppy drive. Without this development, the processing and storage of raster information at a sufficiently high resolution to usefully represent a geographic area would not have been possible.

Processors and Random Access Memory

Technological advances in the microprocessor industry led during the 1980s to a vast decrease in the cost of Integrated circuits (ICs). One of the most important applications of ICs was the development of Random Access Memory circuits (RAMs) which underwent a similar development to that of disk storage technology, allowing the RAM memory sizes of all computers to be increased accordingly. Coupled with a comparable increase in the performance of processor circuits in terms of the number of instructions processed per second, this enabled computers to reduce processing times to a level where the manipulation of geographic data overlays interactively was possible.

Graphics and Output Devices

At the same time as the above advances were taking place, similar electronics technology advances were being made in the quality of output available from computers. To take the example above, the original IBM PC was supplied with a graphics output device capable of 25 lines of 80 characters. Soon afterwards, as an expensive option, it was possible to obtain a `Colour Graphics Adapter' and suitable CRT monitor which was capable of displaying 640 pixels by 320 pixels in four colours. A similar priced graphics subsystem option for a desktop computer ten years later might provide a display of 1024 by 1024 pixels in at least 256 colours (and more expensive systems can render up to 1768 by 1768 in any of many millions of colours). This development of graphics hardware was accompanied by the development of graphically-oriented operating systems for computers (examples include the single-user Microsoft Windows system, the multi-user OS/2 Presentation Manager and the distributed windowing environment associated with UNIX System V and derivatives: the XWindows protocol) which made possible for the first time the interactive visualisation of geographic data at a meaningful resolution.

2.3.4. GIS as a technology-led discipline

It should be apparent from the above, that the development of GIS systems can be characterised as *technology-led.* It could be reasonably argued that the elements for GIS were all developed within separate application software as early as the late 1970s, in the form of image processing systems, Database management systems and CAD/CAM systems. What has happened since that period can be divided into two stages: the improvement of existing capabilities and the integration of the functionality of different systems within one environment.

Improvements to existing systems have mostly been discussed above in relation to hardware. CAD/CAM systems, for example, became capable of handling larger files, processing them more quickly and rendering them in more detail. Image Processing software was allowed to became more sophisticated in the algorithms used for classification and rectification as well as becoming capable of utilising images of higher resolution. Improved Database management systems utilised the increased hardware capabilities to fully implement the relational model of data structure. All of these developments suggest that the development of GIS systems was in fact inevitable during this period, but that the final step towards the creation of GIS was only possible when hardware and software advances allowed the integration of all of the elements into one system.

As discussed above, there seems little point in extending the available definitions of GIS any further, except to emphasise that it is the spatial nature of GIS which set them apart from traditional information systems. Traditional database systems are incapable of easily referencing data on different subjects held about objects which are in the same place, and thus have no inherent concept of place; this also means that non-spatial information systems are not underpinned by the concept of distance, or of proximity, or of area or any essentially spatial phenomenon. GIS are information systems dedicated to georeferenced data and are thus models of physical space in a way in which non-spatial information systems cannot be. As Kvamme (1989) puts it:

> *'GIS are distinct from traditional database management systems because of this spatial referent; indeed it is the geographical structure that give GIS added capabilities over traditional database management systems' (Kvamme 1989)*

Because none of the elements of GIS systems are new, it has been suggested that the GIS is merely an amalgam of the technologies discussed above and does not represent a significant advance in information handling. GIS, it is argued, are simply spatial databases with graphical front-ends. Other geographers however have taken a different view, such as Marble (1990), who comments:

> *'Most tool innovations in science are of a minor nature and, although the cumulative impact of these changes may be impressive, it is only rarely that a change occurs which revolutionizes a field to the point where many of the things we do must be looked at from a completely different viewpoint. It is my contention that GIS is beginning to do this for geography, and that it will also do so for those portions of the social sciences which are concerned with, or at least should be concerned with, the spatial aspects of human society.' Marble 1990, p9.*

and the recent Chorley report undertaken in the UK by the Department of the Environment (DoE 1987) had no doubts about the potential impact of the GIS:

> *'It is the biggest step forward in the handling of spatial information since the invention of the map.' (DoE 1987 p8)*

It must be accepted, however, that none of the four elements of GIS listed above are in themselves novel: data entry for geographic data is part of CAC/CAM, as are data storage systems; manipulation of spatial data has been performed with computers for some time in Image Processing systems and data visualisation for geographic data has also been available for many years.

Here it is accepted that none of the technology which comprises a GIS is in itself original, but it is nevertheless contended that GIS is greater than the sum of its parts, and that the integrated combination of geographic database, spatial analysis system and data visualisation tools which comprise a GIS does consequently represent a technology with the potential to revolutionise the spatial sciences.

2.4. Spatial databases

A wide variety of specific conceptual storage structures and implementations may be utilised in the construction of spatial databases (although see Marble 1988 for general procedures for their design and implementation and Peuquet 1988 for some of the issues involved with the selection of data models). What follows is a discussion of the general characteristics of spatial databases, and a more specific discussion of the most widely used conceptual model of geographic space.

2.4.1. The use of overlays

As suggested above, one of the primary characteristics of GIS systems, inherited from earlier CAC/CAM systems, is the notion that geographic data is organised into different files or overlays. These files each contain georeferenced data relating to a defined subject area or theme. Each overlay can therefore be regarded as a thematic map. Typical examples of thematic overlays which can be stored in a GIS might be:

1. Elevation Model: A Digital Elevation Model overlay contains data describing the physical structure of the surface of the land in a particular area. This may be stored in vector form (as contour lines or sample points) or in raster form (as a regular grid of heights).

Typically the numerical values stored in a DEM overlay are height above sea level in metres.

2. Water Features Overlay: An overlay representing water features might contain a representation of the course of rivers and streams (normally stored as line vectors).
3. Property Boundaries: may be stored as topologically linked polygons, each polygon being assigned the name of the owner as a defining attribute. This area map could then be stored either as a vector polygon map, or as a raster based map, similar to the structure used for an elevation model.

It should be clear from the above that the concept of an overlay is extremely flexible, and can be used to express any sort of thematic data grouping which is found to be convenient. In archaeology, different overlays may be used not only to represent different non-archaeological data themes as above, but also to distinguish different archaeological datasets. Overlays are frequently adopted by archaeologists in order to perform operations on data from different periods (see e.g. Gaffney and Stancic 1991). An alternative approach may be the creation of different overlays for groups of sites which have particular legal status, for example Scheduled Ancient Monuments (SAMs), Listed Buildings or Sites of Special Scientific Interest (SSSIs).

2.4.2. Vector and raster storage

The central component of all GIS is the spatial database, which consists of a digital record of the geographic entities in the area of interest. The representation of these components can be achieved in two fundamentally different ways: either with a raster or a vector method. As in all areas of computer graphical representation, each method has particular strengths and weaknesses. Occasionally systems allow the use of a mixture of both types of representation with utilities to translate data from raster to vector and from vector to raster. This is becoming increasingly common because it offers an flexible way of manipulating spatial data.

In a vector system the locations, courses or boundaries of features are stored as a series of coordinate pairs which are connected to describe fundamental mappable objects. In a raster system, the graphic representation of the features and the attributes of the features are one and the same thing: the study area is divided into a finite number of (usually square) cells in which are placed the recorded attributes of the corresponding geographical location.

Raster systems are typically data-intensive unless the storage is at a low resolution, or mitigated through the use of data compression or a hierarchical storage tactic such as the quadtree structure. Raster systems, however, have the advantage that they store all geographical information in the same uniform and predictable way. Raster systems are therefore suited to the analysis of continuous geographic phenomena such as terrain elevation, while vector systems are more suitable for the storage and analysis of discrete geographical systems such as river networks or pipelines.

Most GIS systems incorporate elements of both vector and raster representation of data. Although the GIS systems IDRISI and GRASS are both generally regarded as 'raster' systems because they both perform the majority of their analysis on raster data, each has the capability to read and write vector data files. Similarly ARC/INFO, which is regarded as an essentially 'vector' GIS system contains modules for creating and manipulating data as rasters.

2.4.3. Georeferencing

Georeferencing refers to those elements of the data which determine the geographic location of data primitives. A spatial database must include this extra georeferencing information in order to determine how to undertake spatial calculations of distances and areas. In some instances of small study areas which have been consistently recorded within an understood projection plane this georeferencing can simply consist of an assumption that the coordinates may be manipulated with simple Cartesian geometry, either ignoring effects of the curvature of the earth as acceptable errors or accepting that the selected projection plane has appropriate Cartesian properties. However, in larger areas or in situations where the locations of entities are recorded in different coordinate systems it is necessary for the system to store both the coordinates of the entities and the projection through which they are defined. If this is not undertaken, and the characteristics of the projection plane used are not incorporated into the information system, errors in distance and area calculations are introduced.

While the mathematics of projection are far from straightforward and are beyond the scope of this thesis, it is worth briefly considering the major projection types in order to understand the basis of the projection used throughout this thesis. Map projections consist of two main elements: firstly the surface of the earth is estimated through the use of a geometric description of a spheroid (really an ellipsoid of which a number are in use) and secondly the surface of this spheroid is projected onto a flat surface to generate the map.

Projections

There are a huge variety of methods for undertaking this second stage, each method producing a map with different properties. In a *cylindrical projection*, for example, the lines of latitude (parallels) of the selected spheroid are simply drawn as straight, parallel lines. Because the parallels become shorter as we move away from the equator, this means that the lengths of the parallels are progressively exaggerated towards the poles. To maintain the right angled intersections of the lines of latitude and longitude, the lines of longitude (meridians) are also drawn as parallel lines. This inevitably means that the meridians never meet at the poles as they do in reality and that the polar points of real space become lines the length of the equatorial diameter of the spheroid in projection space. The cylindrical projection thus maintains the correct length of the meridians but this also means that areas close to the poles become greatly exaggerated in an east-west direction. A *transverse cylindrical projection* is created in the same manner, but the cylinder is rotated with respect to the parallels and is then defined by the meridian at which the cylinder touches the spheroid rather than the parallel.

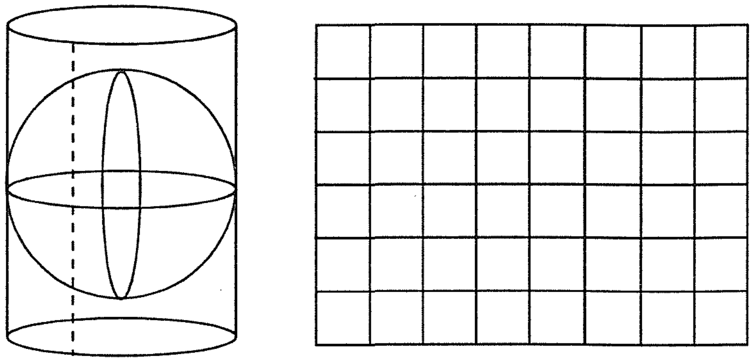


Figure 2.1.Cylindrical projection.

The exaggeration of the areas of the polar regions can be offset by reducing the distance between the lines of latitude as they move away from the poles. This is called a cylindrical equal-area projection or more commonly Lambert's projection. However, this projection distorts the shapes of regions very badly, and cannot represent the poles because they are projected into infinitely small areas.

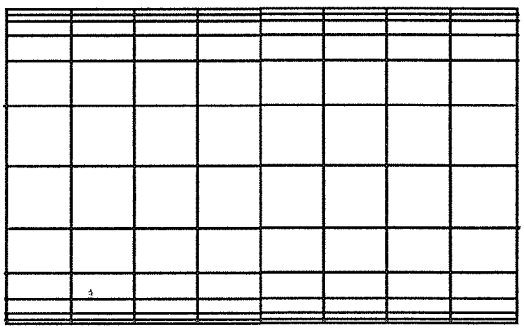


Figure 2.2. Lambert's projection

One other cylindrical projection is worth inclusion. A Mercator projection exaggerates the distance between meridians by the same amount as the parallels are exaggerated in order to obtain an orthomorphic and conformal projection (orthomorphic because shape is preserved so that squares in the projection plane are truly squares, and conformal because the projection plane parallels and meridians always intersect a right angles). The poles therefore cannot be shown because they are deemed to be infinitely distant from the equator, and the projection exaggerates the apparent areas of northerly areas far more than either of the previous projections - at 60 degrees latitude this exaggeration is 4 times, at 70 degrees 9 times and at 80 degrees it becomes an exaggeration of 32 times the same area if it were shown at the equator. A transverse Mercator projection is similar except that it is based on the transverse cylindrical projection. This means that instead of the north-south exaggeration of the traditional Mercator projection, the distance of areas to the east and west of the defining meridian are exaggerated in the projection plane.

Cylindrical projections serve to illustrate some of the distortions which are introduced into the representation of space by projection, but there are a great many other forms of projection which are not based on the cylinder. Conical projections are based on the assumption that a cone is placed with its vertex immediately above one of the poles, and which touches the globe at one of the parallels called a standard parallel. It is also possible to create a conical projection where the cone is allowed to touch the globe at two standard parallels. Conical projections therefore show the parallels as curved parallel lines of decreasing size as they approach the poles, and the meridians as straight lines which meet at the poles.

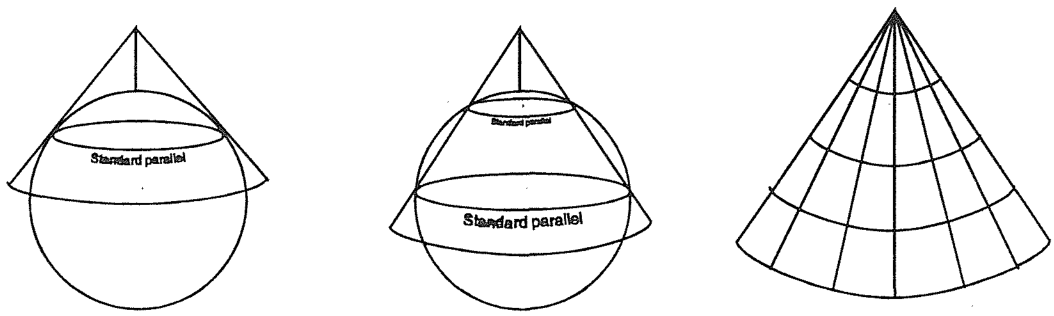


Figure 2.3. Conical projection with one (left) and two standard parallels (middle) and the general form of the resulting projection (right)

All of the above projections represent the meridians as straight lines. In some projections, however, the meridians are curved. For example, a cylindrical projection where all of the parallels are correctly drawn and divided results in an equal area projection plane with curved meridians called Bonne's Projection. Two entirely separate families of projections called two-world equal area projections and zenithal projections are also occasionally used for particular purposes. As none of these has found popular use in the UK, none are described here.

The national grid system of the United Kingdom

In the United Kingdom, a *transverse Mercator* projection (see above) is used by the Ordnance Survey as the basis for all maps. This is projected from the 1849 Airy spheroid which models the earth as an ellipsoid with equatorial radius of 6377563.396 metres and a polar radius of 6356256.910 metres. Transverse Mercator projections are particularly suitable for areas with a north-south extent as the north-south distortion of areas is removed in favour of an east-west distortion. In the particular case of the UK grid the projection is made from the 2 degree (west) meridian (close to the east/west central axis of the island). A false origin is defined by specifying the crossing of the 2 degree (west) meridian and the 49 degree (north) parallel as x

= 400000 metres and y = -100000 metres. An additional scaling factor of 0.996 is used to eliminate the east/west area distortion of the projection.

As a result of this combination of projection and scaling factor, for most areas in the United Kingdom, it is acceptable to use Ordnance Survey national grid references (expressed in metres) as the basic unit for all georeferencing. These coordinates may then be used in all Cartesian area and distance calculations with negligible errors.

2.4.4. Geographic data primitives

The basic types of geographic data which are stored and manipulated within GIS are usually given as *point*, *line* and *area* entities (e.g. Martin 1991), although an additional category of *surface* might be included to distinguish categorical (area) data from continuous spatial variation. These basic entities can be used to represent any mappable object, the essential difference between them being one of dimensionality. Although Martin (1991) refers to *area* entities as 2-dimensional, which is reasonable, he defines *surfaces* as 3-dimensional. This may be problematical as genuine 3D GIS software becomes available, therefore it seems more sensible to regard area maps (choropleth maps) and surfaces as different levels of data relating to 2-dimensional geographic data.

The dimensionality of a geographic data primitive has an affect on the 'legitimate' relationships it may have with other primitives in the database. Because a point is *zero-dimensional*, it cannot contain or enclose other entities. Similarly because lines are considered to have only length, but no width - *one-dimensional*. It is a conceptual requirement of these facts that, for example, all point entities must be either one side of a line or the other and that there is no location which is on neither side of a line. Area entities have both length and width and area referred to as *two-dimensional* entities.

Surfaces present something of a problem within this model: intuitively it may be expected that surfaces would be *three-dimensional* as they have both extent and height (represented as the attribute). However, the same could be said of area data and that is defined as two dimensional. True three-dimensional representation of objects must allow for 3d-surfaces *and* solids to be represented: in other words that true three-dimensional representation should have volume as an inherent characteristic. This requires an extension of the *spatial* component of the data model into three dimensions, not merely the use of an *attribute* as a z-coordinate. Surfaces in 2-D GIS systems are stored with the third dimension recorded as an attribute of a *two-dimensional* entity so to separate surfaces from true 3-d GIS representation surface entities are frequently referred to as *2½-dimensional* entities.

It should be apparent that the method of storage of a data overlay does not in itself define the type of geographic data primitive. Both choropleth maps and surfaces may constructively be stored and manipulated as rasters or as vectors. In vector terms, a choropleth map is stored as topologically linked polygons and surface data as a Triangulated Irregular Network (TIN), while raster systems store both areas and elevation matrices as bitmaps. It is unusual but not exceptional to store point data in the form of single cell values in rasters, and lines as connected sequences of raster pixels, as opposed to the more usual vector representation. In each case the level of geographic data is unrelated to the storage strategy.

2.4.5. Components of geographic data

One of the characteristics which most clearly marks out GIS from other similar computer systems is the use of a fully topological database to store geographic data. The following discussion mainly follows Burrough (1986), but also draws on Star and Estes (1991) and Martin (1991).

In a fully-topological geographic database, the data is divided into three notional components: location, attribute and topology. The *locational* component of spatial data is the information which determines the location of the object in physical (geographic) space. The *attribute* component of geographic data refers to the data model for all other (non-geographic) characteristics of the data. Each of these can be adequately manipulated in CAC/CAM or CAD systems. However the *topological* information refers to the connections between objects and the relationship of geographic primitives to one another. It is this aspect of the data model which is original in GIS.

On a paper map these relationships between the geographical objects are indicated visually. Thus *regions* or *lines* which pertain to the same legend unit are represented in a particular colour or line type. In a spatial database, it is necessary to code the connection between these entities in an explicit way, and although the spatial relations between regions or between lines within one legend unit may be easily interpreted visually, it is necessary to consider these relationships when building data structures for spatial databases.

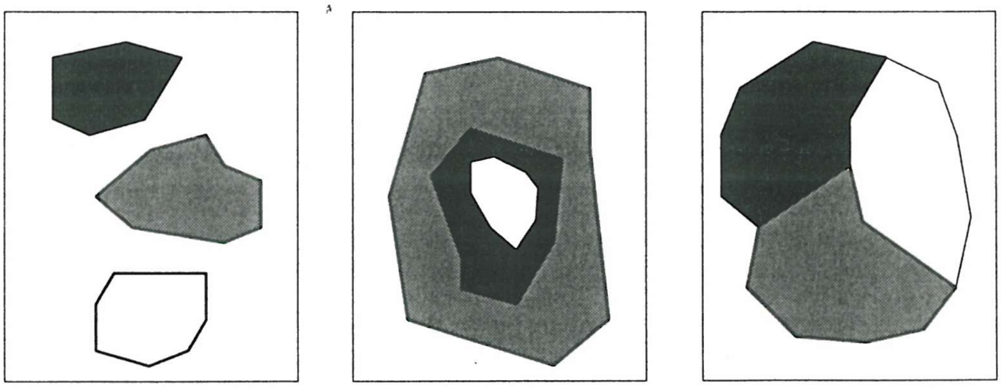


Figure 2.4. Possible topological relationships. Left, unconnected. Centre, regions contained island fashion within other regions and Right, regions connected by shared boundaries.

Consider the three regions shown in figure 2.4. Visually the relationships between the three regions in each case is straightforward: on the left the regions are unconnected; in the centre the white region is wholly contained by the black region which is itself wholly contained by the grey region while on the right the three regions are connected through shared boundaries. Each of these cases may occur in particular types of geographic situation and so each must be explicitly coded in a spatial database system.

The locational component of the data

The basic graphical entities described above may be used to represent real-world objects. Points naturally represent objects which can be regarded as having no spatial extent: archaeological finds for example or, at a larger scale, archaeological sites; consequently they are regarded as dimension 0. Lines generally represent roads, railways, rivers or more theoretical pathways such as least-cost distances, they have a length dimension but no spatial area; thus lines are one-dimensional. The difference between areas and surfaces is slightly less obvious but may be illustrated by considering an example: a choropleth map of parishes, each coded with a unique code is a map of areas, whereas an elevation model is a surface. This difference is essentially similar to that between nominal or ordinal level data, where classification values are assigned to geographic entities but the numeric difference between the values is not meaningful and ratio or interval level data where the values of the numbers themselves relate to reality.

	Raster	Vector	Data level
Point	cell	point	depends on attribute
Line	line of cells	line	depends on attribute
Area	groups of cells	polygon / arc-node	nominal / ordinal
Surface	matrix of cells	network / TIN	interval / ratio

Table 2.1. Representation of mappable objects in a vector and raster system, together with the corresponding aspatial level of measurement.

In the example it can be seen that the code values for parishes cannot be treated as numbers *per se*, because to perform addition or subtraction with them is pointless; the elevation values in a DEM, however, have genuine meaning in terms of their value so subtracting one elevation value from another provides the difference in height between two points, a meaningful result.

It might also be noted that the relationship between representor (graphical primitive) and represented (real-world object) is not fixed but will depend on the context of the study. For example, depending on the study, an archaeological monument may be represented conveniently as either a point, or a line or as an area entity. Central in the decision is the scale at which the study is undertaken. Clearly at a national scale, it is acceptable to regard almost all archaeological monuments (with the exception perhaps of Hadrian's Wall) as point locations. At a slightly smaller scale, however, particularly large or linear monuments will have too greater spatial extent to be regarded as single places and will have to be modelled as line or area entities.

The attribute component of the data

The spatial database must also store *non-spatial* data which describes the properties of the spatial objects. These are referred to as the *attributes* of the entities. Typical examples of attributes which may be associated with different types of entity are given in table 2.2

	Example	Spatial data	Attribute data
Point	site	coordinates	type, date, size
Line	river	location, direction	name, flow rate
Area	soil type	boundary	name, colour
Surface	DEM	location	elevation

Table 2.2. Examples of the spatial and attribute data of some spatial entities.

In many cases a single spatial entity will have a number of attributes associated with it. Thus a point entity which is used to store a house may have attributes to record its length, width, height, colour, orientation, owner and name.

The topological component of the data

For point locations there can be little confusion. Topological interrelations between data occurs in line, area and surface data. It has already been suggested that the relationships between *regions* can be complex, but it is equally important to consider the topological relationships between line entities. Consider a simple road network, for example. The locational component of this can be represented as a series of lines, and attributes can be recorded for each of the roads, for example the name of the road. This is shown in figure 2.5.

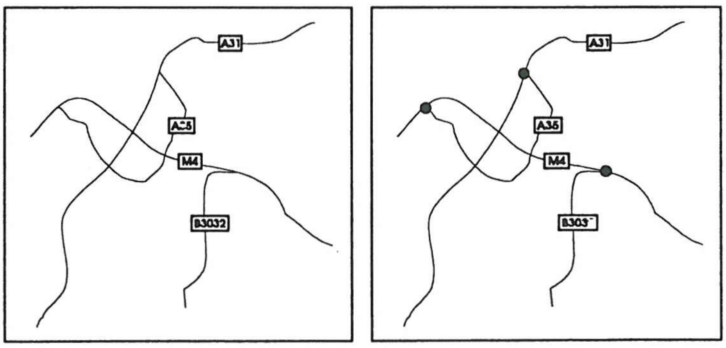


Figure 2.5. Left, roads represented as line entities with attributes. Right, introducing nodes (the black dots) into the model makes the relationship between the roads clear.

However, this in itself is insufficient to record the road network. If we are interested in the possibilities of transport within this system, we need to know whether we can drive from the A31 onto the M4 for example. To understand this, we need to record where the roads meet and where they cross each other with bridges. This sort of information about the spatial connections and relationships between objects is referred to as the *topological* component of the spatial model.

In the case of the road network it is necessary to introduce another spatial object into the model referred to as a *node* to record the way in which the roads relate to one another. In figure 2.5 (right) the nodes have been included - only at the nodes are there road junctions, and we can now observe that to get from the A31 to the M4 we must use the A35.

2.4.6. Simple data structures

Spatial data can be stored and manipulated as simple coordinate values for points, sets of points to define the vertices of lines and closed polygons for areas. This strategy is in fact used by the IDRISI GIS system and such a straightforward method for representing geographic objects has the benefit of being easy to understand. However, there are a number of problems with such a definition. Point entities do not present a major problem, and in fact, more comprehensive spatial databases generally store points broadly as IDRISI does because there are few topological problems associated with point entities (they cannot touch, contain or surround other entities). On the other hand, the storage of lines and areas in the IDRISI model is extremely restricted.

In IDRISI itself this is rarely a major problem as all analysis is carried out on raster files, and the vector data files are used solely for the importation of data, and for display. In more complex situations, however, when it is necessary to perform database and analysis operations on vector data, more sophisticated data structures are required.

Line data

Several limitations apply to the simple lines which IDRISI stores. The main restrictions are twofold: first the connections between lines are not explicitly coded within the data and second there is no directional information recorded in the data.

Recording whether lines connect, as opposed to cross or terminate without connection, is vital within network data as this allows the representation of road or trade networks. Direction within line data may be of considerable importance in rivers or pipelines where the GIS may need to provide analysis of flow direction.

Area data

The problems of recording area data as whole polygons, as in the IDRISI model, are rather more severe than those of line data. It has already been seen how the relationships between areas are not always simple. The main problems encountered are a result of sharing boundaries and failing to record the relationships between island polygons and their surroundings.

In cases where geographic areas border one another, such as soil, geological or political maps, the whole polygon model of area data must repeat each shared boundary in order to fully record the relationships. This can lead to a variety of problems, particularly if data derives from digitised maps. For example, the data file must contain at least two records for each boundary in the data. This effectively doubles the storage requirement for each coverage and dramatically increases the time required to process the data as two boundaries must be updated whenever a modification is made.

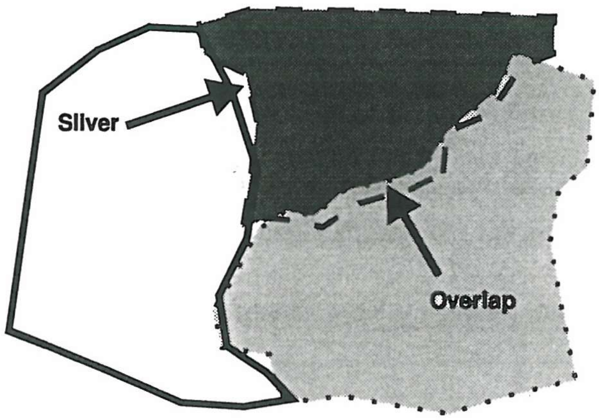


Figure 2.6. Areas represented by whole polygons showing some possible errors which result from shared boundaries

At the data acquisition stage, the shared boundary lines will not have been digitised precisely the same. Consequently these differences between the lines will then show up as *slivers* or *overlaps* in the data coverage (see figure 2.6). Additionally there is no method of checking the data for topological errors such as figure-of-8 shaped polygons, or `weird' polygons.

The whole polygon database also contains no record of the adjacency of different polygons to each other, so there is no way of searching for areas which border each other. Similarly there is no straightforward way of identifying island polygons for point-in-polygon or similar searches.

2.4.7. Spatial databases with topological content

In systems where analysis must be performed on data which is stored as vectors, the limitations of a simple data model are frequently a problem: it has been seen, for example, that use of whole polygons for area (choropleth) data is inefficient and prone to errors. As a result of such limitations, it is necessary to construct more sophisticated data models in situations where either (1) analysis and transformation is mainly carried out on vector data themes or (2) the accuracy and redundancy problems associated with whole polygon data are deemed unacceptable.

If analysis is to be carried out on data stored as vectors, the data must be capable of answering basic questions such as *are these two polygons adjacent* or *does this area contain any islands of different value?* These questions are not answerable from simple line and polygon records.

The key to solving these problems lies in the incorporation of *topological* information into the data structures. It has already been seen that topology describes the information about the relationships which are inherent within geographic data. In order to encode topological information in digital form, many GIS employ data structures based on relationships between *line* entities, *node* entities and *area* entities.

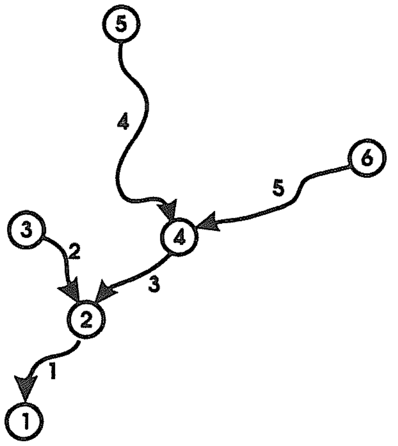
Unfortunately there are few data standards for the storage of full-topology geographic data. Because GIS is a relatively new field, each system has developed largely in isolation and has adopted slightly different ways of storing and representing data. The result is an assortment of proprietary data standards, which can cause extensive problems for the transfer of data from one application to another.

Networks

Simple lines and chains of arcs such as are used by IDRISI carry no information about connectivity or direction. Connectivity information is needed to represent, for example, the difference between roads which cross at a junction and roads which cross by a bridge, while directional information can be important in situations such as rivers, where it can be desirable to know the direction of flow. Network data can be analysed through the application of graph theory, and has found some application within study of ancient road networks (e.g. Earle 1993, Gorenflo & Bell 1993)

To store connectivity information in a digital form it is necessary to introduce the notion of a *node* into the data structure. Nodes are essentially point entities which are used to indicate the end of the chains and to store the characteristics of the ends. The records for node entities include the codes of the lines which are incident (end) at them, and sometimes *which end* of the line chain is incident at each node. If the line object being represented has a direction associated with it then this may be important.

In the example in figure 2.7, two tables have been used to represent a simple river network. The Node table records which lines are incident at each node, and also records *which end* of the line is incident at each node. This has been done by using the number of the line (i.e. `1') to indicate that line starts at this node, and a negative number (i.e. `-1') to indicate that it ends at the node. A quick check through the nodes table should then reveal that each line number has one and only one negative equivalent. In this way, the data structure represents the direction of each line, which in turn represents direction of flow within the river system.



Line	To	From
1	1	2
2	2	3
3	2	4
4	4	5
5	4	6

Node	Lines
1	-1
2	1-, 2, -3
3	2
4	3, -4, -5
5	4
6	5

Figure 2.7. A simple river network represented by a list of lines (chain of arcs) and a separate list of nodes.

Note that the `from' and `to' fields in the lines table are not strictly necessary as they can be deduced from the nodes table, however many systems include them to speed up processing. Note also that no geographic or attribute information has been included in this example. Some systems also store the angle of incidence of each line at the node, which allows the software to make checks on the relationships between lines and nodes.

Area data

Some of the limitations of simple polygons as a model for the storage of area data have been discussed above. In summary the main problems which arise are as follows (see Burrough 1986 p27 for more detail):

1. Data redundancy because the boundaries of polygons must be stored twice.
2. This duplication can lead to errors such as slivers, gaps and weird polygons
3. Islands are impossible to store meaningfully.
4. There is no way of checking the topology of the data for these errors.

Shared boundaries

This arc-node model of network data can be extended to record area data. However a wide variety of file formats have been devised my different GIS manufacturers and there is little standardisation. The following is a simplified example of arc-node structure which is loosely based on the Digital Line Graph (DLG) format of the US Geological survey.

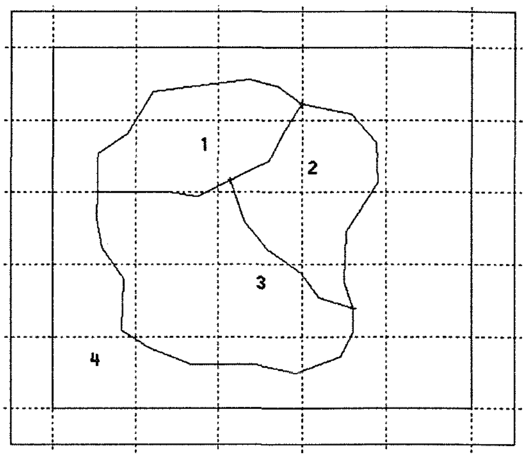


Figure 2.8. A simple map showing three areas of value 1,2 and 3.

Figure 2.8 shows a simple example of an area which might require encoding. This shows three areas of value 1, 2 and 3 which have shared boundaries, while the surrounding area has been coded as 4. In a real situation these areas may be geological substrates, soil types or political boundaries: what is required is a data model which records the boundaries only once, and preserves the relationships between the areas.

As a first step we might give the lines and nodes numbers, and record these as for network data. If the lines and nodes are numbered as in figure 2.9 (left), an initial attempt at encoding figure 2.8 may then be made as follows.

Node	x	y	No. lines	Lines
1	?	?	3	-6, -3, 1
2	?	?	3	2, 4, -1
3	?	?	3	3, -5, -2
4	?	?	3	6, -4, 5

Line	Vertices
1	x,y ...
2	x,y ...
3	x,y ...
4	x,y ...
5	x,y ...
6	x,y ...

Note that this deviates slightly from the network example in that (1) the lines do not record their start and end node (this information can be deduced from the nodes table) and (2) some geographic data has been included (the vertices of the lines). As in the network example, the attribute data has been left out for clarity. As in the network example, the lines have been given a direction by recording each line as a positive number at the start node and a negative one at the end node.

In order to record the area component of the data, however, it is necessary to introduce a third relation into the model. Although not strictly necessary, area information is frequently recorded by creating special points within the relevant areas and then attaching the area data to these. These are then called *label points* and may sometimes be generated automatically by the software.

Nodes	x	y	No. lines	Lines
1	?	?	3	-6, -3, 1
2	?	?	3	2, 4, -1
3	?	?	3	3, -5, -2
4	?	?	3	6, -4, 5

Whether the area data is attached to label points or not, there will be four area entities for the example in figure 2.8. Figure 2.9 (left) shows a the data model diagramatically: the lines and nodes have numbered been recorded as discussed in the section on networks, and numbers have also assigned to the areas. Given the direction of the lines, we can now record which area is to the left and right of each line by extending the lines table as follows (the nodes section remains the same):

Lines	Vertices	left area	right area
1	x,y ...	4	1
2	x,y ...	2	1
3	x,y ...	3	1
4	x,y ...	4	2
5	x,y ...	3	2
6	x,y ...	4	3

Next, a third table can be introduced into the model to record data about the areas. This records the lines which describe each area, coding them as positive for clockwise lines around the polygon or negative for anticlockwise (as before this can actually be deduced from the left and right areas of the lines table and so represents some data redundancy).

Areas	No. lines	Lines
4	4	0, -1, -6, -4
1	3	2, 3, 1
2	3	4, 5, -2
3	3	-5, 6, -3

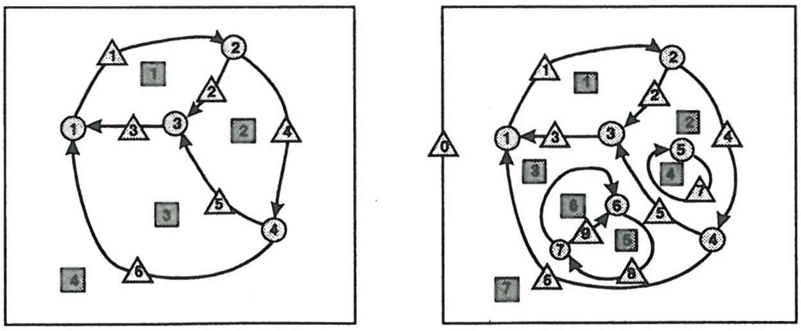


Figure 2.9. The topology of figure 2.8 (left) and figure 2.10 (right). Lines numbers are given in triangles, nodes in circles and areas in squares.

Notice that to record area 4 - the surrounding area - a new line has been introduced into the table (given the number 0). This is because without some boundary, area 4 is notionally infinite and may therefore cause problems when areas are calculated. Most GIS require that a bounding polygon (such as this line 0) is incorporated into the data.

Island and envelope polygons

So far the example has been restricted to polygons with shared boundaries. Figure 2.10 shows a more complicated topological structure, where the same three areas have been drawn but now area 2 contains an island polygon and area 3 contains a complex island. Initially the coding can proceed as before, and the data model, showing the numbers assigned to the lines, nodes and areas is shown in figure 2.9 (right).

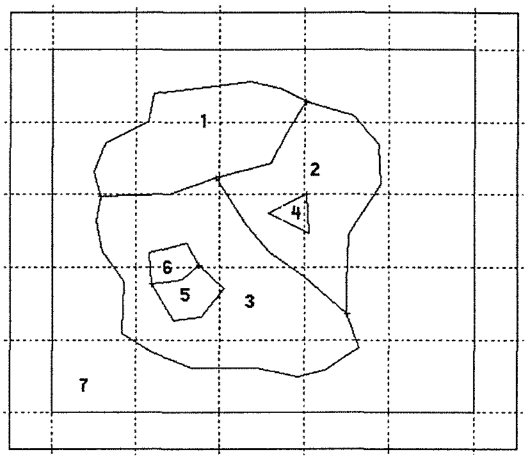


Figure 2.10. A more complex topological arrangement of polygons.

The line and node tables can be constructed as for the simple example, and including the extra lines and nodes as they are numbered in figure 2.9 (right):

Lines	Vertices	left area	right area
1	x,y ...	7	1
2	x,y ...	2	1
3	x,y ...	3	1
4	x,y ...	7	2
5	x,y ...	3	2
6	x,y ...	7	3
7	x,y ...	2	4
8	x,y ...	3	5
9	x,y ...	6	5
10	x,y ...	3	6

Nodes	x	y	No. lines	Lines
1	?	?	3	-6, -3, 1
2	?	?	3	2, 4, -1
3	?	?	3	3, -5, -2
4	?	?	3	6 , -4, 5
5	?	?	2	-7, 7
6	?	?	3	-9, 8, -10
7	?	?	3	-8, 9, 10

The areas table is now modified so that each area record includes the code for any areas it encloses (islands), and the code of any area which enclose it (envelopes). Islands are recorded in a similar way to the direction component of lines, as positive numbers, while envelopes are given as negative numbers. In this way the entire hierarchy of islands and envelope polygons can be recorded within the areas table. For figure 2.10, this would produce an areas table as follows:

Areas	No. lines	Encloses	Lines
7	4	1, 2, 3	0, -1, -6, -4
1	3	-0	2, 3, 1
2	4	-0, 4	4, 5, -2, 0, -7
3	6	-0, 5, 6	-5, 6, -3, 0, -10, -8
4	1	-2	7
5	2	-3	9, -8
6	2	-3	10, -9

Summary

The worked examples given here present a slightly simplified form of the data models which are employed by many GIS such as Arc/Info, SPANS, GRASS or GIMMS. Not all proprietary data structures incorporate all the levels of the model, while some include extra information in each of the tables, building in some data redundancy to speed processing.

Burrough (1986 pp26-32) provides a slightly different treatment of arc-node topology, an approach which involves using the angle of incidence of each line at the nodes. He then describes a method for automatically building up a fully integrated topologically linked polygon network from a set of boundary chains. Burrough's approach has the advantage that the lines may be digitised in any direction, but the method requires that area data be digitised as point entities.

To understand what makes topological data structures different from the simple models discussed previously, it is necessary to try and conceive of automated methods (without looking at the pictures) to use the data to answer questions such as: *is area 1 adjacent to area 2* or *what is the area of value 3?* These types of questions can only be answered with a computer program if the data records the full complexity of the relationships between the geographical objects.

The point of these examples is therefore to illustrate the care which is required to store spatial information in a comprehensive way. Fortunately, most of the work of creating the separate relations and then checking that the logical structure of the database is coherent can be automated within a digitising program. The user then need only remember a few basic rules while digitising a map in order to ensure the complete data structure is maintained. Different GIS digitising programs do this in different ways, so some are better and easier to operate than others. As with data standards generally there is at present no one accepted method of digitising and storing this type of data.

2.5. Digital elevation models

Digital models of terrain differ fundamentally from the choropleth data discussed previously. Geology maps, soil maps, political areas or vegetation types can be regarded as discrete variables, in that there are a finite number of categories into which the land surface is divided. An elevation map, on the other hand, represents *continuous variation* over the land surface. This is usually, but not always, a map of the *topographic height* of the terrain, expressed in metres above sea level.

The storage and manipulation of terrain data is one of the tasks for which GIS are most frequently used. Elevation data in a digital form can be used to

- Display other data in relation to terrain with 3d diagrams (see figure 2.11)
- Generate derivative maps such as slope, aspect and hill shading
- Perform cost-distance calculations and least-cost pathway analyses
- Analyse visibility and intervisibility
- Simulate and predict the effects of processes such as flooding and erosion

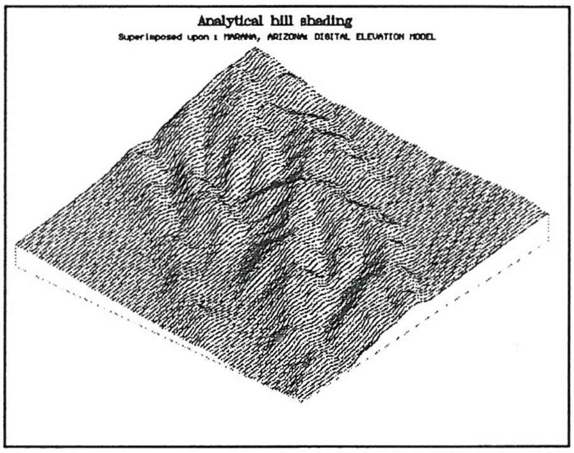


Figure 2.11. An elevation model presented as a block diagram (IDRISI)

The term *digital terrain model* is a specific term which is frequently used to specify a model of the topographic height of land form. The term *digital elevation model*, however, is more general and can refer to a model of any surface which is continuously varying over space. Obviously all digital terrain models are therefore examples of digital elevation models, but some digital elevation models are not of terrain. In fact any empirically measurable *z* variable which varies continuously over a two dimensional surface be regarded as an elevation model - pottery density, distance from water or rainfall for example.

2.5.1. Elevation data in maps

On paper maps, terrain is usually represented by contour lines which are drawn at fixed vertical intervals. Occasionally other representations are used to indicate terrain, such as coloured areas or shading but these involve a substantial generalisation of the topographic information and tend to be used for display purposes only.

The term *contours* is usually used to refer to representations of terrain land form while the more general term *isolines* can be used to refer to other situations, such as rainfall maps, in which the same representation is used (there are terms for other specific types of isoline: isolines of barometric pressure are called *isobars* for example). Although contours are the most common method of representing land form on maps, they are not the only method.

One method of representing topographic variation which is extensively used in archaeology (particularly field survey and in the production of drawings for publication) is the hachure plan. This can be used to indicate very fine variations in topography, including scarps and changes of slope which may not be visible on contour maps unless a very fine contour interval was used. Unfortunately, hachures can only represent *relative variation* of the terrain, rather than absolute height, although clever use of spot height readings combined with hachures can represent a topographic surface reasonably comprehensively. From a GIS perspective, however, the hachure plan is extremely difficult to translate into a representation which is suitable for anything except reproduction.

2.5.2. Elevation data in GIS

Point data

The simplest form of elevation model which can be stored in a GIS is a series of point entities each of which has height as an attribute. This type of data is commonly a product of field survey and can either be in the form of regularly spaced points, usually called *gridded point data* or as irregularly spaced points which are generally called *spot heights.* Spot heights have the advantage that the density of observations can vary according to the degree of variability in the terrain and this can reduce the amount of data redundancy which inevitably occurs in flat areas of a gridded survey.

Contours

Storing contours in a GIS database presents no problems other than those already discussed. Contour lines can simply be regarded as nested polygons each of which has height as a single attribute. Contour plans can therefore be stored in the same way as other types of choropleth data. In fact contours present rather fewer problems in this regard than other overlays because contour lines never cross one another and never meet, which means that there are no topological links to make between them.

Although contours are straightforward to store, they do not constitute an elevation model because a contour map does not form a continuous model of the terrain. For this reason digitised contours are rather unsuitable for generating slope or aspect maps and for investigating visibility. However, because so much topographic data is in the form of contours on paper maps, digitised contours are commonly used as a first stage in the process of generating an elevation model and as an output format for representing continuous variation in printer or plotter output. Terrain data in the form of contours is available from several mapping agencies, including the Ordnance Survey, who provide contours derived from OS maps at a variety of scales.

Altitude matrices

The most frequently used form of DEM is a regular rectangular grid of altitude measurements referred to as an altitude matrix or altitude raster. Such a matrix can then be filtered and manipulated in exactly the same way as all other raster image files, such as satellite images or scanned maps. Many mapping agencies now supply topographic data in this format, including the United States Geological Survey and the Ordnance Survey. Alternatively, altitude matrices can be interpolated from point data or from digitised contours.

Altitude matrices are the most useful form of elevation model for the generation of slope maps, aspect maps, analytical hill shading and for the analysis of visibility but there are some inherent problems with this type of model. Because the sampling size is the same for areas of high variability and very flat areas, there is usually considerable data redundancy in an altitude matrix. This can lead to very large data files, and consequently to very slow processing of data. Also, the choice of resolution is critical when using elevation data: if the cell size is too small then the file will be very large, with massive redundancy while if the cell size is too large, small scale variations which may be significant will be obscured.

Triangulated irregular networks

An alternative to the altitude matrix is provided by the Triangulated Irregular Network (TIN) model of elevation. This was devised in the late 1970s as a flexible and robust method of representing elevation which is more efficient than regular matrices but at the same time allows the calculation of derivative maps such as slope and aspect.

A TIN consists of a sheet of connected triangular faces based on a Delaunay triangulation of irregularly spaced observation points. TINs can thus easily be built up from irregular '*x, y, z*' point observations which may result from survey work. Unlike regular matrices, this allows TIN models to include higher densities of observations in areas where this is most important - areas of high variability - while using fewer observations for relatively flat areas. This allows the representation of complex and highly variable terrain without the need for a great deal of redundant data. Nodes of the TIN can also be arranged so that the model specifically follows ridges or valleys.

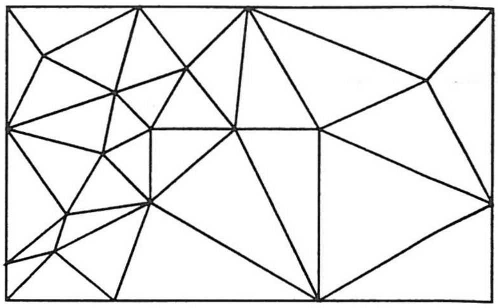


Figure 2.12. Triangulated irregular network

TIN models are stored as vectors, with the nodes storing height values, and can be regarded as a simple form of the arc-node-area data structure which has already been presented. Data structures for TINs can therefore be derived from polygon network data structures with the assumption that (1) the line entities of the model always have only two vertices, (2) all area entities are defined by exactly three lines and (3) there is no need to code islands and envelope information.

2.5.3. Creation of altitude matrices

Interpolation from contours

The most common sources of elevation data are contour maps while the most frequently used form of elevation model is the altitude matrix. Translating between these two forms of elevation model usually requires the following three steps:

1. Digitise the map contours into a vector file
2. Rasterise the vector contours
3. Interpolate the height values between the contour lines.

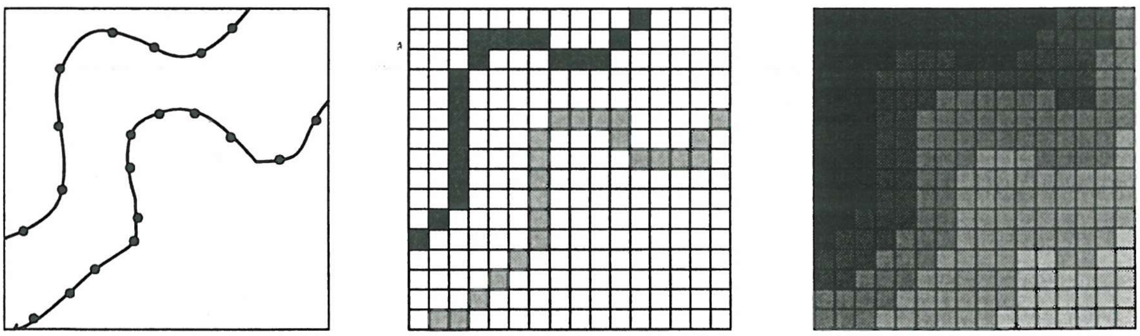


Figure 2.13. Interpolation of an altitude matrix from digital contours

Once the contours have been translated into a raster form, the process of calculating the values for the other cells in the matrix is referred to as interpolation. Different interpolation algorithms use different methods of calculating the values for the heights.

Interpolation algorithms

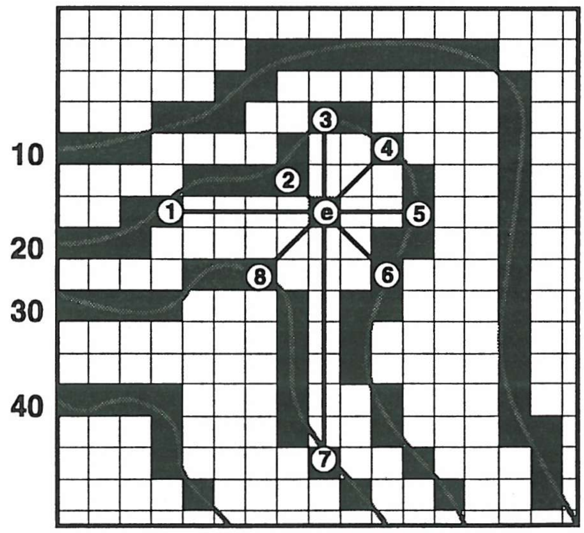
Figure 2.13 shows a typical example of the task of an interpolation algorithm. After rasterising the four contour lines (10, 20, 30 and 40m) the algorithm must ascertain values for all those cells not occurring on a line. The unknown cell is labelled ***e*** while the possible sources of information about its height are labelled 1 to 8. The elevations at each of these points can be given the notation e_1 to e_8 while the distances from the point *e* to each point can be denoted d_1 to d_8 respectively. For points 1, 3, 5 and 7 these are easily obtained but for points 2, 4, 6 and 8 the Euclidean distance must be calculated using pythagorus' theorem.

Perhaps the simplest interpolation algorithms, but least reliable, are *horizontal scan* and *vertical scan* algorithms. These scan the cells of the matrix from left to right, or top to bottom, averaging out the values between known cells. If the elevations at the points 1 to 8 are given by e_1 to e_8 respectively, and the distance between two points x and y is indicated by $d_{x,y}$, then the height of the unknown cell in figure 2.14, by a horizontal scan algorithm is given by:

$$e = \left(\left(e_5 - e_1\right)/d_{1,5}\right) + e_1 \qquad 2.1$$

The possible errors which this type of approach can produce can be deduced by looking at figure 2.14. A horizontal scan algorithm will clearly base the new value only on the 20m contour because of the loop in the line. Using equation 2.1 the slope at the unknown cell is given by:

$$e = 0 \times 5 + 20 = 20 = 20 \qquad 2.2$$



$e_1, e_2, e_3, e_4, e_5, e_6 = 20,\ e_7$ and $e_8 = 30$

$d_1{=}5\ \ d_2{=}1.41\ \ d_3{=}3\ \ d_4{=}2.83\ \ d_5{=}3\ \ d_6{=}2.83\ \ d_7{=}7\ \ d_8{=}2.83$

$s_{1,5}{=}\ 0\ \ s_{2,6}{=}\ 0\ \ s_{3,7}{=}\ 1\ \ s_{4,8}{=}\ 1.79$

Figure 2.14. Interpolating between contour lines

Far more satisfactory are *steepest ascent* algorithms which search in either four or eight directions from the unknown cell, ensuring that the new value is based only on the contours which produce the steepest slope. The algorithm then uses the steepest obtained value as for the vertical and horizontal scan algorithm. If we use the notation $s_{x,y}$ to indicate the slope between points x and y then the steepest slope of the 8 directions can be specified as:

$$e = \max\left[s_{1,5}, s_{2,6}, s_{3,7}, s_{4,8}\right] \qquad 2.3$$

Applied to the case in figure 2.14 this produces a steepest slope between points 4 and 8. This is then used to obtain the new value as follows:

$$\begin{aligned} e &= s_{4,8} \times d_4 + e_4 \\ e &= 1.79 \times 2.8 + 20 = 25 \end{aligned} \qquad 2.4$$

An alternative to this is the *weighted average* algorithm which performs a search as above but instead of using only the steepest slope, replaces the unknown cell with the average slope value, weighted for the distance to the contour lines. This can be expressed as follows:

$$e = \sum_{i=1}^{8} w_i e_i \Big/ \sum_{i=1}^{8} w_i \qquad 2.5$$

Where the weights are designed so that the elevation of the unknown cell is biased towards the closest of the contour measurements. One suitable weight value might be the reciprocal of the Euclidean distance between each known elevation and the unknown location. Using this, the weights can be calculated as follows:

$$\begin{aligned} w_1 &= 1/d_1 = 0.2 & w_5 &= 1/d_5 = 0.33 \\ w_2 &= 1/d_2 = 0.7 & w_6 &= 1/d_6 = 0.35 \\ w_3 &= 1/d_3 = 0.33 & w_7 &= 1/d_7 = 0.14 \\ w_4 &= 1/d_4 = 0.35 & w_8 &= 1/d_8 = 0.35 \end{aligned} \qquad 2.6$$

and the value of $\boldsymbol{e}$ is calculated from equation 2.5 as follows:

$$\frac{0.2x20+0.7x20+0.33x20+0.35x20+0.33x20+0.35x20+0.14x30+0.35x30}{0.2 + 0.7 + 0.33 + 0.35 + 0.33 + 0.35 + 0.14 + 0.35}$$

which simplifies to:

$$(4 + 14 + 6.6 + 7 + 6.6 + 7 + 42 + 10.35) / 2.75 = 59.75/2.75 = 21.72m$$

2.5.4. Smoothing of altitude matrices

It has been shown how different interpolation algorithms can produce different results from the same set of map contours and this is discussed in relation to landscape archaeology by Kvamme (1990c). Particularly it has been shown how simple vertical or horizontal scan algorithms produce a marked stepping of the terrain (see figure 2.15a).

In other situations the elevation model resulting from an interpolation algorithm may resemble a 'faceted' model, with the contour lines appearing as breaks of slope, separated by more or less straight slope facets. This type of result is particularly noticeable if contours are widely spaced compared with the cell resolution, and some form of steepest ascent algorithm is used. Figure 2.15b shows this effect in cross section.

In many applications, such as the calculation of slope or aspect maps, these effects may produce unacceptably distorted results and to circumvent such problems, it is sometimes necessary to smooth the elevation surface before it can be used. This is generally done by using convolution filtering.

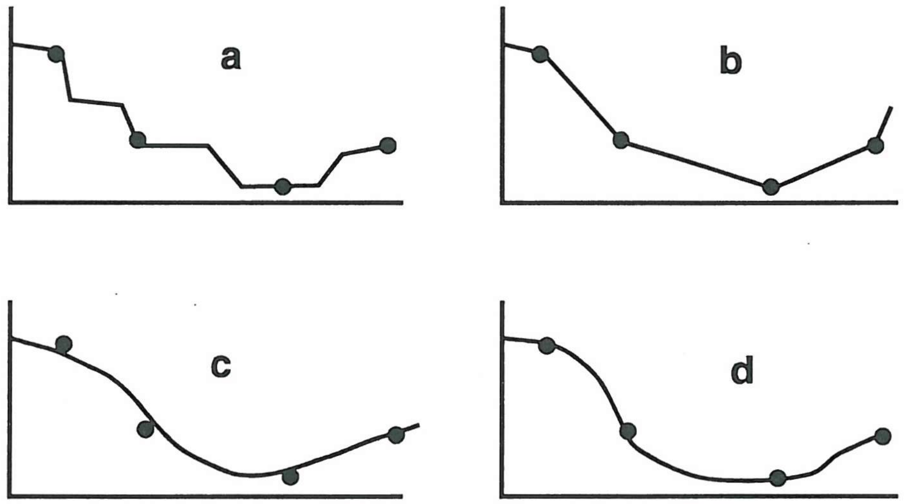


Figure 2.15. Cross sections through characteristic elevation models (a) vertical scan (b) steepest ascent (c) typical effect of filtering and (d) real elevation. Black dots represent the contour lines.

Probably the most widely used filter for this purpose is a mean filter. This can either be a simple 3x3 convolution filter, or a weighted mean filter can be used. Weighted filters generally produce far less drastic effects than unweighted ones.

The effects of filtering an elevation surface can be seen from figure 2.15. Firstly the original `noisy' model is smoothed, and the abrupt changes of slope which can be produced by interpolation algorithms are no longer apparent. However, the smoothing also affects the height of the surface at the contour lines. Figure 2.15c shows how the surface will tend to be lower than the contours on hillsides while higher than contours in valley bottoms. Compare this with figure 2.15d which illustrates the original land form from which the contours were derived.

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

1/20	1/10	1/20
1/10	2/5	1/10
1/20	1/10	1/20

Table 2.3. Coefficients for a 3x3 mean filter (left) and a typical weighted mean filter (right).

Obviously this is an undesirable consequence of filtering the terrain but the loss of absolute accuracy must be balanced against the better representation of the terrain character which results from smoothing abrupt changes of slope. In situations where derived information depends on the land form more than the absolute height of the terrain (such as in the calculation of slope or aspect maps) then smoothing may be an essential step. However, in other situations the absolute height of terrain may be of considerable value and the loss of information might be deemed unacceptable

2.5.5. Products of altitude matrices

Block diagrams

One of the most useful products of an elevation model is the block diagram, a two-dimensional representation of three dimensional terrain form of considerable use in landscape archaeology (e.g. Harris 1987). Block and net diagrams are not restricted to terrain models and can be used to visually represent the variation of any quantity over space, but are most commonly used as a representation of land form. Two main types of block diagram exist, *network* style diagrams and *line* style diagrams. The most time-consuming aspect of their creation is the removal of hidden lines. In either case, the block can be represented either with a parallel projection or by introducing a full perspective.

The usefulness of block diagrams can be increased dramatically when different data themes are overlain on the elevation data. In this way it is possible to visually interpret the association between geomorphological data and terrain: for example between geology and slope, or density of finds and terrain.

Slope and aspect

Calculation of slope maps from altitude matrices is relatively straightforward, and directly analogous to the use of *gradient operators* in image processing. Slope is usually defined by a plane tangent to the surface of the elevation model and in reality consists of two components *gradient* which is the maximum rate of change of altitude within that plane and *aspect* which is the direction at which the maximum rate of change occurs. Many authors use the terms *gradient* and *slope* interchangeably, and from this point onwards the term *slope* can be taken to be synonymous with *gradient.*

Slope can be measured either in degrees - from 0 to 90 - or in percent, in which case 0 degrees is 0 percent and 90 degrees is 100 percent. Slope is generally estimated by passing a 3x3 kernel over the map. For a continuous case, the gradient can be calculated from equation 2.7:

$$\tan G = \sqrt{(\delta Z/\delta X)^2 + (\delta Z/\delta Y)^2} \qquad 2.7$$

In the discrete case of an altitude matrix, this must be estimated from the values surrounding each cell. Either the four cells in cardinal directions can be used (rook's case) or only the diagonals (bishop's case), or all eight neighbouring cells (queen's case) can be employed in the calculation.

Using the notation in table 2.3, for cells of size *dx,* the gradient in the x-direction can be estimated from:

$$g(x) \approx e6\text{-}e4 \,/\, 2dx \qquad 2.8$$

and the gradient in the y-direction from:

$$g(y) \approx e8\text{-}e2 \,/\, 2dx \qquad 2.9$$

Because this gives two slopes at right angles to one another, the overall slope can then be computed as follows:

$$g = \sqrt{[\, g(x)^2 + g(y)^2 \,]} \qquad 2.10$$

which varies from 0 for flat areas to 1 for vertical slopes (an impossible case in an altitude matrix). This can then be scaled into a percentage or as degrees by multiplying by 90 or 100. Alternatively the diagonals can be used instead of the cardinal directions, in which case, distance from the target cell must be corrected by a factor of √2.:

$$g \approx \sqrt{[\, (e9\text{-}e1 \,/\, 2\sqrt{2}dx)^2 + (e7\text{-}e3 \,/\, 2\sqrt{2}dx)^2 \,]} \qquad 2.11$$

The rook's case method given by equation 2.10 is particularly sensitive to small errors in the terrain model, because it relies on only four of the values in the kernel. A more robust estimate of the slope in the x-direction or y-direction can be obtained by using the following:

$$g(x) \approx [(e3 + 2e6 + e9) - (e1 + 2e4 + e7)] \,/\, 8dx$$

$$g(y) \approx [(e7 + 2e8 + e9) - (e1 + 2e2 + e3)] \,/\, 8dx \qquad 2.12$$

and the values from these estimators can then be used in equation 2.10. Note that the application of equations 2.12 followed by 2.10 describes the process for implementing the *Sobel operators* which are widely used in image processing for edge detection.

	e2	
e4	e5	e6
	e8	

e1		e3
	e5	
e7		e9

e1	e2	e3
e4	e5	e6
e7	e8	e9

Table 2.4. Coding of elevation values in the original image showing rooks case (left) bishops case (centre) and queens case (right) values.

Aspect is usually specified in degrees, with 0=360=North. Aspect can be obtained from equation 2.13, where $-\pi < A < \pi$ radians or $0 < A < 360$ degrees.

$$\tan A = \frac{-\delta Z/\delta Y}{\delta Z/\delta X} \qquad 2.13$$

Using one of the estimators of the x and y gradient, the aspect ***a*** can be estimated from:

$$a = \arctan(\, -g(x)\, /\, g(y)\,) \qquad 2.14$$

Aspect maps are usually reclassified from degrees (or radians) into the eight main compass directions (N, NE, E, SE, S, SW, W and NW) with one extra classification for no aspect - where the slope is zero.

Figure 2.16. Analytical hill shading map produced by IDRISI

Analytical hill shading

One of the most attractive products of elevation models is the analytical shading map. These are useful for the visual interpretation of land form, showing the terrain in an intuitive way, and allowing subtle features of the terrain to be picked out by eye. Hill shading maps record the theoretical intensity of reflected light which would be seen by an observer who was situated directly above the terrain. They can be derived from slope and aspect maps, requiring only that the location of the light source be known.

Most shaded relief maps make the assumption that the inherent reflectivity of the elevation surface is constant: in other words that the terrain is not more noticeably shiny in particular places, and therefore base the reflectance value only on the relationship of the terrain slope and aspect to the light source. Results can be extremely convincing, and are especially effective when 'draped' over a block diagram of terrain.

2.6. Analytical capabilities of GIS

The analytical capability of GIS stems from the ability to process, manipulate and combine geographic data overlays in spatially meaningful ways. It is the transformation of geographic data which is the most powerful aspect of GIS, and which allows the realisation of existing but latent information, and the generation of entirely new information. Several attempts to define the fundamental operations of a GIS as a basis for methods of map algebra have been made (e.g. Burrough 1986, Berry 1987). One useful and widely accepted view is characterised by Martin (1990) who divides the basic operations of GIS into four groups:

- **Reclassification** operations result in alterations to the attribute component of geographic objects. This is the simplest class of spatial operation, frequently used to simplify large numbers of classes into more manageable numbers of classes by the application of specified rules.
- **Overlay** operations involve the combination of two or more maps in a systematic way. At its simplest, overlay transformation might involve the addition of one map to another. More complex overlay transformations involve the use of boolean logic.
- **Distance and connectivity** operations include simple measures of distance between points or areas of polygons and more complex operations involving weighted distances such as cost-surfaces and network analyses (see below).
- **Neighbourhood** transformations involve the ascription of values to locations based on the characteristics of nearby geographic entities. Neighbourhood transformations not only include convolution filters (such as mean or mode) which are typically used for smoothing surface data; but also more specialised operations such as that producing the `ridge-drainage index' devised by Kvamme (Kvamme 1992).

GIS operations frequently return a different geographic data type to that used as an input, thus a distance operation might take as its input a map of roads (in the form of line data) and produce a classified map of areas which are specified distances from the roads (a choropleth map). Transformations are also rarely used in isolation, and combinations of transformations are generally used to generate new spatial data in a process which has become known as `cartographic modelling'. Here the main capabilities of GIS are grouped into three main categories: *generalisation* operations, which produce summaries and extract existing information; *transformation* operations which generate new information by acting on single data themes; and *interaction* operations which generate new information by acting on multiple themes. Within this classification, a small subset of the most widely applied GIS operators will be described.

2.6.1. Generalisation

Extraction of implicit information

Once data within a GIS is fully georeferenced and is topologically complete, the database contains a considerable amount of latent information: information regarding the relationships between objects. Extraction of this information is one of the primary functions of the GIS, and a variety of the operations performed by GIS can be regarded in this way.

The simplest example of this type of operation is the addition of attribute information to point data themes, based on data in choropleth or surface themes. For example, in case where a GIS contains soil themes (choropleth), elevation themes (surface), and also site locations in the form of point data it is a straightforward operation to identify the soil class and elevation of each of the sites (this is achieved with a simple lookup function in a raster GIS and with a point-in-polygon function in a vector system). Such operations do not generate new information because the soil class and elevation of each point location is already present in the system, but they do *realise* such latent information. This is frequently referred to as *adding value* to geographic data.

Area summaries

Area operations are generally performed on choropleth themes, such as geological or soil series maps. The input is in the form of a map, the output is therefore in the form of a table.

$$\text{Map of } n \text{ classes} \xrightarrow{\text{Area}} \text{Table of } n \times 2 \qquad 2.15$$

For example, data concerning 20 settlement sites in a region which may be divided into 4 categories of soil. The number of settlement sites occurring on the following soil types are as follows:

Soil type	Settlements
1	2
2	13
3	4
4	1

Table 2.5. Number of settlements per soil class

And it is possible to form the initial hypothesis that the settlements were preferentially distributed on soils of type 2. An initial step in testing this hypothesis would then be to generate an area table from the soil map - clearly if there is much more of soil type 2 in the area, then we would be unable to support the hypothesis. Running an area operation on the soil map described above might produce a table as follows:

Soil type	Area (hectares)
1	2
2	153
3	0.5
4	44.5

Table 2.6. Area of the soil types in table 2.5

Immediately it is apparent that there is far more soil of type 2 in the region. However, it is also clear that soil type 3 is barely present at all and that soil type 4 is extensively represented. With this information, the initial hypothesis can be modified to the extent that soil type 4 seems to have fewer settlement sites than we may expect, while soil type 3 may have rather more.

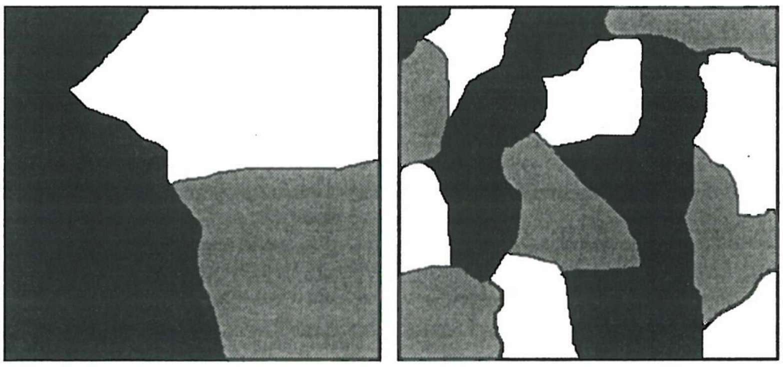


Figure 2.17. Both maps have the same number of classes, and produce the same area table. The map on the left, however, has a different perimeter table to than on the right.

Perimeter summaries

Another generalisation which is sometimes of value is the generation of tables of *perimeter*. This operation is of exactly the same form as the area operation:

$$\text{Map of } n \text{ classes} \xrightarrow{\text{Perimiter}} \text{Table of } n \times 2 \qquad 2.16$$

but the values in the table represent a linear measurement (metres or km) of the boundary of each class in the original map. This can be a very useful measure of the variability of a map: two maps with the same number of classes may produce considerably different perimeter tables if one is far more homogenous than the other. Figure 2.17 shows an example of this.

Area histograms

Summary information need not be presented in the form of a table. If a map contains ordinal or higher level data, then the areas of each class can be presented in the form of an area histogram. This can produce an extremely valuable source of information when investigating spatial relationships with parametric statistics.

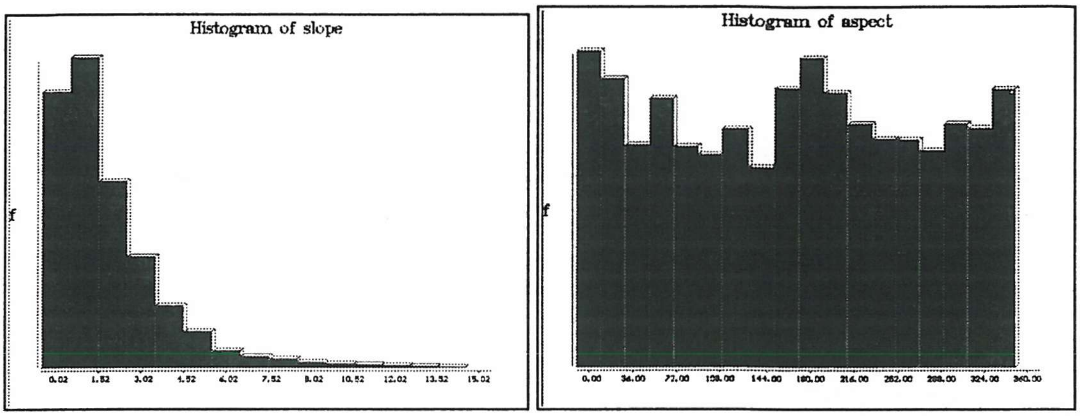


Figure 2.18. Image histograms for a slope map (left) and an aspect map (right)

The two image histograms in figure 2.18 show the areas of a slope map (reclassified into 15 classes) and an aspect map. The different character of each data type can be clearly seen. The slope map produces a histogram which is typical of many spatial data themes, with a large number of low values and an apparently exponential decline in area towards the higher values. The aspect histogram shows that the area has a high proportion of north and south facing areas (the peaks in the middle and at the edges) but a lower area of east and west facing areas.

In neither case could the variable be considered to be normally distributed which is one of the assumptions of many parametric statistics. In the case of the slope image, it may be possible to obtain a normally distributed map by performing a *transformation* on the slope image (see below).

Crosstabulation

Given two maps, it may be important to know to what extent the classes of one map co-vary with the classes of a second map. This operation is referred to as *crosstabulation* of two maps, and takes the general form:

$$\begin{matrix}\text{Map of } n \text{ classes} \\ \text{Map of } m \text{ classes}\end{matrix} \xrightarrow{\text{Crosstabulation}} \text{Table of } n \times m \qquad 2.17$$

For example, we may have a map of soil classes, and a map of geological types. If there are five soil types and four geological types we may crosstabulate them to obtain a table as follows. This indicates the area of each soil which occurs in each geological type. We can see immediately from table 2.7 that Rendsina soils tend to occur in areas of limestone, while

Brownearths seem to occur in areas of clay or shale. We can also note that Gley soils seem to be associated with Alluvial deposits.

	Rendsina	Grey rendsina	Brownearth	Gley	Podsol
Limestone	150	130	15	1	4
Clay	10	12	205	78	28
Shale	0	0	130	28	54
Alluvium	0.5	0	55	104	1

Table 2.7. Result of crosstabulating soil classes with geological substrate classes.

2.6.2. Transformation of single themes

Transformation operations are referred to by Burrough (1986) as *point* operations, because each calculation in each transformation involves only one cell in each theme. This is a similar definition to the definition of a *point processing* operation in image processing, and the principle is exactly the same although in a GIS the concept is not restricted to a raster storage system.

Reclassification

Reclassification of a map into a simpler format is the most basic of all map generalisations. It can be done in a variety of ways either from a legend, or automatically by specifying the class width or number of classes. In general terms, reclassification can be expressed as:

$$\text{Map of } n \text{ classes} \xrightarrow{\text{classification}} \text{Map of } n - m \text{ classes} \qquad 2.18$$

and the ratio of *(n-m)* to *n* is a measure of the degree of generalisation. Reclassification is a useful first step in modelling, or examining spatial relationships. It can also be used to convert between abstract measures, and measures which are give some specific meaning. For example, if we have a map of soil type expressed in terms of the soil series name, with four classes, we may wish to reclassify this into a new map which represents the *quality* of the soil for a particular purpose - say for crop production as good=3, medium=2, poor=1. The classifier would then be in the form of a *table* which related the soil type code, to a soil quality code as in table 2.8.:

Soil type	Soil quality
1	2
2	1
3	3
4	1

Table 2.8. Simple reclassification from 4 categories to 3.

This table can then be used to reclassify the soil type map into a new soil quality map which may be of far more value. Obviously the new map cannot have more classes than the original map, and in this case there is a generalisation from 4 to 3 classes involved in the operation.

Scalar transformations

The transformation of any data theme with the application of a single arithmetic operator with of constant value is referred to as a *scalar* operation, and takes the general form:

$$\begin{array}{c}\text{Map of } n \text{ classes} \\ \text{Constant}\end{array} \xrightarrow{\text{operator}} \text{Map of } m \text{ classes} \qquad 2.19.$$

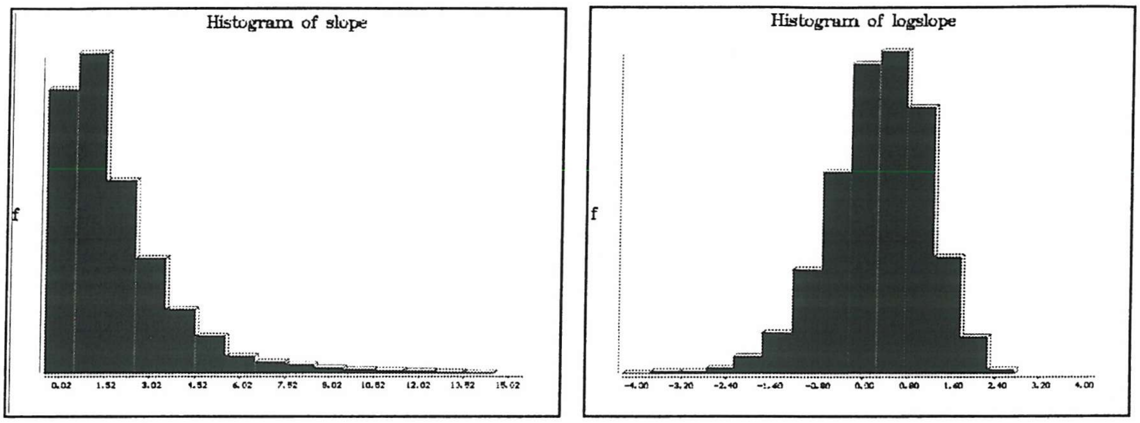


Figure 2.19. Slope image (left) is transformed by taking the natural log of each cell. The result is a far closer approximation of a normal distribution.

The simplest of the operators are *addition*, *subtraction*, *multiplication*, of *division* of a constant number applied to all values in a map. A simple example of the use of an arithmetic operator may be the addition of a number to an entire map. For example, if we have a map of the number of houses per kilometre over a wide area, and we know that the average house is home to 3 people, it is possible to multiply the house density map by 3 to obtain an estimate of the population.

More sophisticated transformations may be used to `normalise' data which is not normally distributed. Thus if the slope map from figure 2.18 were transformed with a *log* operator, the distribution of the new map (of log slope) would appear closer to a normal distribution. The new transformed map may then be appropriate for use in parametric statistics. Figure 2.19 shows the result.

Other operations which come into this category include square, square root, trigonometric functions (sine, cosine, tangent etc.) and exponent operations.

above in relation to products which can be derived from elevation models: thus slope and aspect interpolators are classed as neighbourhood operators, as are the functions for interpolating elevation matrices from contours. Many of the these are specific examples of what, in image processing, would be termed *convolution filters*. Other simple examples of such filters include mean, mode and variance operators which replace the value at a particular location with some mathematical convolution of the values of a given number of neighbouring locations.

Buffers, corridors and proximity surfaces

One of the most useful transformation abilities of GIS is the generation of distance maps, either in the form of distance surfaces or as distance buffers. *Distance surfaces* are maps of continuous variation in which the magnitude at any point of the map is the proximity to a particular geographic entity.

Distance buffers and *corridors* are simply discrete versions of distance surfaces in the form of choropleth maps where the classes represent a range of distances from the entity. The term *buffers* usually means a map with several distance bands, generated from either points or lines, while a *corridor* usually refers to a single distance buffer from a line (see figure 2.20).

Distance operators are an example of a data transformation whose result is of a higher spatial dimension than the original data. For example, a distance operator which produces distance buffers from a point distribution begins with data which is *zero dimensional* and generates a choropleth map which is *two dimensional*. Similarly an operator which generates a distance surface from line data begins with *two dimensional* lines and generates surface data (which can be regarded as two-and-a-half dimensions). In general:

$$\text{0 or 1 dimension} \xrightarrow{\text{distance}} \text{2 or 2.5 dimension} \qquad 2.20$$

for example:

$$\begin{aligned} &\text{Point / line data} \xrightarrow{\text{distance buffers}} \text{Area data} \\ &\text{Point / line data} \xrightarrow{\text{distance surface}} \text{Surface} \end{aligned} \qquad 2.21$$

It should be clear that the generation of buffers and corridors is essentially a simplification of the generation of a distance surface. Many GIS allow the generation of buffers or corridors in one operation, some do not. In the latter, it is therefore necessary to break the buffering operation into two parts, first generate a distance surface, then apply a reclassification to the distance surface to generate the buffers.

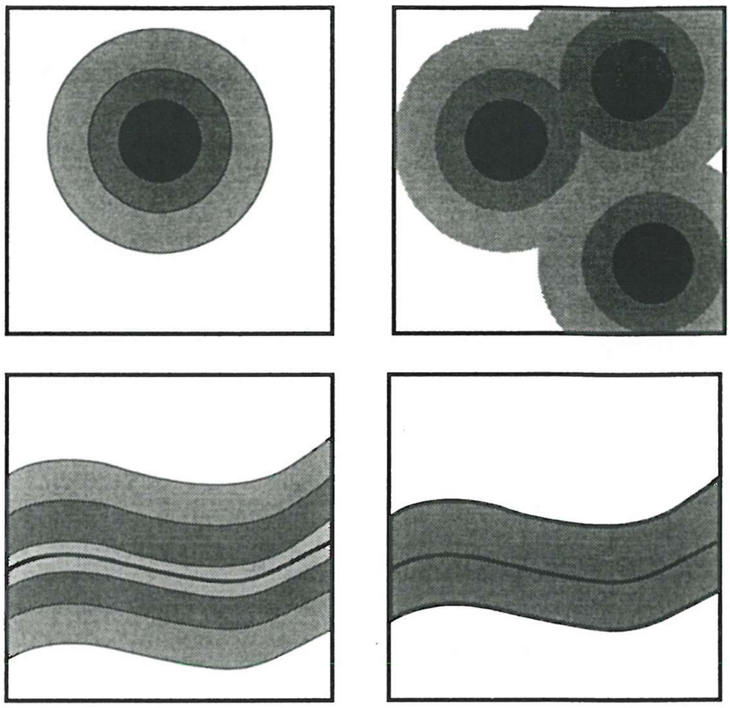


Figure 2.20. Top left: distance buffers from a single point. Top right: distance buffers from several points, Bottom left: Distance buffers from a line, Bottom right: One distance buffer or corridor from a line.

2.6.3. Interaction between themes

Arithmetic between layers

Most GIS will provide facilities to add, subtract, multiply or divide one map from another map. In these operations each location in the new map is calculated from the same location in the source maps - these are still *point processing* operations in image processing terms. A general form of these operations is as follows:

$$\text{Map of } n \text{ classes} \xrightarrow{\text{operator}} \text{Map of } m \text{ classes} \Rightarrow \text{Map of } p \text{ classes} \qquad 2.22$$

For example, if we have a map showing the density of female population and a second map showing the density of male population and want to know the population density of both sexes we would proceed by *adding* the two maps together in this way.

All the other operators which can be used with a constant can also be used with another map, thus maps can be, for example, subtracted from one another, divided, multiplied, one can be raised to the power of another.

Boolean operations

As well as these arithmetic operations, it is possible to define a set of logical operations between maps. These are expressed with the boolean operators AND, OR, NOT, and XOR. The effects of boolean operators can be summarised in the Venn diagram shown in figure 2.21.

Boolean operators return either TRUE (coded as 1) or FALSE (codes as a 0) and so the result map will always be of two classes. In general:

$$\text{Map of } n \text{ classes} \xrightarrow{\text{boolean}} \text{Map of } m \text{ classes} \Rightarrow \text{Map of 2 classes} \qquad 2.23$$

Boolean operations are of particular value in ***unique conditions modelling***. For example, if we have a map of soil depth (in 3 classes where 1= thin and 3 = thick) and a map of slope (in 5 classes, where 0 = flat to 5 = very steep) it is possible to identify areas which will be particularly prone to soil erosion using the following boolean map operations:

[(SLOPE=4) OR (SLOPE=5)] AND (SOIL=1) 2.24

Note that the additional brackets are necessary in this expression because boolean expressions are not *commutative*. For example:

[A OR B] AND C ≠ A OR [B AND C] 2.25

Unlike arithmetic operations, which are commutative. So that:

[A + B] - C= A + [B - C] 2.26

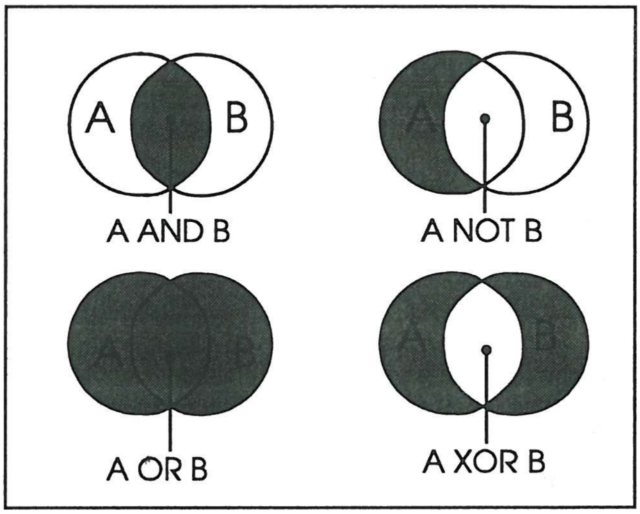


Figure 2.21. Venn diagram showing the effects of the four basic boolean operators

Mask and cover overlay operations

Two of the most useful operations which can be performed between two data layers are sometimes considered as entirely separate operations (although they are specific cases of boolean operations). If A and B are maps then:

- **Cover** A with B- will produce A unless B is 0, in which case it will produce B. In boolean terms this is:

  IF (A =0) THEN B ELSE A 2.27

- **Mask** A with B - where B is usually a binary image, is a particular case of a cover operation. It will produce A where B is non-zero and zero otherwise. This can be expressed in boolean terms as:

IF (B=0) THEN 0 ELSE A 2.28

Line of sight operations

The calculation of a line of sight map or `viewshed' for a point location, given a digital elevation model is a relatively trivial computing problem and is available within the current functionality of many GIS. In a raster system the calculation requires that, for each cell in the raster, a straight line be interpolated between the source point and each other cell within the elevation model. The heights of all the cells which occur on the straight line between the source and target cells can then be obtained in order to ascertain whether or not the cell exceeds the height of the three dimensional line at that point. Figure 2.22 shows diagramatically how this operates: the example top right shows the case for the existence of a line of sight, while the example bottom right shows the case for no line of sight. It is usual, in most cases, to add an additional value to the source cell to account for the height of the human eye above the surface.

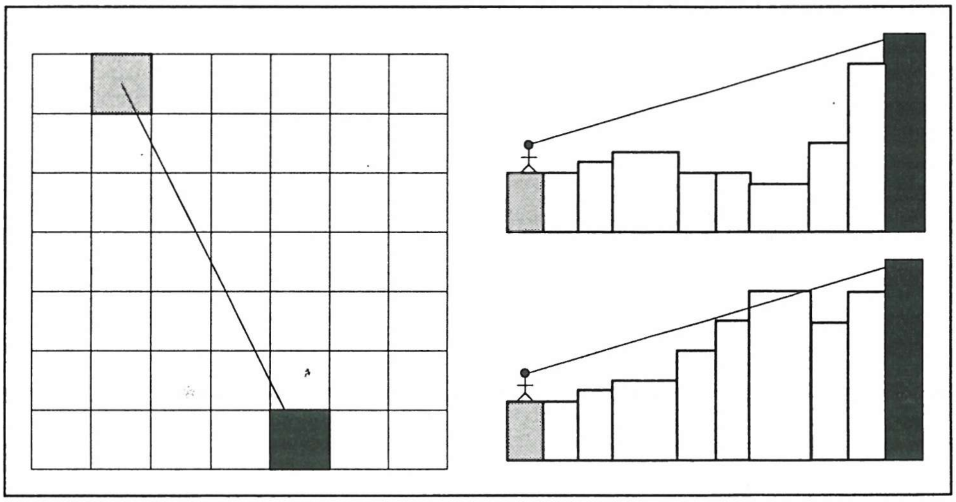


Figure 2.22. Schematic illustration of a line-of-sight calculation. A line is interpolated between two cells of the elevation model (left). If the heights of the intervening cells do not cross this line (top right) then there is a line of sight, if the height of any cell does exceed the height of the line (bottom right) then there is no line of sight.

The result of each of these discrete calculations is either a positive (figure 2.22, top) or negative (figure 2.22, bottom) result, conventionally coded as a 1 for a visible cell or a 0 for a cell which is not visible. When performed for the entire raster, the result is a binary image, those areas of the landscape which have a direct line of sight from the target cell coded as a 1 and those with no line of sight with 0. Such images are commonly referred to as *visibility* *maps* or *viewshed maps* and have already found some application within archaeology (see chapter 6).

Cost surface calculations

The calculation of cost surfaces is an extension of the distance operator discussed in section 2.6.2. Instead of generating operating on a single data input to produce the product, cost surface calculations operate on a data theme input one or more modifier themes. The product of cost surface operators is a map which represents the modified distance from any location to the input location. Inputs may therefore be either points or lines, while the output is usually a surface. Modifier data can take two main forms, depending on whether an isotropic or anisotropic calculation is undertaken (Eastman 1993).

In the simpler, anisotropic case the cost operator assumes that there is a *friction* cost associated with movement across a surface. If the friction is assumed not to be dependent on the direction of movement across the surface, then it can be modelled by a single modifier theme (usually some derivation of slope or terrain roughness) expressed in fractions of a base cost. The cost surface operator will them calculate the distance of each location in the landscape from the target location, cumulatively weighted by the friction value, to produce a cost surface result.

If, however, the friction cannot be considered anisotropic then it is normally necessary to introduce more than one modifier theme to implement the friction. Algorithms which implement this type of operator are referred to as isotropic cost surface algorithms. Although several solutions are possible, not all of which require the use of more than one modifier theme (S. Stead *pers. comm.*), one solution is to store the magnitude of the friction in one theme, and the direction in which the friction has greatest effect in another. If the friction is to be based on slope, then the magnitude of the friction would be derived from the steepness of the slope, while the direction of greatest impact of the slope may be derived from the aspect: the slope should generate greatest friction when traversing straight up, and should produce minimum friction when traversing straight down. The isotropic algorithm may then model the friction by obtaining the angular difference between the direction modifier an the direction of travel (a) and using some variant on equation 2.29.

$$\text{Cost} = \text{Cost}_{\text{base}} \times \text{F}_{\text{magnitude}}^{\cos(\alpha)} \qquad \textit{2.29}$$

This ensures that the friction used for the calculation is the value of $F_{magnitude}$ at a = zero degrees but is $F_{magnitude}$ to the power of -1 at a = 180°. Further modification can be introduced into the process through the use of a constant which is applied as a power function of the cosine (e.g. Eastman 1993).

2.6.4. Predictive modelling

One of the most common applications of GIS has been the construction of what have come to be known as *predictive models* of archaeological resources. Prediction is a central component of spatial analysis, and the goal of all predictive modelling is to generate a model which has predictive implications for future observations.

Predictive models exploit the ability to use GIS overlays as independent variables to predict a desired dependent variable. In the case of archaeology the dependent variable is some measure of the location of archaeological sites, and the technique has become known as `archaeological predictive modelling'. The aim is to construct an hypothesis about the location of archaeological remains which can be used to predict the locations of sites which have not yet been observed. Kvamme (1990a) puts it as follows:

> *`... a predictive archaeological locational model may simply be regarded as an assignment procedure, or rule, that correctly indicates an archaeological event outcome at a land parcel location with greater probability than that attributable to chance.'*

The main impetus for the development of predictive methods was therefore practical and economic: it was believed that such methods could reduce the effort required to create a justified land-management strategy for a large region by reducing the area which needed to be surveyed, and hence reducing the expenditure. Again, Kvamme (1990a) describes the motivation for the development of the techniques well:

> *`If powerful resource location models can be developed then cultural resource managers could use them as planning tools to guide development and land disturbing activities around predicted archaeologically sensitive regions. This planning potential of predictive models can itself represent significant cost savings for governmental agencies.'*

Two definitive descriptions of predictive modelling in archaeology have been recently presented. Warren (1990a) presents a thorough introduction to the principles and describes in detail one approach based on logistic regression, while Kvamme (1990a) provides a similarly thorough overview of the principles underlying spatial modelling in archaeology but relies less on any specific technique. This account derives primarily from these two introductions.

Although there are a number of different methods for generating predictive models, all models of site location have certain characteristics in common:

1. The unit of study is the land parcel, as opposed to the site. These land parcels are generally of the same size and in a GIS context almost always refer to square cells of landscape of known dimensions. It is the characteristics of these land parcels which are measured to generate the independent variables (i.e. soil type, topographic height, distance from water) and it use the characteristics of these which are the dependent variable (usually presence/absence of specified archaeological site).

2. The models all attempt to predict archaeological events in cells of unknown value. These events may be the simple presence/absence of an archaeological site within the land parcel or they may be more subtle as in the density of archaeological finds within the cell, or the type of site present.
3. These events form a mutually exclusive and exhaustive classification of the land parcels, in other words any land parcel can be classified into only one of the archaeological events. This requires that the event of `no archaeological outcome' must always be defined. (Kvamme 1990a).

There are also two distinct approaches to predictive modelling: deductive and inductive. In a deductive model, the theory comes first and this theory states that a set of rules may be applied to a set of input unknowns to determine their status. In an inductive model, a series of measurements are taken and then a method is used to identify the pattern within these. It is this pattern which then decides the rules for the prediction. It is theoretically possible to generate a model in a purely inductive manner or in a purely deductive manner, but in practice it is far easier to generate an entirely deductive model than a purely inductive one. This is because there are an infinite number of possible independent variables to consider in a purely inductive method and the selection of a subset of these for inclusion in the practical application is necessarily a theoretical deduction. The commonest approach is therefore to deductively identify a series of independent variables for the model which theory suggests may influence the status of land parcels, and from there to use an inductive methodology to generate the model itself.

Although statistical methods have been commonly used in predictive modelling, for reasons discussed below, it is not actually necessary to rely on statistical regression techniques to derive a predictive model. It is, for example, possible to utilise the map-algebra notions of GIS in order to identify locations within a study area which conform to certain theoretical specifications (see below). However, many applications of predictive modelling have used statistical procedures in order to generate a result which represents the probability of a given cell containing an archaeological site.

Most popular among these tests have been variants of regression analysis, which in its simplest form (bivariate linear regression) can be used to define the relationship between two variables for a given sample of cases. Multiple regression techniques extend this principle to allow one dependent variable to be predicted from two or more dependent variables (see e.g. Shennan 1987 for archaeological applications). Also used is Discriminant Function Analysis which is a multivariate procedure which utilises observations made on members of two distinct groups to generate a function which minimises the number of cases which are classified into the wrong group. This function can then be used to predict the group membership of new cases for whom the original variables are known. Perhaps the most suitable statistical method for predictive modelling is logistic regression analysis, as Warren (1990a) puts it:

> *'... logistic regression analysis is clearly the method of choice for empirical predictive modelling of archaeological site location.'(Warren 1990a p96)*

Logistic regression has been used in a number of archaeological applications such as the study by Kvamme (1988) discussed below, Warren's investigation of the western Shawnee National Forest, Ilinois (Warren 1990b) in which the author demonstrated the use of logistic regression to generate maps of expected site density for the region and by Carmichael to generate site sensitivity maps for his study area in north-central Montana (Carmichael 1990).

Deductive and inductive approaches

It should be clear from this that predictive modelling is not a single method, but essentially describes a wide range of approaches to one specific problem. Predictive models can be based on two different sources of information:

1. Theories about the spatial distribution of archaeological material and
2. Empirical observations of the archaeological record

Conceptually, it should be possible to create a model based either on pure induction, or on pure deduction: for example we may theorise a relationship between site location and environmental resources and then implement this relationship as a predictor in a GIS without any recourse to the actual locations of the sites. Alternatively it may be possible to derive a mechanical method for using the environmental characteristics of a group of known sites to deriving such a relationship without recourse to a hypothesis about the function or meaning of the sites.

It should be recognised, however, that such a distinction between data and theory is not universally recognised, and most archaeologists accept that the two are not independent - data is collected within a theoretical context, and so may be regarded as theory-laden, while theories are generally based to some extent on empirical observations.

However, although it is practically not possible to devise a predictive modelling method which is based entirely on either of these tactics, it is useful to maintain the distinction between the two approaches at a methodological level.

Inputs and outputs

Modelling is generally the production of one or more outputs from one or more inputs through the use of a specified rule. In the case of archaeological predictive modelling, the outputs are always archaeological in nature while the inputs are generally not.

It should be noted (as stated above) that `unit of currency' of archaeological predictive models is the landscape location, not the archaeological site - in a raster system this means that the object of the model is the raster cell of specified resolution. The model then works on this unit, to assess the output from the inputs.

Inputs to models

Input to predictive models is usually in the form of non-archaeological information about the landscape locations. Various inputs have been used, including:

1. Physical environment characteristics: which are the most easily obtained inputs in GIS contexts, and therefore most widely used inputs to models. Which variables are selected should be determined by knowledge that they are related to the location of the sites in question, which may be determined either by the experience of a particular archaeologist or more formally by a statistical analysis of existing sites. A huge variety of characteristics may be used as the predictors, and those which have been employed include elevation, land form derivatives such as slope, aspect, indices of ridge/drainage, local relief, geological and soil data, nominal classifications of land class (such as `canyon', `plain', `rim' etc.) and distances to resources such as water (rivers) or raw material sources (flint). In cases where the vegetation is considered not to have changed significantly, vegetation classes may also be used.

2. Spatial parameters: In cases where sites are known to have spatial relationships - in other words they tend to cluster or to be noticeably dispersed - this information can then be used as a predictor in the model. For example, if it is assumed that a particular type of site is likely to form a clustered distribution, then the likelihood of an undiscovered site occurring in any given location will decrease with distance from known sites.

3. Cultural features: In some cases the cultural features of a landscape can be of use in predicting the location of sites, for example it may be that sites occur in close proximity to road networks or central places.

Outputs from models

The output from the model may be very simple or fairly complex, and may take a variety of specific forms, such as:

1. *Presence/absence of site* is the most common type of output required from a predictive model. These types of models usually represent one specific site type with general characteristics which can be reliably claimed, but these site classes may be defined by function or by chronology. Output in this case is binary.

2. *Site class* can be output from more sophisticated models. In this case a decision is made about each location which places it in one of several site classes or a special class of `no site'. Output from these models is categorical.

3. *Densities* of sites/artefacts. In situations where the resolution is large enough so that several sites or artefacts tend to occur in each location then density can be considered as the output variable of the model. In this case the output may be of ordinal or real number data.

4. *Site significance.* Occasionally resource management applications have attempted to use prediction to classify landscapes according to the perceived importance of the archaeological remains which may be present. In this case the output may then be of any type, depending on how `significance' is defined. The most notable example of this approach (James *et al* 1983) attempted to predict a complex interval scale dependent measure of site significance from environmental measurements. This application has been heavily criticised on the grounds that the dependent variable is a composite, and therefore is a misuse of the regression method.

5. *Probability* models. One particularly useful type of model produces as output an indication of the probability of a site occurring at each location in the landscape. In these instances the output is usually a real number, either ranked between 1 and 0 or expressed as a percentage.

Approaches based on site location only: spatial interpolators

It is possible to build a prediction of the location of undiscovered sites purely from the locations of existing known sites, a procedure which is clearly essentially inductive in nature. Procedures which estimate the magnitude of a spatial variable at an unknown location from the values at known locations are called *spatial interpolators*.

Use of spatial interpolation for archaeological predictive modelling requires that two assumptions be made. The first is that the known sample of sites constitutes a reasonably random sample of the undiscovered site population (which can often be justified) but secondly assumes that the sites are related in a rational way solely in terms of the space in which they occur. This is undoubtedly often the case for a specific group of sites in that they may exhibit clustering, or more complex spatial patterning such as the hexagonal distribution predicted by central place theories. However, such procedures are always a simplistic and unreliable predictors of sites because the prediction ignores all other types of information.

Trend surface analysis

One of the simplest techniques for predicting site location is trend surface analysis which is a polynomial regression technique. Regression procedures are not generally suited to the modelling nominal level variables such as site classes (although logistic regression does allow the use of nominal data) but they can be used to model site densities. However, trend surface analysis is most meaningfully applied to modelling the spatial distribution of a measurable characteristic of a group of sites; in other words to the modelling of continuous dependent variables. Examples of this include the analysis of the Lowland Maya sites by Kvamme (1990d), where the terminal dates of Maya settlement sites were shown to exhibit a degree of spatial autocorrelation allowing a trend surface to be used to describe the trend, and to predict the terminal dates of undiscovered sites. Similar studies of the distribution of length/width indices of Bagterp spearheads in northern Europe, and of percentages of Oxford pottery in southern Britain can be found in Hodder and Orton (1976).

Convolution interpolators

It is straightforward to interpolate unknown values from known values using any of a variety of image processing convolution techniques. In cases where the majority of the spatial area is of known value, and only a few areas are unobserved then a simple mode filter may be appropriate. This simply estimates the value at the unsurveyed location as being the most commonly occurring value with the filter window. Alternative methods which may be of use in specific circumstances include mean or weighted average filters.

This approach to spatial interpolation has one serious disadvantage, in that each filter may produce widely different results depending on the choice of weights for the filters, and on the size of the window.

Kriging

More recently, alternatives to the use of polynomial regression type trend surface analyses have been applied to archaeology. Kriging is a technique which is based on the use of a moving filter, and which can be used for the same type of data as polynomial trend surface analysis - that is x,y,z data where z is an attribute value of the coordinate, usually site densities - but produces (for a variety of reasons) better results than both trend surface analysis and filtering.

Kriging was developed in the 1970s within the mining industry (the technique is named after DG Krige a South African mining engineer), but has found wide application in other areas. Kriging is a method for spatial interpolation which is based on the recognition that spatial variation within certain variables (such as geology, soil type or hydrological property) is too irregular to be modelled by a smooth mathematical function (such as a polynomial trend surface or spline). Instead these are `regionalized variables' which are better described by a stochastic surface. This is achieved by modelling the stochastic aspects of a `regionalized variable' (which translates into a `choropleth theme' in GIS jargon) using regionalized variable theory. This makes the assumption that any spatial variable is composed of three components:

1. A structural component associated with a constant value: described by a function m(x)
2. A random component spatially correlated with (1) called $\varepsilon'(x)$
3. A random noise element not spatially correlated with (1) called $\varepsilon''(x)$

The mathematics are not straightforward (Burrough 1986), and are beyond the scope of this text but the main characteristics of Kriging are as follows:

1. It is very appropriate for modelling unknown values in cases where the spatial variables are known to be regional in nature - a situation in which, for example, polynomial trend surface techniques would be wholly inappropriate.
2. It is an exact interpolator, in other words the resulting interpolated data theme will pass through the data points. This can be contrasted with polynomial trend surface analysis which is not an exact interpolator, the surface not being constrained to pass through the data points.
3. A by-product of the Kriging equations, the estimation error can be mapped as well as the data values themselves - this can be a useful indicator of how reliable are the interpolated values. Obviously the estimation error rises with distance from the observations, reflecting the fact that the estimation becomes less reliable.
4. It is heavily computational, and in many circumstances the results may not be substantially better than those obtained from using a simpler method.

Logistic trend surface analysis

One other alternative which has recently been applied to archaeology is logistic trend-surface analysis. This is based on logistic regression analysis and is particularly suitable for archaeological problems because it is designed for nominal level data. Kvamme (1988) presents an archaeological model of lithic scatter presence/absence using this technique. Logistic regression techniques are an example of probability models and apart from the ability to use nominal level data, logistic regression techniques have the advantage that the results can appropriately be interpreted as a probability of a particular outcome, ranging from 0 to 1. This result is also an interval level result rather which means that the results from different land parcels can be directly compared: for example a probability of 0.6 is twice that of 0.3.

In Kvamme's example the output is the probability that a particular land parcel (50x50m cell) will contain a lithic scatter. The is also a corresponding output for the other case that each land parcel does not contain a lithic scatter, but this is simply the `negative' of the first result. The technique can also be extended beyond this simple two-class model to provide probabilities for several alternative outcomes, consequently logistic trend-surface analysis could also be used to generate a series of probability surfaces for different site classes.

There is, however, one possible disadvantage of a probability model for prediction of multiple site classes. Because the result of the logistic regression must be a series of probabilities of mutually exclusive categories (so that the probabilities for each of the cases must always sum to one for each land parcel), the technique is not applicable to circumstances where the site classes are not mutually exclusive. For example, the technique is perfectly acceptable to the prediction of the two-class case of site present/site absent because there can be no land parcel which both has a site and does not. On the other hand, it cannot be applied to the three-class case for prediction of Neolithic settlement sites, Bronze age settlement sites and no sites if it is possible for a land parcel to contain both Neolithic and Bronze age settlement sites because the three categories are no longer mutually exclusive. To circumvent this restriction it is possible to create a new category for land parcels with both and then use this as a four-case analysis - `Neolithic', `Bronze age', `none' and `both' are now mutually exclusive categories as no land parcel can fall into more than one class.

Deductive rule approaches using map algebra

The simplest way to construct a predictive model is to generate a *decision rule* which describes the conditions under which a particular location is likely to contain an undiscovered site. This rule can be based on one or more variables, and can be derived either entirely from theoretical hypothesis or from observations of an existing sample of sites.

If a simple decision rule is used, then the outcome of each implementation of the rule will produce a yes or no answer - yes if the rule determines that a site is likely to be present, or no if it is not. More sophisticated decision rules may be able to produce a variety of outcomes, representing different classes of archaeological outcome - typically different densities of remains, or different types of sites.

The inputs to a decision rule are generally a subset of the non-archaeological characteristics of the region in question. Thus the first assumption of the decision rule is that the locations of archaeological sites are related in some way to the non-archaeological characteristics of the region. In many cases this can be justified by observation of an existing set of sites, but occasionally no sites, or only a very small sample of existing sites are available. In these cases a sound archaeological case should be established for the assumption.

At their simplest, decision rules can be regarded as functions of the variables (themes) which are used as inputs. This can be described as in equation 2.30, where M represents the case for the model indicating the presence of a site, and M' indicating the opposite (that no site occurs).

$$\begin{aligned} \{f(x_1, x_2, x_3, \ldots) > 0\} &= M \\ \{f(x_1, x_2, x_3, \ldots) \leq 0\} &= M' \end{aligned} \qquad 2.30$$

Because the function produces a binary, mutually exclusive classification using more than one input, this type of function is called a multivariate discriminant function. Another approach is to use set notation to define M as the intersection of a variety of variables, in which case M' is the complement of that action. A typical example of such a decision rule might be that a particular class of sites tend to occur in flat areas (slope of less than 10 degrees), not too far from a source of fresh water (less than 1km), on a particular type of soil (class A) and in south facing aspects. This can be expressed as a decision rule, using set notation as follows:

$$\begin{aligned} M &= (Slope > 10) \cap (Distance < 1km) \cap (Soil = A) \cap (Aspect = South) \\ M' &= (Slope > 10) \cup (Distance \geq 1km) \cup (Soil \neq A) \cup (Aspect \neq South) \end{aligned} \quad 2.31$$

Some GIS systems will allow rules to be specified in their entirety in some high-level language, examples of this include the MapInfo `mapbasic' language and the SPANS internal language. Many simpler systems, such as IDRISI, however require that the decision rule be implemented in a series of steps.

Problems with predictive modelling

Groups of recorded sites are rarely a random sample of the population of sites. Therefore a predictive model based on the location of known sites, may simply identify those types of areas in which known sites occur. (extreme example: half of an area is buried by alluvium, sites occur on the half which isn't buried : the model will give high probability for no alluvium, and a low probability for alluvium). If this problem isn't addressed, a model can be mis-used with disastrous consequences - a unit may concentrate on looking for sites where they already occur, thus compounding an existing bias in the record. The finds of sites would then 'strengthen' the existing model, so the unit would go on looking for sites in the same type of area - recursively refining the bias in the archaeological record.

To some extent, this can be offset by recognising the potential pitfalls and adopting 'red flag' strategies (Altschul 1990). These entails identifying the general pattern (as above) but then flagging those sites or which deviate from the observed norm. Once flagged, the characteristics of these 'red flag' sites can be analysed, and a more sophisticated strategy for survey can be developed.

2.7. Summary

This chapter has established what constitutes a GIS by identifying those features which most clearly distinguish GIS from other forms of information technology. Some background to the devepopment of GIS was presented in section 2.3, to illustrate that GIS technology did not spring fully formed from software vendors and universities. Instead, it was argued, the origins of the different elements of GIS technology can be traced to a variety of different information technologies which have developed over the last twenty to thirty years. GIS, in a sense, was not invented but manufactured out of existing components.

However, although GIS does not represent a major technological breakthrough in its own right, the remainder of this chapter has argued that this integration of different components has produced a tool which is capable of revolutionising the way in which disciplines concerned with geographic information work. This has begun to happen within geography and within archaeological practice in the United States (see chapter 1) and it is the contention of this thesis that it will have a similar impact in the sphere of British archaeology.

The method and theory which underlies the creation of properly topological spatial databases was described in section 2.4 and 2.5 and the analytical capacities which must accompany the implementation of such a database in 2.6.

Chapter 3
Archaeological background

3.1. Introduction

This chapter will establish the background to the case studies in the two chapters which follow. The chapter will review the evidence from the Neolithic period of the study areas, and critically assess some of the interpretations advanced within the archaeological literature. The aim of this chapter is to provide the archaeological context for the case studies: it is argued that an interpretation of Neolithic economic history as a colonisation by migrant farmers, followed by a period of sedentary agriculture is no longer valid and that there is a need to rethink the dominant interpretations of the social organisation of the Neolithic period. With this in mind, the chapter will then identify archaeological questions which might be approached with GIS methods.

3.2. Archaeology and chronology of the Neolithic

3.2.1. Origins of the Neolithic

The land bridge, connecting Britain to the Europe, was probably removed by the mid 7th millennium bc, leading to a breakdown of long distance contacts between southern Britain and the continent. This in turn may have led to some reduction of the resources available to people in Britain, and certainly to a divergence of material culture traditions. The vegetational sequence of the British Isles continued from the earlier period, and was probably mixed oak forest from around the 6th millennium bc. The artefact sequence of the British Isles during this period begins with similar forms to that of the continent, but over the following two millennia, the pattern is of divergence - more geometric artefact styles appear in Britain, not matched on the continent, and later the continental traditions of blade-and-trapeze forms fail to appear in the British record.

Inland, there is evidence of expansive occupation, typically of larger lowland sites and smaller sites on higher, dryer, and therefore possibly more marginal land. An example of typical land use in Wessex during this period may come from Cranborne Chase, where there is convincing evidence (Barrett *et al* 1991) of a preference for river valleys and for the clay-with-flints deposits which overly the chalk. The latter is probably due to the obvious availability of raw material, but the degree to which flint implements were made at source, rather than elsewhere from transported raw material, will clearly effect this interpretation. Nevertheless, there is no evidence for Mesolithic activity in Cranborne Chase on the chalk downlands themselves, and it seems likely that a real preference for valleys and areas of clay-with-flints is revealed in the record.

Pollen evidence shows episodic clearances, and there is some evidence (although poorly dated) from Oakhanger in Hampshire for use of Ivy as fodder or lure for animals. There is little evidence of plant or animal exploitation on inland sites, except for the occurrence of the split shells of hazelnuts which must have been an important food resource, supplementing the staple foods of aurochs, cervids and pigs, and their growth may have been aided by the periodic small-scale clearances shown in the pollen record.

Shell middens dating from the later Mesolithic have been identified in a number of places in Britain, although as Jacobi points out (Jacobi 1976), coastal erosion may have destroyed many of these sites in Wessex. At these coastal sites, occupation may have been semi-permanent. Otoliths from Oronsay, for example, indicate that there is no consistent season of occupation for these sites: the four sites investigated produced four different seasons of occupation. The abundant resources exploited in these coastal ecotone environments might be typified by Morton in Fife where remains include land animals (roe & red deer, aurochs, pig), shellfish (cockles, crab) and fish (cod, haddock, turbot, sturgeon and salmon). The occurrence of spatially distinct artefact categories, possibly implying the existence of separate social territories has been identified in the British record (Jacobi 1979) and from the 6th/5th millennium, relations between these territories may have been mediated by exchanges of raw materials and portable artefacts. From about 5200bc, a realignment of economic land use then led to increased use of coastal resources (Jacobi 1979).

Generally, therefore, the later Mesolithic may have been a period of considerably less mobility than some traditional models of hunter-gatherers suggest, with fairly large coastal sites exploiting on a more or less permanent basis the wide range of resources available to them, while inland, arboriculture and horticulture may have been practised to some extent, reducing the density of forest cover and increasing the availability of browsing animals and edible plants.

Dating of later Mesolithic sites is based mostly on lithic typology with few radiocarbon dates. However, the microlith industries characteristic of the later Mesolithic have not been dated to later than about 4000bc. If this is coupled with the general absence of dates for characteristic Neolithic material before about 3300bc (Gardiner, 1984), this leaves an apparent gap of around 700 years in the archaeological record. To some extent this gap may have been artificially created by the classification of later Mesolithic and earlier Neolithic flint industries: published accounts generally fail to highlight the fact that where leaf-shaped arrowheads (or other 'diagnostic' pieces) are missing from Neolithic assemblages, or microliths from Mesolithic ones they can be extremely hard to distinguish.

3.2.2. The nature of the evidence

Regardless of the nature of the transition, by around 3300bc the evidence from the Wessex area is different from that of the later Mesolithic. The phenomenon of the elm decline is visible in most pollen diagrams, although in light of such widely varying dates as c.3000bc from Rimsmoor and c.3600bc from Winnal Moor (Waton 1983) it cannot now be regarded as a synchronous event throughout Britain (Smith 1984). Nevertheless, the pollen and snail evidence leaves little doubt that a major change occurred to the environment resulting in a marked opening of the landscape. This is not to imply either that the landscape was predominantly open after the elm decline or that the pre-elm decline landscape consisted of uniform dense forest; rather that the small patchwork of clearings which must have already existed before the Neolithic now became, as Smith (1985) argued, enlarged and conjoined in selected areas, probably mostly on what is now the chalk downlands. Direct pollen evidence for the Wessex chalklands is scanty, but the core from Winchester (Waton 1982) suggesting that the elm decline dates to around 3600bc, also has cereal pollen appearing in the record at the same point. Both of the dates described above for Winchester and Rimsmoor in Dorset were 'projections' onto a time-depth curve, and so should not be granted the same credibility as dates for specific artefacts.

The distribution of lithic material in Southern England is markedly different in the early Neolithic from that of the preceding period, there are fewer sites and a heavy emphasis on the chalk (Gardiner 1984). Sites on the Greensand and gaults (generally with clay soils) are clearly preferred during the later Mesolithic, but are almost entirely avoided during the Neolithic period in favour of sites on the chalk. In addition to the characteristic leaf-shaped arrowheads, sometimes unfinished or broken, early Neolithic flintwork generally comprises small cores, a variety of scrapers, denticulated flakes and simple retouched pieces, with a reduced proportion of blades (Gardiner 1984).

For the first time in the British record there is evidence that large effort was expended on the construction of earthworks: the 'causewayed enclosure' sites and there is a visible burial record, generally in the form of multiple inhumations, often disarticulated within long mounds. Pottery is found for the first time in the British record, in the form of undecorated course vessels with round bases, commonest in Wessex being Hembury wares, dating from around 3330bc at Hembury to about 2580bc at Windmill Hill. Some of this pottery was imported from the Lizard, Cornwall (gabbroic wares) or from areas of Jurassic limestone ('oolitic ware'), while the majority seems to have been locally made, frequently copies of the imported pottery (Smith 1974, Peacock 1969).

There is some evidence that the ritual sites may have been deliberately located in areas not used for domestic purposes. In Cranborne Chase, for example, there seems to be a separation between the areas of greatest monument-building activity and of highest concentration of flint scatters (Bradley *et al* 1984). In addition, Gardiner has observed that, for Sussex, the distribution of early Neolithic leaf-shaped arrowheads clusters significantly around long barrows as compared with enclosure sites (Gardiner 1984), and that virtually all surface scatters and stray arrowheads are located on or near to clay-with-flints deposits, as are the long barrows in Sussex. Gardiner continues to suggest that the clay-with-flints deposits may have been foci for early Neolithic agricultural activity, as they may represent the remnants of fertile loess soils, easily cultivated. It remains to be seen whether similar patterns of spatial organisation exist in other areas, including Wessex.

3.2.3. Long barrows and other burials

The funerary monuments of the period can broadly be divided into two groups: 'earthen' long barrows and 'chambered' long mounds, although there is some evidence that some of the barrows which were made predominantly of earth in fact contained timber chambers which have subsequently collapsed (Ashbee 1984). There are probably a few early Neolithic round barrows, of which Mere 13d has been cited as an example (Thorpe 1984), although whether these represent a variation or a later development is unclear.

Some, such as Normanton Down (Vatcher 1961), covered the remains of pre-existing structures, interpreted as mortuary houses or exposure platforms (Ashbee 1984). Although it is clear from remains in other barrows that corpses were frequently deposited in a disarticulated state, it is not certain that the pits interpreted by Ashbee as 'mortuary houses' were always for this purpose. More likely, there was some considerable variation in the treatment of bodies, either contemporaneously or through time. Some barrows, such as West Kennet (Pigott 1962) exhibit elaborate sorting and selection of the bones (Shanks and Tilley 1982), suggesting that they were not only placed in the barrow un-fleshed, but that the bones were moved subsequent to inclusion. In some cases, bones were burned, although not cremated, probably to remove the flesh. As discussed above, the location of the excarnation is unclear, but probably includes mortuary structures such as Normanton, some Enclosure sites such as Hambledon (see below) and also simple pits such as those recovered beneath some barrows.

The normal orientation of long barrows is with the entrance facing approximately east or south-east (Ashbee 1984), with some deviation, particularly in cases where articulated remains are included. Those with disarticulated burials, or many burials tend to be smaller. It is clear that barrows with an unusual form (bank barrows, U-ditched barrows, oblong or oval barrows etc.) or orientation tend to include few bodies, and these are more likely to be articulated (Thorpe 1984).

The number of burials within long barrows varies considerably, from barrows such as Fussell's Lodge (Ashbee 1966) where a large number of individuals are represented, to those such as Beckhampton Road and South Street (Ashbee *et al*, 1979) with no burials at all. A pattern can be observed in that where the bodies are disarticulated, there are almost always more of them; all single inhumations, for example, are articulated (Thorpe 1984).

It seems likely that those barrows, such as Fussell's Lodge, with many disarticulated bodies are earliest (Fussell's Lodge is dated to around 3230bc), and that there is a relationship between date and number of individuals included: those with few bodies are later, and those with single articulated remains or none at all are the latest. It is also probable that deviant barrows, (oval, round, bank etc.) are later (Thorpe 1984).

Spatial patterning in the distribution of long barrows has been noted by a number of writers including Grinsell (1958) and Ashbee (1984), who broadly divided the extant barrows into five groups, the Dorset Ridgeway group, the Cranborne Chase group, the North Wiltshire (Avebury) group and two groups on Salisbury Plain east and west. While the distribution maps of long barrows certainly seem to suggest that these groups represent a real grouping in prehistory, this hypothesis has never been properly tested. It is plausible that the areas visible as 'groupings' in Ashbee's distribution map (Ashbee 1984. p10) reflect not the distribution of barrows in prehistory but areas with a high survival potential for such monuments. Indeed one of the features of these barrows often commented on is their tendency to occur on higher ground, while they are not visible in the valley corridors of the region. This may, as it is usually assumed, represent a desire by the builders to use marginal land not required for farming, or to site the monument in an area of high visibility, but there has been no comprehensive attempt to test this.

3.2.4. The enclosure sites

Variously known as Causewayed Camps, Causewayed Enclosures, 'causewayed enclosures', and even 'interrupted ditch enclosures', these sites have proved enigmatic, and most discussion has centred on their function as domestic, defensive or ritual monuments. That they have been classed as a unit is misleading: there is a huge variation in the nature of these sites. Firstly, the size varies; broadly they can be divided in to 'large' and 'small' enclosures, Maiden Castle, Hambledon Hill and Windmill Hill being the largest, all enclosing over 17 hectares, while all others enclose less than 10 hectares, except Whitehawk, Sussex, which is slightly larger (Smith 1971).

Site	Grid Ref	Construction Date?
Knap Hill	SU122636	2760+115 bc
Robin Hood's Ball	SU102460	?
Rybury	SU083640	?
Whitesheet Hill	ST802352	?
Windmill Hill	SU087714	2950+150bc
Hambledon Hill	ST849122	2790+90 bc
Maiden Castle	SY669885	?

Table 3.1 Excavated Causewayed Enclosures in Wessex (After Smith, 1971).

The only factor that these sites can really be said to have in common is their construction method. This generally comprised the digging of an interrupted ditch, probably by discrete groups or gangs of people (Smith 1971), from which the spoil was used to create an internal bank. The bank frequently contained fewer interruptions than the ditch (as with the outer ditch at Windmill Hill), and may in many instances have been supported or topped by timber palisading.

The sites, however, vary considerably in form, Knap Hill comprises only a single interrupted ditch, Robin Hood's Ball, Maiden Castle and Coombe Hill have two concentric ditches, Windmill Hill has three and Hambledon Hill has two enclosure ditches and outworks. Some (such as Hambledon, Windmill Hill and Robin Hood's Ball) have produced evidence for posts or timber revetment of the bank, while at Coombe Hill post-holes were sought but never found.

Some enclosures have connections with domestic activity, a settlement (c. 2950bc) preceded the enclosure (c. 2550bc) at Windmill Hill, domestic activity was indicated earlier at Hembury, Devon (c.3300bc to c.3150bc) and some activity was identified at the Maiden Castle enclosure. Organic layers at a number of enclosures have been interpreted as settlement refuse (Whittle 1978), although the organic debris in the ditches at Hambledon cannot be interpreted in this way due to the structured nature of the deposition and the high proportion of human skeletal material.

The role of enclosures in the burial record has been discussed above; examples include Hambledon Hill (Mercer 1980), where a number of human skulls had been seemingly deliberately placed in the base of the ditch. All were without mandibles or cervical vertebrae, indicating that they had been de-fleshed when placed. Also at Hambledon, two complete and articulated crouched child burials were located in the base of the ditch, beneath flint cairns. The excavator considered that the position of the bodies could only have been obtained if they had been bound or placed in a container before interment. A third burial at Hambledon comprised the articulated remains of the trunk (femurs, pelvis and lower vertebrae) of a young man, again covered with flints. This deposit showed traces of animal gnawing, and suggests that the largest part of a body in a fairly advanced state of decay had been placed in the ditch. Quantities of human bone were also contained within an organic deposit in the base of the ditch, suggesting that more human bodies were present at the site. At the nearby smaller Stepleton Enclosure, the articulated skeleton of a young man was discovered with a leaf-shaped arrowhead within his thoracic cavity. He had been deposited with a considerable amount of charcoal in the terminal of the southern ditch, next to what the excavator interpreted as a gateway.

At Windmill Hill, a crouched inhumation was recently discovered in the ditch (Whittle *pers comm*) to add to the two young child burials and scattered bones in the primary levels of the outer ditch, already known from Keiller's excavations (Smith 1965). A notable feature of these places is that all eight of the recorded Neolithic child inhumations in Britain are from enclosure sites. Connections between enclosures and Long Barrows have been suggested for Windmill Hill and West Kennet, where the under representation of long bones and skulls at the barrow can be contrasted with the over representation of exactly those bones at the enclosure (Piggott 1962, Smith 1965). A similar connection has been suggested for the Hambledon Hill enclosure and Barrow.

Early interpretations of these sites assumed that they, like the Hill Forts of the Iron age, had a defensive function; perhaps as defended villages. This was dismissed by, for example, Smith (1965, 1971) on the grounds that the construction method lacked strategic soundness - the interruptions in the ditch were said to be far to frequent for effective defence, and the monuments did not always take account of the natural topography in an effective way. Nevertheless, it cannot be entirely coincidental that sites such as Maiden Castle, The Trundle and Hambledon Hill were all subsequently adopted for use as Hill Forts, monuments for which a defensive interpretation has never been seriously questioned. Windmill Hill and Knap Hill also dominate the local landscape, at Knap Hill for example:

> *'The position ... is a commanding one, with fine views to the east along the scarpment and the south east to the Vale of Pewsey. On these sides it is extremely steep, especially on the east ...'* (Connah, 1961)

These locations command a highly strategic position within the local landscape; at Hambledon and Maiden Castle particularly there is an exceptionally good view in all directions from the monuments. Not all of the enclosure sites are in such positions, as has been shown by the discovery of a number of apparently similar sites on the gravels of the Thames Valley, but for those that are the possibility that they are sited strategically for easy defence, and the implication that conflict between groups was a very real part of Neolithic society, should be given more credibility. It has been shown that there were generally fewer interruptions in the banks than in the ditches, and suggested that the interrupted nature of the ditches is an artefact of the construction method. Neither of these characteristics therefore seem to preclude the possibility that the sites at Hambledon and Maiden Castle, for example, might have had defence of people and livestock as a significant part of their functionality.

There is, of course, other evidence of conflict during the earlier Neolithic, and associated with the enclosures, in the form of leaf arrowheads. Although these may in many cases have been deposited with a ritual intent, the sheer number of examples at Crickley Hill, and their association with burnt timbers strongly suggests a more functional explanation; and, as noted above, at the Stepleton Enclosure, Hambledon Hill, an individual was found in the base of the ditch, seemingly where he died, with a leaf arrowhead in his chest (Mercer 1980).

Enclosures have suffered, it seems, from over-classification. As discussed above, about the only things they have in common are their approximate date and construction method. The argument as to the function and significance of 'the enclosure sites' can therefore now be disregarded, as it is clear that this is a sterile topic: there is no one function for 'the enclosure sites', and it should be clear that they must each be treated on their own merits.

3.2.5. Other monuments

One other monument class can be shown to belong to the first half of the fourth millenium BC: the cursus monuments of which there are several in Wessex. The most impressive of these is the Dorset Cursus running through Cranborne Chase between Thickthorn Down and Martin Down. The monument was built in two sections, first from Thickthorn Down to Bottlebush Down, where the first terminal can be identified, and later from here to the new terminal at Martin Down. Excavation of sections through the ditches of the earlier section showed that this ditch had begun to silt before settlement associated with Mortlake and Fengate Wares. More precisely, bone from the primary silt provided dates of 2625±77bc and 2620±120bc, broadly confirmed by the date from an antler pick, also from the primary silt, of 2540±60bc (Barrett *et al* 1991).

The monument consists of two shallow linear ditches, 50-100m apart, about 1.5m deep, 3m wide at the top tapering to 2m at the base. The remains of an internal bank, probably about 3m wide, possibly revetted, have also been identified and probably accompanied this ditch for the full length of the monument. At various stages in its length, the cursus incorporates or encloses long barrows, most notably it encloses the barrow on Gussage Cow Down which is visible from both terminals, and incorporates a barrow into the north bank between here and the Martin Down terminal. At Martin Down, it seems that an pre-existing long barrow, abutting the cursus terminal was extended into a bank barrow after the cursus was constructed, the extension being added to the 'front' of the long barrow instead of the 'back' because the natural extension direction was blocked by the cursus (Barrett *et al* 1991). In addition to this, the monument can be seen to be generally axial to the distribution of barrows in the area, emphasising the close relationship between the cursus and the barrows.

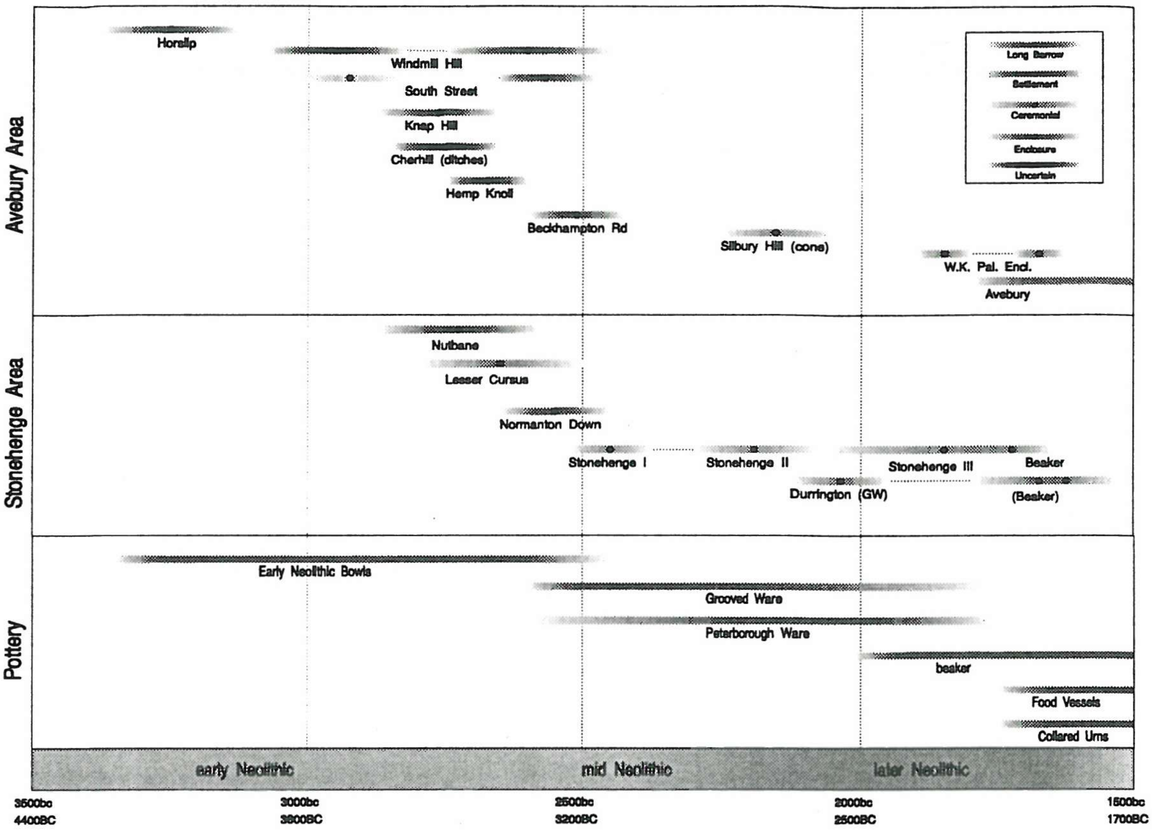


Figure 3.1. General chronological sequence for the Avebury and Stonehenge areas and approximate chronologies for the related pottery.

3.3. Interpretations of the Neolithic

3.2.1. The Neolithic economy

The 'earliest Neolithic'

The traditional view of the period between around 4000bc and 3300bc is that it represents an 'earliest Neolithic' phase (e.g. Case 1969), during which mixed farming was practised by migrant settlers. To explain the lack of physical remains from this phase, it was claimed that monument building was not possible due to the pressures of establishing the new, agricultural economic base. In this model, settlers arrived from the continent around 4000bc, bringing the techniques of cereal cultivation and animal husbandry with them. It is normally supposed that the method of farming adopted would have been long-fallow or 'slash and burn' during this pioneer phase. The early farmers would undertake the clearance of an area of forest, possibly by felling with stone axes or by burning, which has the supposed advantage of releasing potassium (potash) fertiliser into the soil. This would be followed by a few years of intense cereal cultivation until the soil fertility declined to an uneconomic level. At this point the location would be abandoned to be reclaimed by the forest (long-fallowed) and the cycle begun again somewhere else, leaving little trace in the archaeological record except in pollen diagrams. After some 500 to 1000 years of this an economic surplus was built up, and permanent clearings in the forest cover established. The population was then free to expend energy on other activities including the building of earthworks and burial monuments.

There are two primary assertions in this model. Firstly there is the explanation offered for the absence of monuments dated to this pioneer phase. Bradley (1984a) has argued that this explanation has problems, most notably that there is no ethnographic evidence to suggest that monument building would be postponed by a society in the process of establishing a surplus. Drawing on the ethnographic work of Meillassoux (1972) Bradley has observed that, in societies which are economically similar to those which probably existed in the earlier Neolithic (in other words early agricultural societies), there is competition for resources such as agricultural land. Although such competition would also exist in non-agricultural societies, among agriculturalists the work of one generation would have a direct effect on future generations; for example, land cleared and cultivated by a community would become a valuable resource to their offspring in the future. In this way, Bradley argues, rights to resources such as land would become linked to the notion of ancestors, and communities would be inclined to stress links with ancestors as a strategy for retaining and increasing their economic resources. The interpretation of the tombs as evidence of mortuary ritual related to ancestor rites rather than as burial ritual is also advocated by Barrett (1988) who contrasts the architecture of the long mounds with the later round barrows: at chambered long barrows, bones of the dead could be accessed and manipulated collectively after initial burial, suggesting that the rite was concerned with the continuation of the dead into the realm of ancestors. In the later monuments, the bones were generally interred permanently indicating a burial rite.

Bradley also argues that it is less credible still that an immigrant population arriving in about 4000bc would entirely abandon its material culture, only to resume it 700 years later. In addition, and contrary to the traditional view, it may be argued that times of greatest economic stress are exactly those times when societies are most likely to emphasise social cohesion by the construction of burial monuments and other ceremonial structures (Chapman, 1981).

Secondly, however, there is the claim that the earliest farmers would have been forced by their inability to conserve soil fertility to adopt 'slash and burn' strategies. This long held belief has also been recently questioned; for example, Jarman *et al* (1982) dissolve a number of the assumptions underpinning it. Firstly, there seems little justification in the assumption that early farmers were incapable of maintaining or enhancing soil fertility other than by burning forest. Neither the manuring of fields nor the rotation of crops require any special equipment, and neither would have left any significant trace in the archaeological record.

Jarman *et al* (1982) also observe that the earliest crops, such as einkorn (*Triticum monococcum)* and emmer (*Triticum dicococom)* are substantially less demanding of soil fertility than modern varieties. This suggests that early farmers could have effectively cultivated good soils for many years with little deterioration in yield without a long-fallow strategy and even without resort to any soil improvement such as manuring. Further, they make the point that the loess soils favoured by the earliest farmers in Europe (to judge by the distribution of Linear Pottery settlements) are characterised by very high base fertility, considerable soil depth and good drainage. Although not obvious today, it is clear that the soils overlying the Wessex chalklands would similarly have had a substantial loessic content in the past (Catt 1978, PF Fisher 1991), so there seems little to support the assumption that crop yields would quickly deteriorate, whether or not the earliest farmers used manuring.

The implication that early farmers lacked even moderately sophisticated land management tactics may also have been responsible for the claim that infestation by weeds and parasites would have compelled a shifting system of cultivation: for example bracken, which must have been one of the main agricultural weeds of the chalkands, may have been controlled by the use of pigs to root out bracken rhizomes (Smith 1985), and this may be only one example of the range of agricultural strategies available to early farmers. It also seems that neither the drainage characteristics nor the base fertility of these types of soil would gain any significant benefit from the addition of potash (Jarman *et al* p141) which is the only other significant benefit of long-fallow agriculture.

The developing Neolithic

Whatever the truth about the earliest period of farming, by around 3300bc, there seems little doubt that both cereal cultivation and animal husbandry were practised in Wessex. Which form of farming was practised at any given location in the landscape would be dependent on the nature of the particular local environment, on the previous land-use history and on the cultural history of the people living there. It seems probable that arable and pastoral land-use were complementary, in that nitrogen depletion of the soil by cereal cultivation could be partially or wholly remedied by manuring. Grassland (maintained by grazing) is more suitable for cultivation than cleared woodland both because, even allowing for the energy requirement of cutting turves, there is a lower energy requirement to begin (or restart) cultivation, and grassland also has a higher humus content than cleared woodland. Pigs in particular may have played a vital role in the clearing of partially regenerated land by rooting out bracken from scrub land prior to re-use as arable (Smith 1984).

It is also to be expected that Neolithic land use would have been variable, with periods of ploughing and cereal cultivation inevitably perhaps followed by nitrogen depletion of the soil, abandonment of cereal cultivation followed by either adoption for pastoral use or regeneration of scrub; at any given time small clearings and larger areas within a small region would exhibit considerable variation in their use depending on their stage in the cycle. This sort of pattern can indeed be observed in Wessex, for example in the buried soil beneath the South Street long barrow (Ashbee *et al* 1979), where the earliest pollen evidence is for woodland with some clearings (but not cultivation) fairly close to the site, whereas the molluscs indicate cultivation in the immediate location (also attested to by the ard marks at the base of the pre-barrow soil). The pollen and molluscs from the upper part of the profile indicated pasture, although with considerable quantities of bracken. This last element may represent, as Whittle (1979) has suggested, regeneration of scrub after abandonment of the land, but it should not be taken as a direct indicator of the domination of the bracken species because, as Smith (1984) has observed, bracken spores are persistent in the soil and will tend to be over represented whilst other species characteristic of thornscrub are naturally low pollen producers.

Recently Entwhistle (1989) has challenged the accepted view (of e.g. Whittle 1979) that cereal cultivation was very widespread during the later fourth millennium, and claimed that the environmental evidence for clearance could equally as well be for animal husbandry as for cultivation. To what extent the economy was based one or other of these elements is unsure, but Entwhistle's view that there is essentially *no* evidence for cereal cultivation does ignore a significant body of evidence: plough marks beneath earlier Neolithic long barrows have been found in a number of locations, grain impressions have been found in pottery from Windmill Hill and elsewhere and cereal pollen (or weeds associated with cultivation) have been identified from buried soils and peat sequences.

If a genuine colonising phase can be discounted, there still remains unresolved the problem of the apparent gap in the archaeological record between c. 4000bc and 3300bc. The solution to this problem may lie in the observation by Jacobi (1979) that the later Mesolithic saw considerable emphasis on coastal resources, and by Bradley (1985a, 1984b) that those coastal sites from this period which have been found occur in areas of isostatic uplift, in the north of England and Ireland. It might be expected, therefore, that the very nature of coastal sites in southern England and Wales, areas which have not been subject to uplift, may render them especially prone to destruction. If the last phase of the Mesolithic is really characterised by coastal settlements, now in the sea, then it is perhaps less attractive to push back the origins of the Neolithic to c.4000bc, and rather more likely that the Mesolithic period can be extended some way forward towards 3300bc. Also, it may be that the differences between the 'Hunter-gatherers' of the Mesolithic, and the 'Farmers' of the Neolithic have suffered from the over-application of precisely these stereotypes, and that there is rather less of a clear distinction between the two lifestyles than these would imply. There is merit in the more sophisticated classifications of Jarman *et al* (1982) of terrestrial animal and plant exploitation each into stages:

Terrestrial animal exploitation

1. *Random predation* - Wholly opportunistic, no attempt is made to control animals or predict regularities in animal behaviour.

2. *Controlled predation* - A degree of control is exerted on animal movement, although only at certain times, such as game drives or coralling.
3. *Herd following* - A human group or part of a human group maintains some contact with particular animal populations.
4. *Loose herding* - Control of animal movement, at least at some times of year. Initiation of movement from summer to winter grazing, for example, as in many pastoralist economies.
5. *Close herding* - Close control of animal movements all the year around. May involve use of fences etc.
6. *Factory farming* - Animals are maintained in a wholly artificial environment.

Terrestrial plant exploitation

1. *Casual Gathering* - opportunistic exploitation of nearby plants with no attempt to husband or increase the resources.
2. *Systematic Gathering* - particular plants become the focus of systematic activity, site location or seasonal movements may be influenced by location of plants, and some husbanding may occur.
3. *Limited Cultivation* - Productivity is encouraged by transplantation, weeding, irrigation etc. Commonly horticultural rather than agricultural.
4. *Developed Cultivation* - Propagation of crops under human control, frequently in situations where the plants would be unable to propagate themselves.
5. *Intensive Cultivation* - Sophisticated systems such as selectively bred plants, greenhouses etc. where the entire plant environment is under human control.

With this in mind, it is possible to look again at the later Mesolithic and earlier Neolithic without attempting to force the evidence into one or other economic stereotype. For example, the evidence of clearance during this period, pre-dating the classic elm-decline dates of around 3300bc, and the sparse environmental evidence that some agriculture was practised (Groenman-van Waateringe 1983) may not be associated with migrant farmers, and could perhaps be tentatively assigned to adaptation of the existing economy. Barker and Webley (1978) have suggested that the forest cover during this period on the southern chalklands may have been substantially less uniform than was earlier believed, and they argue for a mosaic of dense forest, lighter vegetation and clearings, depending on local variations of drainage and aspect. Although they may have overstated the case for an open landscape on the chalk, these observations have value and would not require the substantial clearance of the forest cover which the elm decline seems to represent in order for local communities to herd aurochs and pigs, and to grow some cereals to supplement their existing diet of hunted red deer, aurochs and pig and some edible wild plants such as hazelnuts.

There are some other reasons for postulating a degree of native development of agriculture during this early phase, many of them deriving from the new perspectives recently brought to the later Mesolithic, which suggest a far more static native community than previously thought. This native population may have already practised clearance of forest cover to encourage browsing animals and edible plants, and the increasing emphasis on coastal communities during the period would have inevitably led to an increase in contacts with the continent, facilitating the importation of agricultural technology. Bradley has rightly drawn attention to the chronological gap between the initial use of agricultural techniques and the appearance of monuments and pottery in the British record, and has observed that when these features do appear, the style zones which have been recognised within them and postulated as social territories (e.g. Bradley 1984b, Renfrew 1973) are similar to those recently identified in the later Mesolithic (Jacobi 1979). Conversely, there is substantial evidence from earlier Neolithic sites that the population was still heavily dependent on wild resources: aurochs, pig and deer have been found at all sites, and hazelnuts, crab-apples and sloes have been identified at Hemp Knoll and Windmill Hill among others.

In terms of the Jarman *et al* classification, the period between 4000bc and 3300bc may be seen as representing a shift from *limited gathering* to *limited cultivation* and from *controlled predation* to *herd following/loose herding*. The 'Neolithic proper', beginning around 3300bc can then be seen as an economic shift from to *developed cultivation* on the one hand and *close herding* on the other. Neither of these economic changes need then be seen as so dramatic as to imply rapid or large-scale population movement. However, while the model of indigenous development seems fairly strong for coastal areas, Whittle (1990), has urged caution over the uncritical acceptance of native development models. Reviewing the evidence for the Kennet Valley near Avebury, he accepts some evidence of continuity in flint working and site use (e.g. Holgate 1988), but concludes that:

> *'... from the start of the Neolithic the Upper Kennet area was exploited in a new way, which amounts to a process of infill or colonisation.'*

He observes the occurrence of pits beneath Horslip long barrow, and of pre-monument activity at Windmill Hill and South Street and offers these as a candidates for ritual activity associated with early (pre-3300bc) migrant agriculturalists. Drawing on recent work in the Kennet Valley (Evans *et al* 1988, Mount 1991), he claims that there is no evidence that the wet, flood-prone land in valley bottoms was avoided by later Mesolithic communities, and may in fact have been extremely valuable as a permanent water source, attracting animals.

It seems possible, therefore, that neither of the currently popular models of agricultural beginnings are entirely correct. There are good reasons for rejecting the idea of agriculture arriving with a migrant population in the early fourth millennium bc, and to expect that this period may have been dominated by a continuation or even expansion of the exploitation of coastal resources visible in the later fifth millennium, combined with the introduction of new subsistence techniques including cereal growing in these areas. At the same time it appears that inland areas, where indigenous communities were probably exploiting the ecotones in the valley bottoms and to a far lesser extent the higher ground, may have experienced a degree of colonisation by new people during the earlier part of the millennium. Whittle's argument does not preclude the possibility of local Mesolithic communities gradually adopting the new economic practices and he admits that:

> *'There may be continuity with the late Mesolithic, or there may even be a gap in occupation.'.*

3.3.2. Social organisation and social evolution

The earlier Neolithic was clearly a period characterised by a different social organisation than the later Mesolithic. Specifically, however, it has been characterised as a period of emerging 'Chiefdoms'. The geographic groups of long barrows discussed above have been used by Renfrew (1973) to argue for the existence of social territories which can be equated with 'chiefdoms' in the earlier Neolithic of Wessex. Renfrew observed that these groups of barrows were each associated with an enclosure site (although in the case of the Avebury group there are three enclosures) and that these represent the territories of the chiefdoms. He also argued that the pattern persisted into the later Neolithic with the association of larger ritual monuments such as Henges with broadly the same territories.

Although hampered by an almost complete absence of settlement evidence, Renfrew proposed a model of Neolithic society derived from the location of ritual monuments in Wessex and based on a hierarchy of monuments deduced from the labour requirements of each. The hypothesis was that the observable groupings of long barrows (The Dorset Ridgeway, Cranborne Chase, East and West Salisbury Plain and Avebury) could each be associated with a Causewayed Enclosure site, and that these areas represented the political units of the earlier Neolithic. Renfrew further suggested that this pattern persisted into the later Neolithic when, except in two cases, a major Henge site could be substituted for each enclosure, and drew further support for this interpretation from the work of Fleming (1971), who showed four main concentrations of early bronze age round barrows which broadly corresponded to Renfrew's Neolithic polities. Renfrew suggested that the polities he had identified were in fact 'chiefdoms', which he defined as:

> *'a ranked society, hierarchically arranged, sometimes in the form of a conical clan where the eldest descendent in the male line from the clan founder ranks highest, and the cadet branches are ranked in seniority after the main line.' (Renfrew 1973 p.542)*

Drawing on the socio-evolutionary work of, for example, Sahlins (1958), he claimed that chiefdoms, could be distinguished from tribes by the presence of *'centres which coordinate economic, social and religious activities '*. This, then, is the function of the enclosure sites within Renfrew's model of the early Neolithic communities.

The model therefore requires that the enclosure sites share a single common function within each of the different areas, a view which has already been shown to be misleading. Although certain activities do seem to have occurred at several of the enclosures, it is clear that no single function can be ascribed to all the sites. The most that can be said is that they *'saw a number of activities, including feasting, exchange and the exposure of the dead, and are thought to have been used for seasonal aggregations ..'* (Bradley 1991 p50). However, while Bradley is undoubtedly correct, and feasting, exchange and exposure of the dead are all attested from enclosures, this does not imply that all sites were used for all of these purposes, and Bradley admits, contrary to Renfrew, that *'... it no longer seems as if they were the centers of social territories ...'* (ibid. p50).

If the long barrows mark 'family patches' in the way Renfrew assumes, then the model of hierarchical organisation of territories, with the territory of the chiefdoms (equated with the enclosures) being a superset of the territories of smaller social units (equated with the long barrows), also makes the assumption that the people of the period were entirely sedentary. It has been suggested above that the notion of a sedentary agricultural population is a simplification and that the economic evidence for the early Neolithic period is perhaps better interpreted as indicative of an only partly sedentary lifestyle.

There seems no good reason to assume *per se*, as Renfrew's model does, that the barrow groups represent the 'body' of political territories on the chalk uplands, and the enclosures the centres. It is possible that the barrows are the boundary markers for far less visible settlement centres in the valley corridors, and that the enclosure sites, far from representing centres of social territories, may have been far more marginal. Smith (1984, 1985), for example, has argued against the assumption that settlement centres were on the uplands. He has observed that sites in valleys are less visible than those on the uplands for taphonomic reasons: valley areas have been more extensively used than the uplands subsequent to the Neolithic and have often been subject to deposition of colluvium, burying sites. Combined with a lack of research in valley areas, even though when this has been undertaken it has shown considerable evidence of valley settlements in this period, Smith concludes that the weight of earlier Neolithic settlement was probably in the valleys.

There are also methodological grounds for questioning whether socio-evolutionary classifications of societies into 'tribes' and 'chiefdoms' are useful at all in the understanding of social organisation in this period. Chiefdoms have never been satisfactorily defined in a way which allows the concept to be universally applied: every case study of a so-called chiefdom presents a different definition, and generalisation from these cases produces a definition so broad and all-encompassing that it says nothing about the specific case. Renfrew's approach, like many socio-evolutionary studies, begins by accepting Service & Sahlins' view that chiefdoms exist and then proceeds to find out if they can be identified in the archaeological record by checking off the characteristics against a pre-determined list of 'chiefdom characteristics'. Nowhere in this approach is there any scope for asking whether chiefdoms really exist or not, and the reader is left with the impression that, if you look hard enough, a chiefdom could be identified in most societies. This is not to say that the 'characteristics' of chiefdoms are not in their own right relevant to the study of non-state societies: population density, agricultural productivity, ritual authority and craft specialisation, for example, recur in the ethnohistorical record as important factors in shaping social organisation but the arrangement of these component parts into a whole and they way they transform through time differs from culture to culture, depending on environmental and social circumscription.

But there are profound theoretical as well as methodological reasons for rejecting the idea that the Neolithic communities can usefully be characterised as 'tribes' and 'chiefdoms'. While some archaeologists do persist in trying to 'pin down' the nature of the chiefdom and correlate this with archaeological material (e.g. Earle, ed 1991), others (e.g. Shennan 1993, Yoffee 1993) have criticised the entire social evolutionary framework and begun to proffer alternatives. Barrett (1994), for example, argues that the type of hierarchical, direct control which is implied by these models is more appropriate to early state societies than to early agrarian ones:

> *'Institutional forms of surveillance such as the census, taxation, legality or military terror, operating over a claimed territory of jurisdiction from one or more centres within that territory, both characterize early states and distinguish them from non-state systems'* (Barrett 1994 p162)

Barrett also echoes the position of many current theorists that the 'tribe' and 'chiefdom' concepts are more a product of an ethnocentric desire to find hierarchical arrangements of power within Neolithic communities than a useful analytical tool. Implicit in the use of socio-evolutionary classes is the idea that the history of social organisation has been one of increasing complexity through time, culminating finally in the western industrial complex state. Many would not view social development in this way, and may even question that the western industrial complex state represents the peak of human achievement.

The rejection of the 'chiefdom' model need not present an interpretative problem. As Shennan puts it, the acceptance that social evolutionary approaches are ideological

> *'... helps us to escape from the deeply ingrained view that they* (non-state societies) *are evolutionary stepping stones, and from the associated tendency to look at them from an unsatisfactory teleological point of view as containing the seeds of future states'* (Shennan 1993 p53)

and their rejection, far from producing a theoretical and methodological vacuum actually holds greater theoretical and methodological merit than their use.

3.4. Case studies

3.4.1. Location of the case studies

Two areas are used as the basis for regional scale (around 20x20km) case studies: the Avebury area and the area of Salisbury Plain around Stonehenge. Each of these is defined broadly by a concentration of monuments, and are claimed, unlike Wessex as a whole, to represent regions with some archaeological coherence. Geographically, both the Avebury region and the Salisbury Plain region are at the core of an expanse of chalk landscape, and form discrete ecological areas. The Avebury region comprises the Marlborough Downs, which are bordered to the north and west by the steep scarp of the chalk, and to the south by the Vale of Pewsey, and is served by the upper Kennet River valley. The Stonehenge area is similarly bordered to the west and south by the edge of the chalk, and to the north by the Vale of Pewsey, and the core of this area is also defined between the valleys of the River Avon and the River Till.

During the earlier Neolithic, the two main monument forms can be seen to form spatially distinct clusters within these geographically defined areas, with little evidence for activity in the areas immediately surrounding them. Thus, in the Avebury region there are about 26 long barrows on the chalk around the Kennet Valley, but none in the Vale of Pewsey between the Marlborough Downs and Salisbury Plain. In addition, a distinct group of stone long barrows, which do not occur south of the Vale of Pewsey, occur in the Avebury region, in addition to the earthen ones which are found to the south, thus typologically this monument cluster can be sustained as distinct. A similar clustering of barrows occurs on Salisbury Plain itself, and supporting evidence for the clustering of long barrows is supplied by the distribution of the larger enclosures. As observed by Renfrew (1973), each of these areas of barrow concentration also contains a large enclosure monument - Windmill Hill in the Avebury region and Robin Hoods Ball on Salisbury Plain.

It must be admitted that there are two objections to the use of this pattern to determine study areas. Firstly, there a few monuments which occur between the two: Hatfield Long Barrow, for example, and the Marden Henge site in the Vale of Pewsey seem to be situated almost equidistantly between the claimed 'core' areas. In simple numerical terms, however, such sites are few compared to the multitude of sites which cluster within the core areas. In the case of Marden, this is a later Neolithic monument and so does not argue against this division for the earlier Neolithic. More significantly, perhaps, it might be argued that the observed distribution of monuments is more an artefact of differential preservation due to later land-use strategies than a reflection of the original distribution of these monuments. It is possible, for example, that agriculture has caused the greatest destruction of such sites and that the chalk uplands have not historically been as extensively used for agriculture as the lower areas. Further, it is likely that the bulk of colluvial and alluvial deposition has occurred in the lower areas, a process which has concealed many smaller monuments. Even if these effects are accepted, however, and it would indeed be surprising if neither of these processes had any effect, it is contended here that they are insufficient to fully explain the monument distribution described above.

It is also argued that despite the location of Marden between the two areas, this division persists into the later Neolithic. The construction of the major monuments of this period - Avebury in the northern area, and the two cursus monuments the various phases of Stonehenge on Salisbury Plain - takes place within the core areas. Round barrows follow a similar distribution to the earlier long barrows, and can be seen to cluster around these monument centres. In this later period, however, there are clearly extensions of the exchange networks evidenced previously: prestige goods such as polished stone axes and, later still, amber suggest longer distance trade, and it is for this reason that the second defining statement recognises the possibility that these two areas, and possibly the other areas of southern Britain, would have shared some broader cultural affinity. There is, however, sufficient evidence of regional differentiation of material culture and monument construction to support the contention that the two areas continued to represent discrete entities.

Figure 3.2. General location map for the Wessex region.

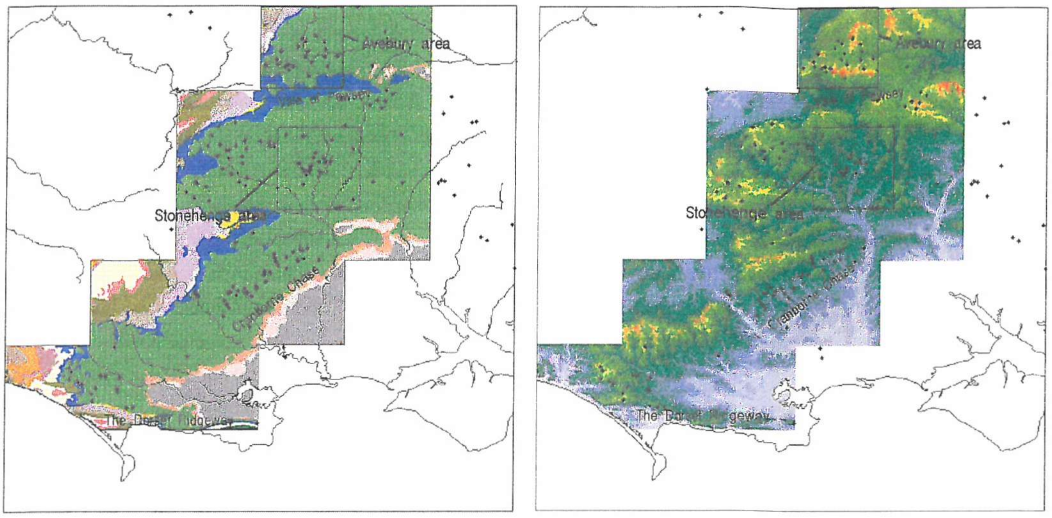


Figure 3.3. Location of the Avebury and Stonehenge areas, and other areas mentioned in the text. Showing the geology (left) and topography (right) of the study area.

3.4.1. General aims of the case studies

The above discussion has criticised some aspects of current interpretations of the Neolithic of the Wessex chalklands. Specifically the case has been made that the economy of the period need not be characterised exclusively as one of sedentary agriculture, and one consequence of this is that interpretations of the monuments as representative of the hierarchical nature of 'chiefdom' society require reconsideration.

The case studies in chapter 5 will address the interpretation of the spatial organisation of the same monuments which Renfrew (1973) used in the development of the chiefdom model, with the aim of taking advantage of GIS methods to develop new explanations for the monument locations. The study will not attempt to force the monuments into a pre-determined social-evolutionary category, but will seek to explain the individual practices which are evidenced by their location and how these changed through time. Two specific GIS methods will be applied in chapter 5. Firstly, the ability of GIS to manipulate distance and to introduce 'cost' into distance calculations will be used to explore the extent to which the interpretation of the causewayed enclosures as 'central places' may be supported. If not, alternative explanations for the location of the enclosures will be examined. Secondly, chapter 5 will use the ability of GIS to perform line-of-sight calculations to examine the relationship between the long barrows of the two regions with the aim of providing an alternative explanation for the locations of the long barrows.

Chapter 6 will explore the use of GIS at a different spatial scale of enquiry. Where chapter 5 is concerned with a regional scale of analysis, chapter 6 will address the relationship between the monuments and activity within a part of the Stonehenge study area of chapter 5. Although much of the aim of chapter 6 is concerned with development of methods for manipulating extensive survey data, and hence is methodological rather than explicitly theoretical, this chapter at the same time will explore the nature of flint-working activity during the period and the way in which location influenced the types of activity which took place.

Chapter 4
Construction of the study areas

4.1. Introduction

This chapter describes the sources of primary archaeological and geographical data which will be used during the case studies and explains the methods which were used to translate it from its original format into the GIS. It also outlines how this primary data was used to generate useful data themes through GIS transformations.

The data used in the studies will be classed here as either archaeological data, which relates to the locations and characteristics of cultural monuments and materials from the Neolithic, or as geographical data, which relates to the other characteristics of the study areas. Thus geographic encompasses all data relating to soil types, geology and landform. The two terms are used for convenience only as it is clearly the case that some archaeological information is a type of geographical data, while some geographical data is also archaeological.

Two GIS systems were used in the experiments in chapters 5 and 6. The IDRISI system (Eastman 1992) is a collection of programs written for the MS-DOS operating system. IDRISI was developed at Clark University, USA in collaboration with the United Nations Environment Programme Global Resource Information Database (UNEP/GRID) and the United Nations Institute for Training and Research (UNITAR). SPANS is a commercial GIS product produced by Tydac Technologies which makes extensive use of the quadtree data structure. Quadtree structures are a modified raster storage method based on successive division of a matrix into quadrants (Burrough 1986), and is particularly suited to storage of choropleth data in a raster format. SPANS also provides extensive vector capabilities and a modelling language which is used to produce maps of predictive models in chapter 6. SPANS versions 4.0 to 5.2 for the OS/2 operating system were used in this study. In addition to these, some analysis was undertaken with the GRASS system. GRASS is an extensive raster-based GIS system developed by the US Corps of Engineers and placed in the public domain. GRASS may be compiled for a number of variants of the UNIX operating system, and the version used in this study was compiled for the Linux operating system.

4.2. Availability of data

The extant sites of the Wessex chalklands together represent an unparalleled selection of visible monuments for the period in question. The geographic locations of the monuments are relatively easily accessible from documentary sources, from computer databases such as county Sites and Monuments Records and from the National Archaeological Record maintained by the RCHM(E). The existence of such data provides an essential starting point for GIS analysis.

Primarily because of the visibility of the monuments, particularly the henge monuments of Avebury and Stonehenge, the area has attracted field workers for many years. Data is consequently published regarding field surveys, aerial surveys, excavations and environmental sequences. This provides a large database of attribute data about the monuments themselves, and in some cases about the spaces between the monuments. In addition to this, the area is well covered by non-archaeological data sources such as the Ordnance Survey, Geological Survey and the Soil Survey. These sources provide essential environmental information on which to base the study.

Research within the chalklands has an extensive history and the area consequently provides a rich and varied archaeological literature from which to select areas of archaeological interest which are appropriate for the application of GIS. This archaeological literature is discussed more fully below and the precise areas of research which are selected for investigation are identified in the introductions to the case studies themselves.

4.3. Archaeological data

Three sources of 'primary' archaeological data were used during this project: a small subset of the computerised National Archaeological Record, a subset of the Wiltshire County Council Museums Service computerised Sites and Monuments Record, and some fieldwalking data deriving from the Stonehenge Environs Project (Richards 1990).

4.3.1. National archaeological record

The locations of the two major Neolithic monument types in the study areas (Long Barrows and Causewayed enclosures) are used in chapter 5. The Royal Commission for Historic Monuments maintains a computerised National Monuments Record (NMR) for England. The NMR includes a range of data sources including a database of historic buildings, a collection of aerial photographs which provides complete coverage of England and a record of archaeological sites and monuments. Data from the sites and monuments record was used as the primary source of monument locations.

The subset provided for use in this project comprised ASCII data listings from three relational tables, each selected through geographic and chronological constraint. The tables are linked by a unique site identifier referred to as the NAR_PRN code. The three tables are as follows:

- Period-Type-Form - contained period, type and form of each site
- NGR - contained the grid reference(s) of each site
- Land-Class - was designed to contain land class data but contained no information

The apparent relationship between these three data entities is shown in figure 4.1. All three tables contain multiple instances of NAR_PRN codes: it is possible for a single NAR_PRN to have more than one grid reference, to have more than one land class and to refer to more than one site type. This relationship is problematical in that it does not represent a normalised relational data structure.

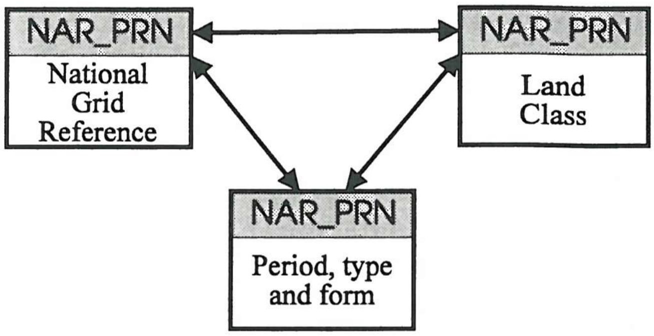


Figure 4.1. Entity-relationship diagram relating the three data tables from the NMR. The arrows indicate a one-to-many relationship between entities, while the text is a summary of several fields..

The listing files were imported into the dBase III+ database program. The land class table proved to contain no data for sites within the study area, and was therefore discarded. The remaining tables were normalised through the creation of a third entity to store a single record for each NAR_PRN code. This results in a normalised data structure as shown in figure 4.2.

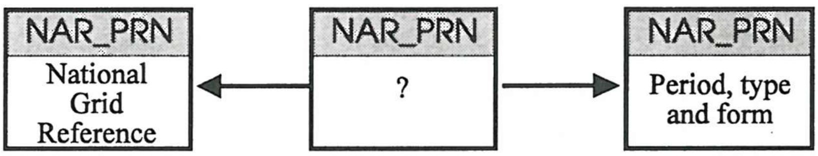


Figure 4.2. Normalised entity-relationship diagram including the nar_prn entity.

Figure 4.2 reveals a problem in regard to the data. The lack of an explicit NAR_PRN entity in the supplied database has allowed an uncertainty to become established within the data model. Multiple grid references are required for each NAR_PRN, to allow for monuments with a large geographic extent (linear banks for example) while multiple period-type-form descriptions are required because several 'sites' may share the same location. Linking these *only* through the NAR_PRN, however, means that it is sometimes impossible to identify which site form (i.e. 'bank' or 'long barrow') belongs with which reference. Fortunately all instances of long barrow and enclosure records within the data are linked to a single geographic reference.

The locations of all recorded long barrows and enclosures were then exported to (a) the IDRISI GIS vector data format (point entities) using a short interpreted script program and (b) to comma-delimited ASCII text files, from where they were uploaded as distinct point-entity data themes into the SPANS GIS.

The resulting monument locations were then checked against archaeological sources (e.g. Barker 1985) in order to remove erroneous records, and against Ordnance Survey maps to check location data. This revealed a number of omissions and errors within the data derived from the NAR, and the themes for barrows and enclosures were interactively modified and updated within the GIS to account for this.

4.3.2. Wiltshire SMR

Data regarding the location of archaeological monuments on Salisbury Plain, notably the locations of Round Barrows used in the prediction of lithic density in chapter 6 was obtained from the Wiltshire sites and monuments record, maintained by the Wiltshire County Council Library and Museums Service. This data was supplied in 'flat table' arrangement, in ASCII format and was simply imported into the dBase III+ database program using the existing facilities of dBase itself. The fields supplied and their format were as shown in table 4.1.

Field	Field Name	Type	Width	Dec
1	GREF_E	Numeric	6	
2	GREF_N	Numeric	6	
3	SITENAME	Character	36	
4	SMRNUMBER	Character	12	
5	GRIDREF1	Character	2	
6	TYPE	Character	30	
7	SUBTYPE	Character	23	
8	PERIOD	Character	2	
9	ALTITUDE	Numeric	3	
10	GEOLOGY	Character	6	
11	SAM	Character	6	
12	SAMCODE	Numeric	1	
13	CONDITION	Character	10	
14	CONDCODE	Numeric	1	
15	LANDUSE	Character	10	
16	LANDCODE	Numeric	1	

Table 4.1 Fields and format of data supplied from the Wiltshire SMR database.

As for the NAR data, a short dBase program was written to extract the locations of round barrow monuments as point entity vector files in vector formats suitable for both spans and idrisi. A vector overlay of round barrow locations was then available for generation of the barrow density surface used in chapter 6.

4.3.3. Stonehenge environs project fieldwalking data

Attributes of fieldwalking data

Monument location data, obtained from the NAR and from the Wiltshire SMR was regarded exclusively as vector format data at the construction stage of the project. The lithic density data which was the result of field collection, however, was treated in a different way, with the aim of storing the primary data themes for fieldwalking as raster data. This is essentially because, while monument locations can be considered as a complete record of the location of a monument (albeit as modelled by a simple graphical entity), the fieldwalking data represents a regular grid of *sampled* areas of the material. This is analogous to the regular sampling undertaken by a digital scanner, or a remote-sensing camera, each of which produces raster data. Table 4.2 lists the fields supplied.

Field	Field Name	Type	Width	Dec
1	EASTING	Numeric	6	
2	NORTHING	Numeric	6	
3	CORES	Numeric	3	
4	CORE_FRAGS	Numeric	3	
5	TOTCORES	Numeric	3	
6	FLAKES_NO	Numeric	3	
7	BROKEN_FLA	Numeric	3	
8	TOTFLAKE	Numeric	3	
9	BURNT_FLIN	Numeric	3	
10	RETOUCH_FL	Numeric	3	
11	SCRAPERS	Numeric	3	
12	OTHERTOOLS	Numeric	3	
13	TOTFLINT	Numeric	3	

Table 4.2 Fields and format of data supplied from fieldwalking records.

Although the Stonehenge Environs project (SEP) collected both flint and pottery for these areas, only flint counts are currently available. The easting and northing fields record the NGR of the start of each 25m transect, while the database consists of counts of artefacts in the following classes (Richards 1990 p16):

1. Number of whole cores (pieces showing traces of flake removal by percussion)
2. Number of broken cores
3. Number of whole flakes (pieces removed by percussion)
4. Number of broken flakes
5. Number of burnt flint (examples of worked flint which had been subsequently burnt)
6. Number of retouched flint (flakes with traces of consistent secondary flake removal or modification by percussion)
7. Number of scrapers
8. Number of other tools

Inspection of the database revealed that the last two categories (scrapers and other tools) were incompletely present and therefore analysis of these categories was not possible. This was not regarded as a serious restriction on the analysis as one of the aims of these experiments was to develop methods for investigating the material which was not dependent on 'type fossils'. The numbers of unbroken artefacts were generally rather small, and therefore some further generalisation was felt to be desirable in order to ensure that the data classes used contained sufficient variation within for meaningful regression analysis (see below). The analysis therefore uses the following categories of flint data, which were derived from the original classes:

1. Total number of cores (whole and broken)

2. Total number of flakes (whole and broken)
3. Number of burnt flint pieces (as above)
4. Number of pieces with retouch (as above)
5. Total number of all flint artefacts recovered

Resolution of fieldwalking data

Drawing on the experience of Maskell (1993), the data for the five categories of flint was converted from counts per transect into counts per 50m square. This was felt to be the appropriate resolution for the analysis of the data, requiring a minimum of alteration to the original record while providing a conveniently square unit of analysis for translation into raster overlays. The use of hectare units, as attempted by Maskell, was felt to be too great a generalisation at this stage of analysis, and the maintenance of as much of the existing record into the GIS was felt to be desirable. Maskell's conclusion (p20) that the 50m resolution (4 units per hectare) provides a good compromise between aggregation of the data and efficiency of processing was therefore accepted. The transects were regarded as systematic samples of a 50x25m unit of landscape and a transformed data file was generated as follows:

- For 50m units containing two transects, the sum of the two transects was used as the value for the new unit.
- For 50m units where only one transect was taken it was necessary either to discard the single transect, or to scale the value for the 50m unit appropriately. In order not to discard potentially useful data, it was decided to use these instances but to scale the value (double it) in order to maintain the same unit of measurement in each land parcel. Although preserving the extent of the surveyed areas, this strategy may have inadvertently exaggerated some edge effects (see chapter 6).

Once this was achieved the database represented a regular grid of samples of 50m quadrats. These were then exported from the database as ASCII text and imported into the GIS as a point data file containing 2587 observations of the five selected attribute variables. This was exported from a database by a short program which generated a file in the correct format for the SPANS GIS table input, and this was in turn imported into the SPANS GIS as point data, the counts of the five flints forming the attributes of the points.

4.3.4. Translation between IDRISI and SPANS vector files

One of the tasks frequently encountered during this thesis, was the translation of this type of vector data from one GIS system to another - some tasks were easier to undertake in the IDRISI system than in the SPANS system and others easier with SPANS than IDRISI. While the raster storage files for each system are sufficiently similar to allow fairly painless movement of data from one to another, the two GIS systems do not provide export/import facilities to directly convert vector data to each other's format. Consequently, a short utility program for translation of IDRISI vector format files to SPANS format vector files was written in the Clipper language, and compiled for the MS-DOS operating system to facilitate movement from IDRISI to SPANS. The program listing is included in appendix I, and a compiled binary may be obtained via anonymous ftp from the Department of Archaeology, University of Southampton (further details including internet address in appendix I).

4.4. Elevation data

4.4.1 Contour data

By far the greatest effort in the establishment of the gis experiments in chapters 5 and 6 was dedicated to the generation of the elevation overlays. This data contains within it the essential information which forms the basis for the analyses of visibility and intervisibility and the analyses of cost/perception in chapter 5 and the information from which several predictor variables used in the regression analyses of chapter 6 were derived.

Supply and extraction of data

Data for elevation was licensed from the Ordnance Survey. This data is vector in nature, and describes contour lines at 10m vertical intervals, digitised at a high horizontal resolution, with a claimed rms error of around ± 3m vertical height for derived products. This level of error is clearly perfectly acceptable for the derivation of topographic surfaces at resolutions of up to 20 or 30m resolution, for archaeological applications. Greater accuracy would almost certainly be entirely spurious for the applications intended here, as the land surface may have altered by two or three metres in some areas due to soil erosion or colluviation. Contour data was supplied in 20km square 'tiles', the origin of each tile being an even number of the OS national grid.

The data was supplied on 3.5 inch magnetic tape reels, coded in EBCDIC format text. This comprised contour data records in the Ordnance Survey's internal data transfer standard National Transfer Format V1.0. The tapes were mounted on the University of Southampton IBM3090 mainframe computer and read from there onto the IBM3090 disk. Utility software on the IBM3090 was used to translate the EBCDIC coding to ASCII standard coding and the resulting files were then downloaded via coloured book protocol file transfer to the project microcomputer. Although this is noted here for completeness, it is worth noting that this procedure is now obsolete (as is the IBM3090), because Ordnance Survey Contour data can now be supplied in a variety of more useful media than 3.5 inch reel-to-reel tape. The speed of technical change in information technology makes such events an occupational hazard.

Once the contents of the tapes had been transferred into files on the microcomputer, some extraction software was written to perform two tasks. First it was necessary to extract each 'tile' of data from the tape file. This was essential because contours were not fully georeferenced to

OS National Grid, but instead were recorded by tile. Each tile contained the NGR coordinate of the bottom left hand corner, and all contours were then recorded in terms of their position relative to that point.

A utility to extract the individual tiles was therefore written, using the Clipper programming language. This accepts arguments to indicate the source file, the 20km square to extract, and the target file. A complete listing of this program called GETCELL.EXE is given in appendix II. This extracts single 20km squares from the tape file and at the same time re-registers the coordinates into true NGR coordinate references. The program listing is included in appendix II, and a compiled binary may be obtained via anonymous ftp from the Department of Archaeology, University of Southampton (further details in appendix II).

Conversion to GIS formats

Neither of the gis systems available to the project contained import routines for NTF data; it was therefore necessary at an early stage to develop utility programs to perform this task. Two utility programs were therefore written to undertake this task although strong structural similarities in the read routines reduced the time and effort required by allowing some duplication of code.

The two resulting programs are called NTF2IDR, which translates NTF to IDRISI vector format and NTF2SPA which translates NTF to SPANS vector archive format. SPANS itself maintains an internal binary vector format which it uses for all processing and analysis and provides internal routines to translate between binary vector and vector archive formats. The program listing for both utilities are included in appendices III and IV, and compiled binary versions may be obtained via anonymous ftp from the Department of Archaeology, University of Southampton (further details in appendices III and IV).

4.4.2 Interpolation

Once imported into the GIS systems, the contour data were used to generate a raster elevation model. The IDRISI intercon routine was used for this task. Documentation is generally unsatisfactory for this routine, with no detailed description of the algorithm provided (this may be for commercial reasons). However some simple experiments suggested that the program implements an eight-connected weighted average algorithm for inter-contour interpolation (see chapter 2. Weighted distance algorithms are generally robust and reliable, although some slight problems became apparent at a later date (see below).

The intercon program requires that the contours be rasterised prior to the interpolation and produces best results under certain particular conditions. Firstly the resolution of the raster must be sufficient to prevent large numbers of contours occurring in the same raster units, and secondly the algorithm produces best results when around one-third of the grid cells are occupied with contour values. In order to provide these conditions, the interpolation of elevation models was undertaken at a very high resolution (10m raster units which required a resolution of 2000x2000 pixels for 20km square units). Although this provided the optimum conditions for the algorithm, this caused the interpolation to take a very long time and to produce extremely large elevation products. The resulting model was then generalised to a lower resolution according to need (80m or 50m depending on the experiment).

4.4.3 Smoothing and quality of elevation models

The interpolation results were visually inspected for smoothness and for any obviously deviant results, and were found to be acceptable. In places there was some slight evidence of straight facets in areas of the model with few contours, and all the elevation models showed a tendency to produce some low spine-like radiations in very flat areas. It seems likely that this is an inherent problem of the IDRISI GIS system and may be related to the use of an eight-connected algorithm for interpolation in conflict with the line rasterising algorithm which produces eight-connected lines.

These effects were not dramatic and, although visually unattractive, were not regarded as severely detrimental to the estimate of absolute elevation which the model represented. However, because such effects may have a severe effect on elevation products such as slope and elevation, their character was minimised through the use of mean filtering of the elevation surface prior to the derivation of slope, aspect and other secondary data.

4.5. Derived data

Several important landscape characteristics were derived from the elevation models but may still be regarded as 'primary' data sources for the analyses. Even in situations where the elevation itself may have had little impact on the organisation of archaeological material, it is distinctly possible that the slope of the terrain and the aspect of the land were significant factors. Both IDRISI and SPANS allow the calculation of slope and aspect from elevation values, producing identical results. Each was used at different times, depending on which was to hand.

4.5.1 Slope

The elevation models were smoothed by the repetitive application of a mean filter to minimise the effects of 'faceting' which cause the slope derivative to be discontinuous. Experimentation suggested that two applications of an unweighted mean filter greatly improved the continuity of the slope product, while the result was progressively less improved with subsequent passes. Two passes were therefore used on the original elevation models prior to the calculation of slope values as a percentage (zero for flat ground, 100% for vertical). Visual inspection of the result in comparison with maps suggested that the result then bore a close resemblance to the actual values, with slope for the study regions varying from flat to a maximum of around 20%.

Highly localised extreme values of slope which occur through the study regions are not well represented, although they could have been if the scale of the studies had been significantly smaller.

4.5.2 Aspect

Aspect calculations proved particularly susceptible to the effects of faceting within the models, and experimentation suggested that three passes with the mean filter was the ideal preparation for the elevation model. Aspect was then calculated in degrees (1 to 360) with zero representing on aspect calculated. A simplified aspect map was also produced by reclassification of the aspect product into sixteen directions (N, NNE, NE etc.).

4.5.3 Stream locations from runoff modelling

The smoothed elevation model was also used to model the water runoff within the regions, allowing the generation of the location of streams. Automatic generation of the streams is possible using hydrological modelling to measure the flow of water through each cell in turn. By setting a threshold value on this to represent the volume of water which must cross a cell in order for a stream to flow, the local drainage may be calculated. At its simplest, this method does not allow for the porosity of the land surface which reduces the initial runoff from each cell (porous areas will have less runoff than non-porous areas), and assumes that the percentage of rainwater which runs from each cell is identical. Although it is possible to model this, appropriate data was not available.

Although the location of streams and rivers was also digitised from current maps, the derivation of stream locations by this method was felt to be productive because the location of small streams in the study regions may have been greatly altered since the Neolithic by (a) extraction of water from the chalk aquifers for drinking water (b) agricultural irrigation and (c) the construction of pumping stations for the Kennet and Avon canal in the early 19th century. For each of these reasons, it is likely that small streams were more prevalent in the study areas in the Neolithic period than they are now.

Hydrological modelling of this type is not available in the IDRISI or SPANS systems, and was undertaken by transferring the elevation model to the GRASS GIS system, then running the hydrological model, then transferring the results back. Fortunately both IDRISI and GRASS provide ASCII raster data formats which allowed fairly uncomplicated transfer between these two systems. Due to memory restrictions, the derivation of hydrology was undertaken at a 50m resolution, using the GRASS 'r.watershed' module. Experiments with the threshold suggested that a value of around 5km maximum watershed size provided the best estimate of stream formation when compared with the number and extent of current streams.

4.6. Line data

Digtitising was initially undertaken using the AutoCAD program, which provides no facilities for generation of full topography data and no facilities for export of data to either of the GIS systems used in this project. As a result, it was again necessary to write software for export of data to a GIS format. This was undertaken using the internal version of the LISP programming language associated with AutoCAD and called AutoLISP. In addition to the export function, a series of utilities were also developed to undertake repetitive tasks using AutoCAD. The IDROUT function is listed in appendix V while the utilities are contained in a single AutoLISP file called DIGITOOL.LSP, which is listed in appendix VI.

4.6.1 Digitised rivers and streams

Both rivers and streams from the study area were digitised as line entities and imported into the GIS. Visual inspection and judgement was used in the separation of rivers from streams. If it was estimated that, during the Neolithic, a waterway would have been capable of carrying even a very small vessel such as canoe, then it was treated as a river.

4.6.2 Large monuments

Although most of the analyses in chapters 5 and 6 model archaeological monuments as points, larger monuments were digitised as line data. This data was used to indicate visually the locations of the main monuments of the study areas in the production of map output in chapters 5 and 6. Although this data was only used for display purposes in these studies, it is worth noting that the generation of derivative data and the analysis of the locations of the monuments within the study areas is possible with the data in this format.

4.7. Choropleth data

Two choropleth data themes, soil type and geological substrate, were required for the project, primarily for use in the predictive modelling experiments in chapter 6. Initially, the AutoCAD system was used for digitising area data, and the utilities described above were used for export. Late in the history of the project, a full-topography digitising program (the GIMMS DIGIT II program) became available. This was used to digitise the soil overlays with full arc-node topography and allowed the export of data directly from there to the SPANS vector archive format. Although of considerable value for the primary production of digital data, this program does not allow import of data supplied as, for example, dxf (drawing exchange format) files. As dxf is a *de facto* standard in which much spatial information is supplied and exchanged, the use of the AutoCAD utilities remains the most flexible method of generating GIS overlays for many tasks. All choropleth data overlays were stored as vector data within the GIS, and converted to raster (or SPANS quadtree) formats at an appropriate resolution for the analyses.

4.5.1 Soil

Soil types for the study areas were digitised from 1:250,000 scale soil maps. The scale is rather generalised for use in study areas of only around 20km square but was the only source of soil data available. Therefore it was carefully digitised to minimise error and included in the study. Some problems may be associated with the use of contemporary soil data in an archaeological context.

The work of, for example, Limbrey (1975) and Catt (1978) suggests that the distribution of present day soils does not correlate exactly with the 'natural' distribution of soils which occurred in the Neolithic period. There is also some controversy as to the nature of the changes which have taken place since the Neolithic. However, although changes in specific soil structure and tractability may have been quite profound, it seems likely that the general distribution of soil types, particularly with regard to pH (essentially the variation between alkaline rendsina soils and more acidic argillic soils) has not been dramatically altered because this characteristic depends largely on the source material. With reservations, therefore, the distributions of present day soils may be seen as being related to those of Neolithic soils.

However, the distribution of contemporary soils may be held to be valuable in the interpretation of archaeological remains even if this does not prove to be true. Contemporary soil type may hold important information as to the taphonomic processes which create the archaeological record: for example soil type will affect visibility of artefacts during fieldwalking and different soil chemistries will affect the survival of pottery within the soil. With this in mind, only the major soil types were digitised to reflect the distribution of soil characteristics. The subdivision of this into soil associations was regarded as too specific to be of any value.

4.5.2 Geology

Unlike soil, the surface geology of the study areas is unlikely to have changed dramatically since the Neolithic period. While valley sediments have certainly been deposited throughout the intervening time span, their distribution has obviously remained within the valleys, and the distribution of valley gravels and alluvial deposits visible today may be held to be a good indicator of the distributions of such deposits in the Neolithic at the scale of these case studies. The geological substrate will have had a substantial effect on the soil formation and on the drainage of the land. Consequently the surface geology may have affected the selection of locations for cultivation or settlement, and would clearly have affected the availability of raw material for building or flint working. Geological data was digitised from 1:50,000 Geological Survey sheets for both study areas. The main geological units (e.g. Upper Chalk, Lower Chalk, Clay-with-flints, Valley gravels) were digitised.

Chapter 5
GIS and the interpretation of monument locations

5.1. Introduction

The aim of this chapter is to examine the use of GIS in the interpretation of ceremonial archaeological monuments. A series of spatial analyses of archaeological monuments within the study areas will be undertaken in order to explore the application of existing GIS methods in a new context, and to develop new methods where necessary.

It is one of the central goals of this chapter to show how GIS technology may be applied to the types of ritual and ceremonial archaeological remains which characterise the Wessex area. This will be achieved through methods of spatial analysis which have been developed in response to archaeological questions, rather than imported wholesale from other arenas, and which take account of the need for methods which recognise the importance of perception in the representation of landscape. To this end the locations of the major earlier Neolithic monuments in two regions of the Wessex Chalklands are analysed here.

5.2 The location of causewayed enclosures

As discussed in the previous chapter, the most popular interpretation of the archaeological evidence from the study areas is that they represent early farming communities, characterised by Renfrew (1973) as 'chiefdoms'. This specific argument for chiefdoms in Wessex rests on a number of assumptions, firstly that the enclosure sites form the topmost tier of a hierarchy of earlier Neolithic monuments which are characterised in terms of the increasing labour investment which they represent; below the enclosures are the long barrows, each enclosure being associated with a group of barrows. This hierarchy of sites is supposed to be mirrored by nested territories: each enclosure is supposed to control a territory which comprises a series of smaller 'family' territories each of which contains a long barrow. Consequent to this, it is assumed that the enclosures share the common function of 'central places' as described by Smith (1965, 1971). Although a 'central place' in these terms need not necessarily imply geographic centrality, it is nonetheless reasonable to expect that the enclosure sites within Wessex might tend to be more physically central to the 'territories' they are associated with than physically marginal.

5.2.1. Centrality estimated with linear distance

The most obvious definition of a geographic central place is that it is a location such that the sum of the distances to all the other places within the region is minimised. In the context of the enclosures, this immediately raises a problem of what they might be expected to be central to.

In fact there is only one class of observation which can be thought of as broadly contemporary with the enclosure sites: the long barrows. The locations of the long barrows will be examined in some detail below, but here it is assumed that the distribution of the barrows broadly represents the distribution of the people within the region. This is not to say that each barrow corresponds with a settlement territory as assumed by Renfrew (1973), it has been suggested in chapter 3 that there may have been no permanent settlements in the way that many models of early farming communities have assumed and that the areas may actually represent rather more mobile communities, albeit within a defined geographic region. Additionally the barrow distribution may not be equated with settlement distribution because it may be that the places of ritual, in the form of barrows were deliberately located away from domestic activity - this will be examined further in the next chapter.

However, there are good reasons for using the barrows as a starting point for the exploration of the centrality of the enclosures. Firstly, some of the activities at the barrows and at some of the enclosures may have been closely connected. Bones from Windmill Hill enclosure and the West Kennet long barrow (Piggot 1962) show a complementary distribution of body parts and both enclosures and barrows were used for disposal of the dead, although in neither case was this necessarily the primary function of the site: other activities which took place at both included digging, deposition of animal bones and pottery and burning things. If the activities at enclosures and barrows are so closely connected, it would seem likely that their 'centrality' or otherwise should therefore be in relation to the barrows. With this in mind, centrality may be defined as that extent to which a location minimises the distance to all the places to which it is supposed to be central. From this, it is possible to measure the degree of centrality of any part of the landscape with respect to a particular group of monuments. In this case, we wish to obtain the geographic centrality of the study areas with respect to the long barrows.

To do this, distance overlays were generated for each of the long barrows. The result of this is a set of maps each in which each location contains the distance in metres from a long barrow. Once a distance overlay is obtained for each barrow, the entire series of maps may be summed to obtain the total distance to long barrows for each location in the landscape. It follows that this is now a representation of the centrality of each location with respect to the long barrows. This operation was performed for the Avebury region and the Stonehenge region study areas, with the results shown in figures 5.1 and 5.2.

The results are interesting. Compared with the other major Neolithic monuments, particularly those which are generally later than the enclosures and the barrows, the causewayed enclosure sites are less 'central' in both the Avebury area and in the Stonehenge area. Around Avebury, both the Avebury Henge monument and Silbury Hill are far more centrally placed than Windmill Hill, although Windmill Hill is itself obviously more central than the two enclosures at Rybury and Knap Hill. On Salisbury Plain also, the enclosure at Robin Hood's Ball seems to be distinctly marginal compared with the Cursus and Stonehenge itself, although the complex at Durrington Walls is comparably placed in these terms. From this we may make the claim that, in geographic terms at least, the later Neolithic monuments show a far greater tendency to be 'central' to the long barrows than do the enclosure sites in both regions.

5.2.2. Centrality and 'cost distance'

Although the results of this experiment are interesting, the linear distance measurement used to generate the central tendencies of the two study areas relies wholly on an abstract Cartesian measure of distance, and takes no account of the perceived or experienced distance between locations within the landscape. The distance which separates any two locations is not experienced in terms of linear distance, but as time taken to travel between locations or as effort expended in travel between places. Thus two locations which are separated by a flat, easily traversed area will seem to be closer than two locations separated by highly variable terrain. Crossing variable terrain involves climbing and descending slopes which takes longer and requires more energy. This is particularly the case for transportation of goods and materials within the landscape, and must have been a major influence on the 'cognitive maps' which members of Neolithic communities had of their landscape. If the causewayed enclosure sites can be interpreted as central gathering places, as Renfrew's (1973) model suggests, then it is reasonable to assume that they may be situated so as to minimise not the absolute distance to the other locations within the landscape, but the experienced distance.

A modification to the procedure may be made to estimate this. Instead of generating a distance map for each barrow, a *cost distance* map is generated. This modifies the distance calculation according to the cost of traversing each location in the landscape, and requires that a *friction surface* be introduced to represent the cost of travelling across the surface. A short series of experimental 'timed walks' within the Avebury area, suggested that for the same effort expended, speed of travel is roughly proportional to the percentage slope divided by three. This implies that traversing a one-in ten incline is three times slower than travelling on flat ground, and a one-in-five incline reduces speed to about one sixth of normal. On this basis, the elevation map for each region was used to derive a slope map, and this in turn was used to derive a friction surface which varied from 1 (for flat ground) to about 6 for the steepest slopes in the regions (about 20%). The friction surfaces were then used to generate cost surface maps for each barrow, which were summed as in the experiment above. The results of this experiment are shown for each region in figure 5.3 and 5.4.

For the Stonehenge area, the cumulative cost distance map still suggests again that, compared with the later Cursus monuments and Stonehenge itself, the enclosure site is distinctly marginal to the geographic centre of the area. Similarly, for the Avebury region, the two enclosures at

Knap Hill and Rybury appear to be on the periphery of the area, although Windmill Hill is again far more centrally placed than either of these. As for the linear distance experiment, monuments which date from the later peiod, Avebury itself and Silbury Hill, both seem to be far more central to the region than any of the enclosures.

5.2.3. Rivers and cost calculations

One of the differences between the linear and cost-distance models of geographic centrality presented above is that the cost maps suggest that some valley areas are more 'central' than was suggested by the Cartesian distance versions. This is because the valley bottoms are characterised by flat floodplains which therefore produce low values for traversal cost. These low values may be misleading, however, as the valley bottoms might have been boggy, wet and generally less penetrable than the uplands, although for the Avebury region at least, the work of Evans *et al* (1988a) & Mount (1991) suggests that the Kennet valley was not particularly prone to flooding during this period. This does, however, suggest another misleading element of the maps: that they take no account of the effort involved in crossing rivers or of the assistance afforded to travel by water transport such as canoes. One method of representing both of these aspects of river networks (Limp *pers comm*) is to generate a line of cells three wide for each river, in which an 'investment cost' is stored in the outer two bands to represent the effort of locating a boat, and the central line of cells represents a corridor of minimal cost to represent the cost of travel once this 'investment' has been made. Although this is an elegant proposal, it effectively side-steps the real issue which is one of *scale*.

At a larger, inter-regional scale, the rivers may have had a very positive influence on movement. For example it is possible that the River Kennet acted as a transport corridor to and from the Thames valley, and that any of the rivers which flow from Salisbury Plain could have acted as transport corridors southwards into the river Avon and from there to the coast at what is now Bournemouth. There seems little doubt that the larger rivers, such as the Avon below Salisbury and the lower reaches of the Kennet, would always have been of sufficient size to have allowed small vessels, but how far upstream it would have been possible to navigate even by canoe is impossible to determine because of alterations to the hydrology of the areas. Today, for example, it is not possible to take even a canoe much beyond West Overton on the Kennet and, because the Kennet is a winterbourne, even this is unpredictable. This was probably not the case even in historic times, however, because the river carried substantially more water and flowed further up the valley before the construction of pumping stations for the Kennet and Avon Canal at the start of the 19th Century. For this reason it would be extremely difficult to model the transport viability of such rivers, and in any case there is no firm evidence that they were in fact used in this way.

The scale of these experiments, however, is intra-regional: concerned with local travel, probably on foot. In this context, the rivers clearly represent an obstacle rather than a corridor. None of the rivers of the region are too large to wade or swim safely across, and many of the rivers could have had fording places or even bridges. Unfortunately there is no way of knowing where such conduits were, or even if they existed at all and consequently it is not possible to accurately model the effects of the rivers in these terms.

5.2.4. Windmill Hill, Knap Hill and Rybury

Another problem with the interpretation of the enclosure sites as the top tier of a hierarchy of sites, is that while the barrow groups on the Dorset Ridgeway, Cranbourne Chase and around Stonehenge may seem to be associated with a single enclosure, the Avebury group of barrows seems to be associated with not one but three enclosure sites. The largest of these, at Windmill Hill, has received extensive attention from archaeologists, including numerous excavations, and is usually considered to be the 'central place' of the earlier Neolithic community of the region later served by the Avebury Henge monument. However, there are also enclosures at Knap Hill and Rybury on the southern scarp of the chalk for which an explanation is not provided in these terms. This is a problem recognised by Renfrew (1973), who noted that the enclosure at Knap Hill was 'anomalous' because it does not correspond with a later Henge monument, and assumed that Rybury was associated with a separate region, based in the Vale of Pewsey to the south and later served by the Marden Henge. However, there is no group of long barrows in the Vale of Pewsey to support this assumption, as Renfrew shows in his own figures, and the existence of the later Henge monument at Marden seems far better interpreted as the extension of activity during the later Neolithic than, perversely, as evidence of earlier Neolithic activity.

Another possibility is that the three enclosure sites were never in use at the same time, and that they each successively represented the single 'central place' of the model. This is not supported by the dating evidence from Windmill Hill, however, which seems to have been in use for a long period of time, successively improved and enlarged (Smith 1965, 1971). It seems most likely, therefore, that all three of the enclosure sites form part of the Avebury complex of monuments and, while the possibility that they are external to the region must be entertained, their relationship with each other must be of considerable interest.

It is possible to explore the geographic relationship between these three sites through a similar use of cost surface analysis. Using the same assumptions about the relationship between slope and walking time, cost surfaces may be generated for each of the enclosures in turn (shown in figures 5.5, 5.6 and 5.7). The cost in these examples is directly correlated with time taken to walk to the monument, and it is therefore possible to 'calibrate' the surfaces in terms of time. This has been done using the same field observations as above, and the maps show 'time contours' at intervals of one hour return-to-base time, in other words it takes approximately one

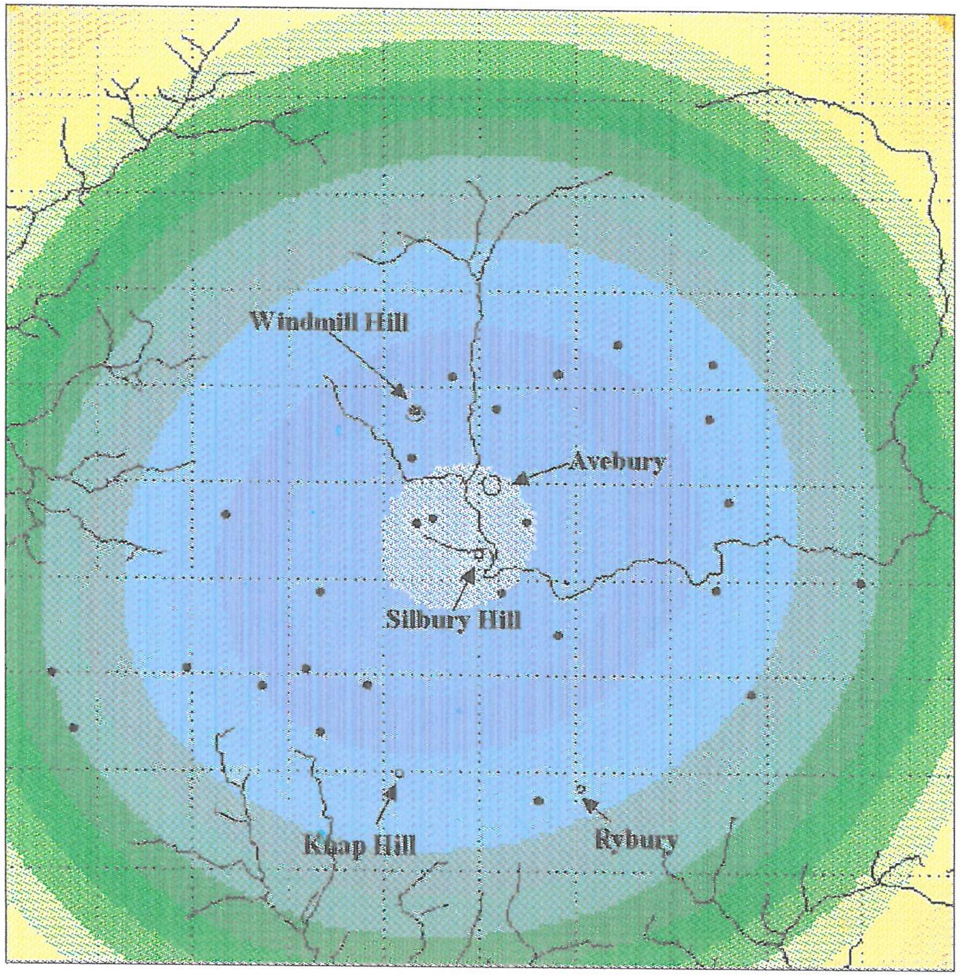


Figure 5.1. Centrality to the Avebury long barrow group as represented by cumulative linear distance.

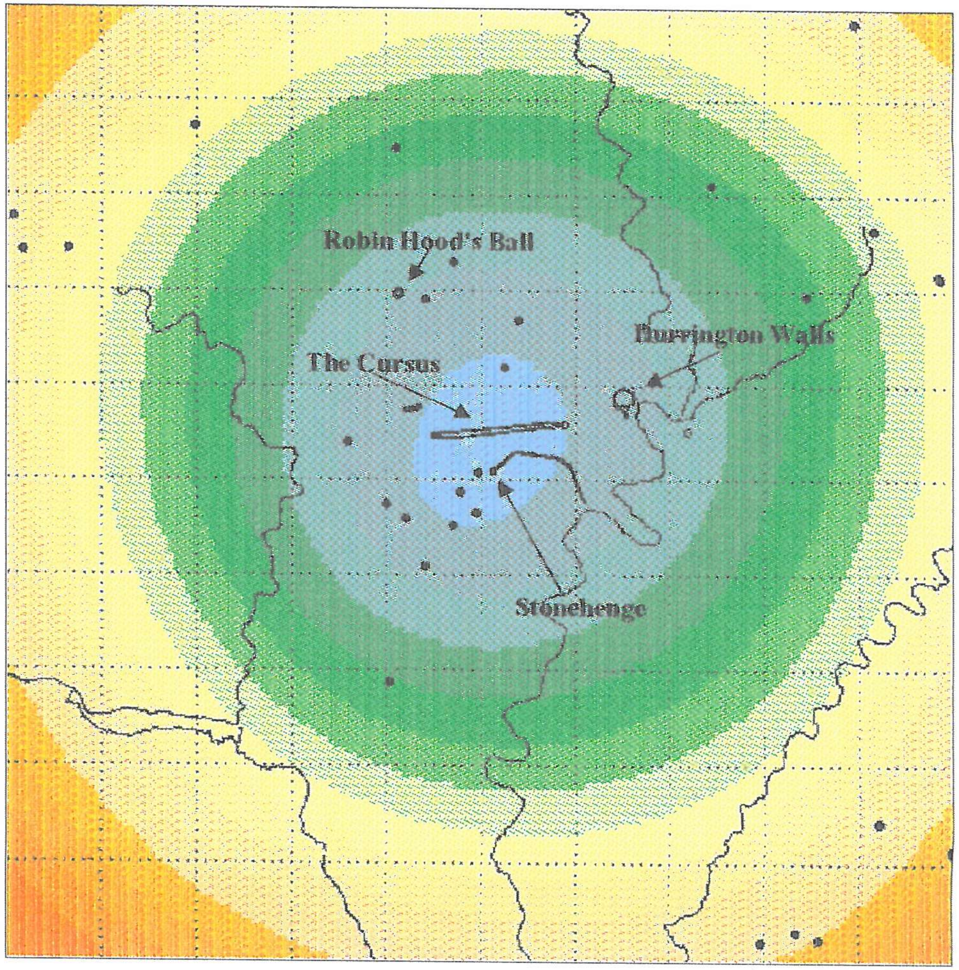


Figure 5.2. Centrality to the Stonehenge long barrow group as represented by cumulative linear distance.

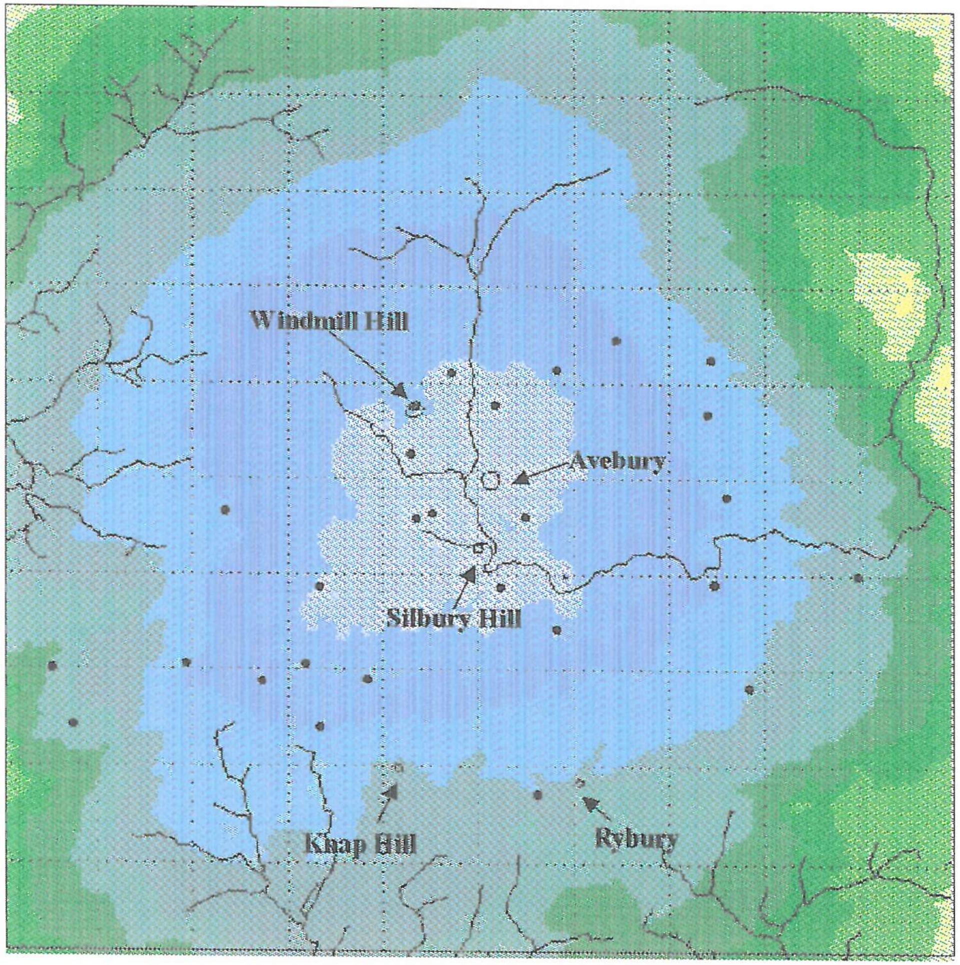


Figure 5.3. Cumulative cost distance to the Avebury long barrows.

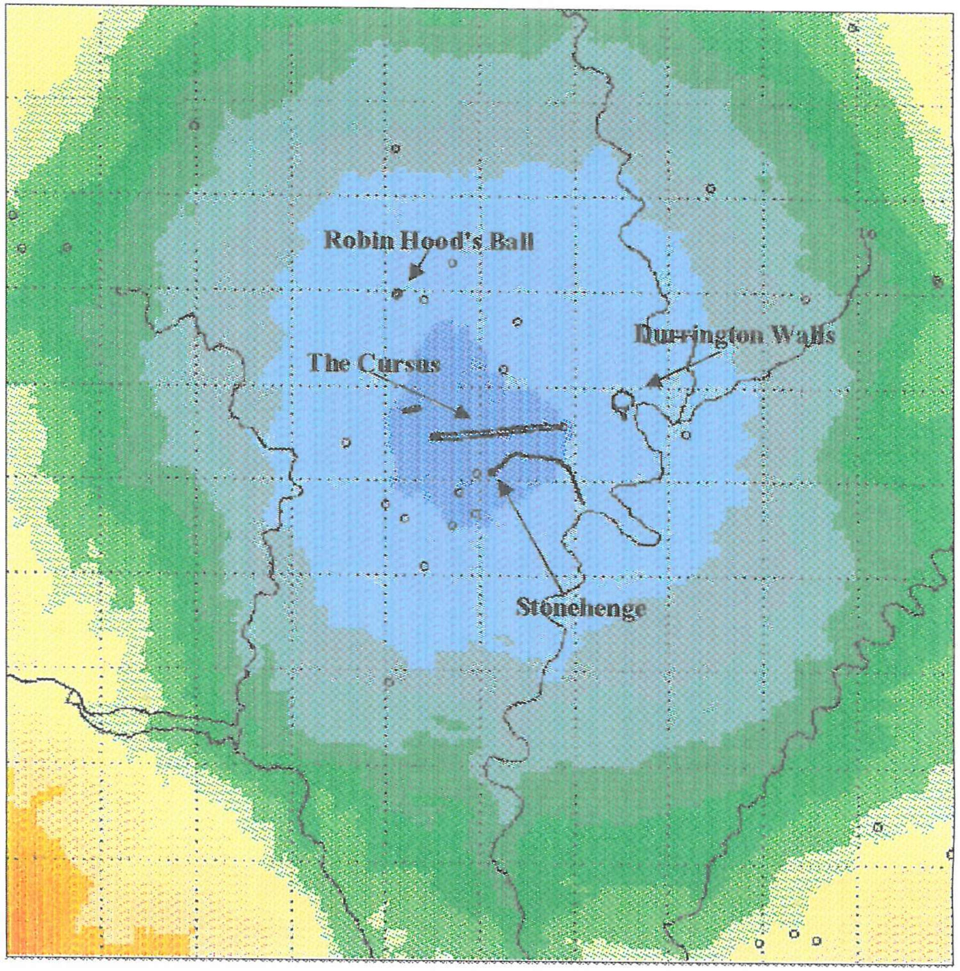


Figure 5.4. Cumulative cost distance to the Stonehenge long barrows.

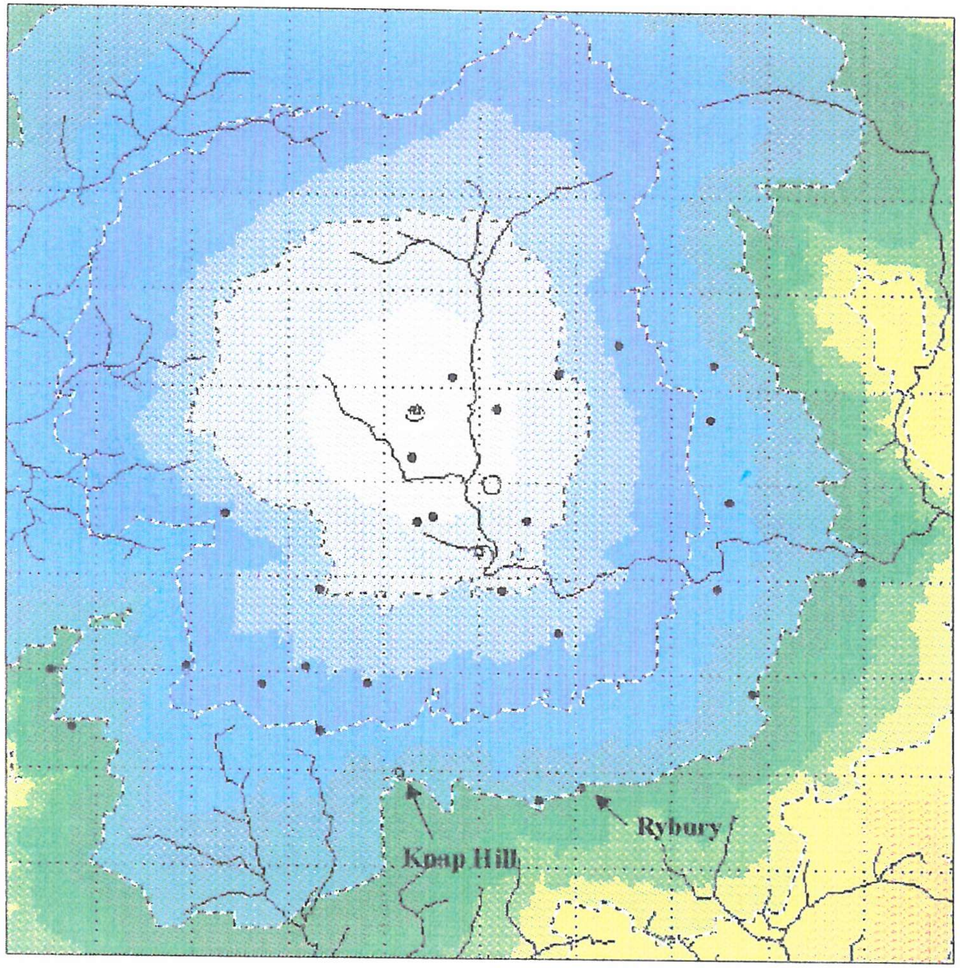


Figure 5.5. Cost distance (time) from Windmill Hill, dashed lines represent 'time contours' at one hour return trip times.

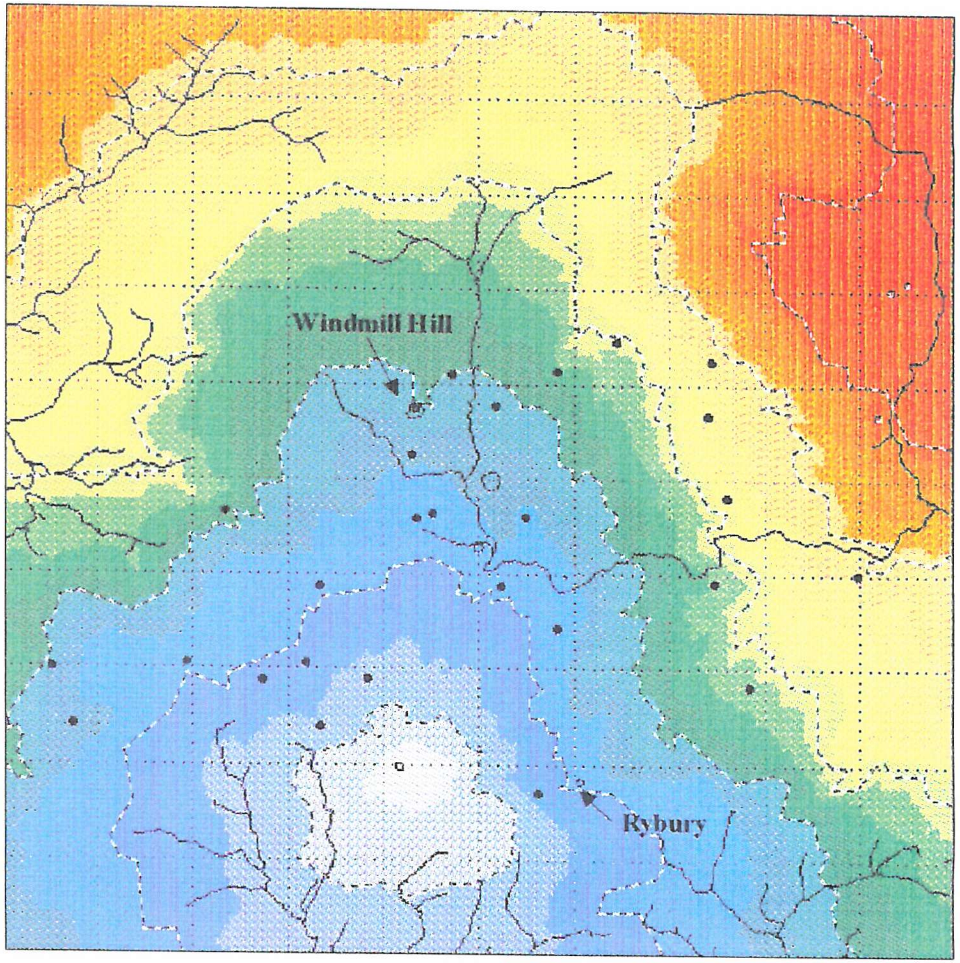


Figure 5.6. Cost distance (time) from the enclosure site at Rybury, dashed lines represent 'time contours' at one hour return trip times.

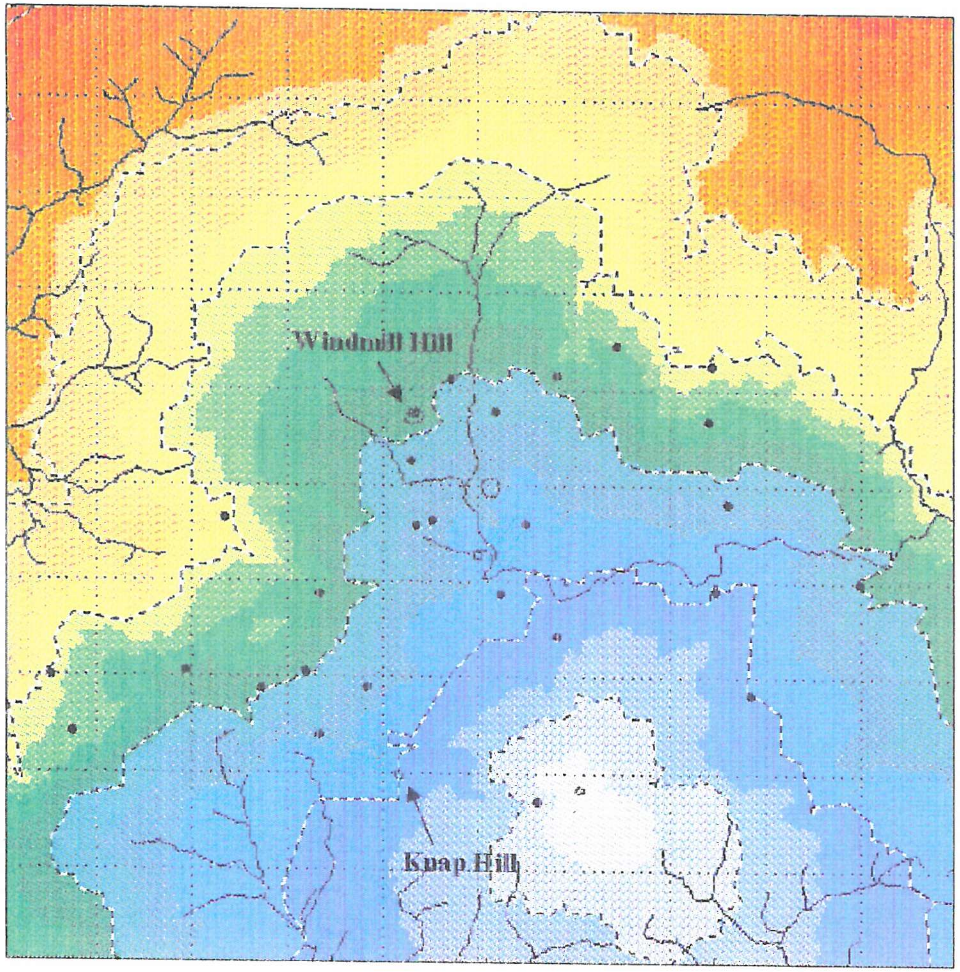


Figure 5.7. Cost distance (time) from the enclosure at Knap Hill, dashed lines represent 'time contours' at one hour return trip times.

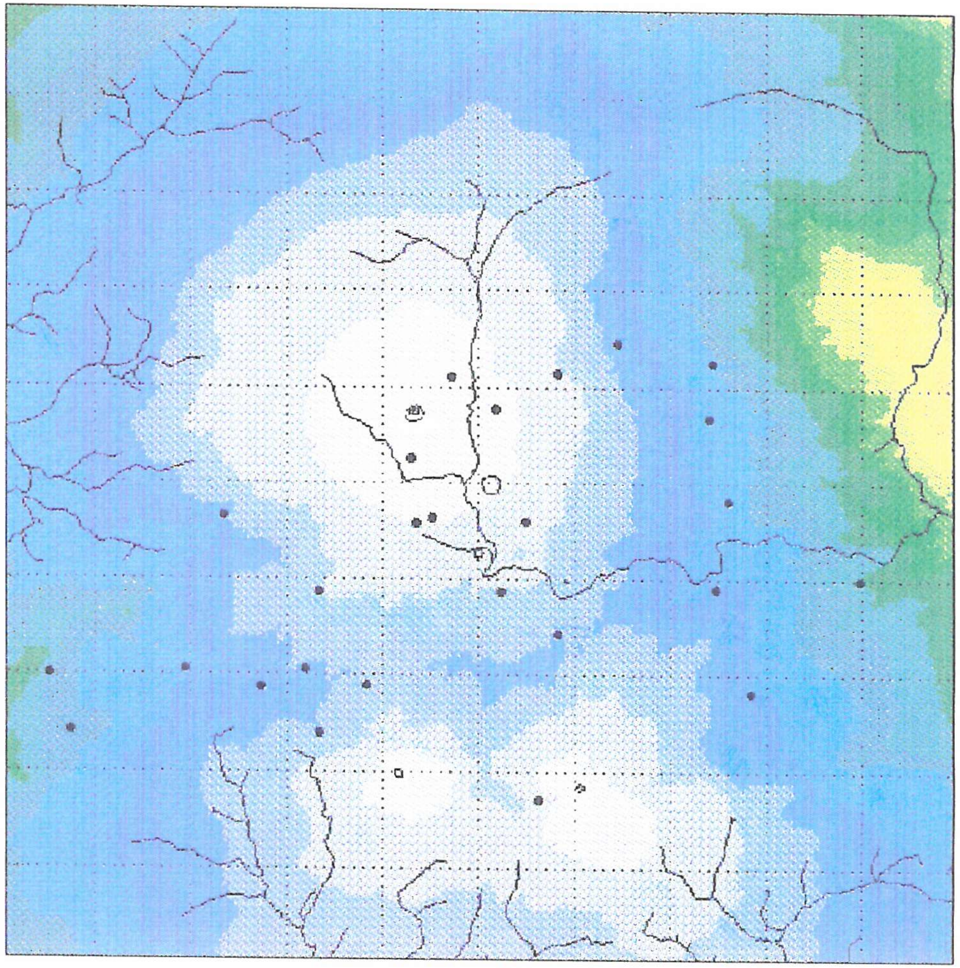


Figure 5.8. Minimum cost distance (time) to any of the three causewayed enclosure sites of the Avebury region.

hour to start at an enclosure, walk to the first contour and then return to the enclosure. Obviously this also means that it take about one hour to start walking from anywhere on the contour to the enclosure and back again.

The isotropic nature of the analysis means that a one hour return-to-base contour need not correlate with a half hour trip in each direction (the terrain may be downhill in one direction, and uphill on the way back) although experience of walking in the region suggests that this is a reasonable estimate. It is also possible to generate a 'gravity map' for the three sites by combining the three maps with a boolean map algebra operation to obtain the minimum value for each location. This then represents the time taken to walk to the nearest enclosure, nearest being defined in time not distance terms but in time. The result of this operation is shown in figure 5.8.

The individual 'time contour' maps shown in figures 5.5, 5.6 and 5.7, can be regarded as suggesting the geographic 'catchment' of the site concerned and, if the sites do in fact represent the central places of earlier Neolithic territories, then these might be interpreted as the areas over which the sites exert some control. This is not to claim, however, that the catchments may be interpreted as they would be in site catchment analysis, where the characteristics of a site are analysed in terms of the economic resources available within its catchment. It is far from clear that the enclosure sites represent permanent dwelling places, although shallow pits containing seemingly domestic refuse were found beneath the bank at Windmill Hill. Therefore the assumption of site catchment analysis, as used by in this context by Barker and Webley (1978) for example, that the sites are economic bases is not valid. The area of land from which they may easily be approached must be regarded not as an economic catchment but rather as indicative of the area which the site may have served in a ceremonial role. Although this is obviously simplistic, the maps might reveal some of the different character of the places which were chosen for the construction of these monuments.

Figures 5.5, 5.6 and 5.7 do reveal a difference between the 'catchment' of Windmill Hill, and that of the two southern enclosures, a difference which is clarified by figure 5.8. Windmill Hill, the largest and most long-lived of the three sites, is situated so that it 'serves' a far larger area of the landscape and a larger number of long barrows. The one hour time contour for Windmill Hill encloses an area roughly two to three times the size of the one-hour contour for Rybury and for Knap Hill. Interestingly, the combined one-hour catchment for Rybury and Knap Hill is around the same as that of Windmill Hill, a fact which mirrors the sizes of the sites themselves. Windmill Hill is around 1200ft at its greatest diameter, Knap Hill and Rybury are both around 600ft. It is also worth noting that, while *all* the earlier Neolithic barrows and monuments of the Avebury region are within about a three hour return trip from Windmill Hill, the same cannot be said of either Knap Hill or Rybury. The relationship between the sites also shows an interesting pattern. Both Knap Hill and Rybury fall almost exactly on the three hour time contour from Windmill Hill - the same distance as the farthest long barrows - while they are themselves about two hours apart (this is 'return trip' time so represents perhaps an hours walk from one to another).

Perhaps the most interesting image, however, is figure 5.8. This shows clearly how the barrows do not cluster around the enclosures themselves, but instead occur in a broad band *between* Windmill Hill and the southern enclosures. While the 'territory' of Windmill Hill extends northwards and westwards rather more than it does south and east, there are no barrows in this area. The same can be seen to be true for both Rybury and Knap Hill, whose 'territories' extend southwards into the Vale while all the barrows occur to the north of them. While the later Neolithic sites at Avebury and Silbury Hill might therefore be argued as 'central' to the group of ceremonial monuments, all the evidence is that the enclosure sites are in fact entirely marginal to the long barrows, standing not at the *centre* but at the *edges* of the man-made landscape.

5.2.5. Interpretation

It must be accepted that there is a distinction between physical and social centrality, and that the physical inaccessibility of some of these sites is not wholly incompatible with their function as ritual or social centres for associated territories. But physical centrality is clearly not a consistent feature of these sites and it is perhaps time to seek alternative explanations for their location. Some enclosures may indeed serve the function which Renfrew claims, others clearly do not. It is worth noting that the enclosures are not particularly uniform in construction or size, and that the variation between sites and sub-regions has tended to be subsumed in the search for a generalised explanation of social organisation in the Neolithic.

There is huge variation in these sites. Their size, for example, varies from less than 10 hectares at Robin Hood's Ball to Windmill Hill, Hambledon and Maiden Castle all of which cover more than 17 hectares (Smith 1971). Their form also varies considerably: Knap Hill (Connah 1961) comprises just one interrupted ditch, Robin Hood's Ball (Thomas 1963) and Maiden Castle (Wheeler 1943) have two concentric ditches, while Windmill Hill (Smith 1965) has three. Hambledon (Mercer 1980) has two enclosure ditches and outworks. Some enclosures were probably revetted (Windmill Hill and Robin Hood's Ball) while at others, such as Coombe Hill, evidence of revetment was sought but not found. In truth, the only thing these sites have in common is their construction method of segmented ditches - most probably reflecting their construction by gangs or groups of people. To interpret them as a group with a shared function seems then to be the equivalent of treating all buildings made from bricks as serving the same purpose.

It is not surprising, therefore, that the sites are geographically dissimilar also. It is not necessary to invent new earlier Neolithic regions from later Neolithic evidence in order to accommodate the three enclosures of the Avebury region within an explanation. It must be accepted that Windmill Hill, while far less central than Avebury or Silbury, could have served an equivalent social function during the earlier Neolithic as the Avebury Henge did during the later Neolithic. The change in location from Windmill Hill to the Valley location of Avebury might be explained as just that - a movement from hill to valley. However, if the location of the main 'central' enclosure at Windmill Hill was simply a compromised centrality to accommodate the desire for a hilltop location, then why was the enclosure not built on Waden Hill which is a wholly central place? If the three Avebury sites, and Robin Hood's ball have anything in common, it is that they are situated on the margins of the group of monuments. Two explanations might be put forward for this geographic relationship.

Firstly, we may persist with the belief that earlier Neolithic territories were centred on the enclosures themselves. If this is the case, then we must accept that the long barrow monuments, used by Renfrew to suggest 'family territories', are no such thing, but are located between the Neolithic communities. Some solace for this belief might come from the later development of a henge at Marden, in the Vale of Pewsey which might be interpreted as evidence of such a community which was centred in the Vale itself, perhaps related to the Rybury enclosure. It was argued above, however, that the existence of later Neolithic activity in the Vale cannot be taken to suggest earlier Neolithic activity in the absence of any other evidence. The henge at Marden could rather better be explained as a later Neolithic expansion. The alternative is to accept that the distribution of the long barrows does represent the distribution of use of the Avebury area in the earlier Neolithic, and that the division of the landscape into 'ritual' and 'domestic' areas was at a smaller scale than the above requires if it existed at all. The existence of cultivation marks beneath some long barrows also suggests that there was no rigid distinction between places suitable for agriculture and places suitable for monumentalisation.

A better explanation of the enclosure sites may then be that they are places of passage between a physical and social context which is 'outside' the Neolithic community and one which is 'inside' it. Such places might have been locations for activities which involved the passage of raw materials or goods from other geographic areas into the community. It is possible that they were places where ritually 'dangerous' activities, associated with non-local people and objects took place away from the centre of the community.

It may be far better, therefore, to regard the enclosures firstly as individual sites rather than as a coherent group of sites: the differences in activity at each site and of size and form are also reflected in their geographic position. However, if the sites do have any common features other than their segmented construction method, it may be that were not 'central' to territories in the social sense that Renfrew (1973) implied or the economic way that Barker & Webley (1978) intended, but that they functioned in as *marginal places* rather than *central* ones.

5.3. The location of long barrows

5.3.1 Spatial organisation of long barrows

As discussed above, one of the claims of the traditional interpretation of the monument clusters of the chalklands, proposed by Renfrew (1973), is that the long barrow monuments represent some form of territorial marker, a claim Renfrew supported by the generation of Theissen polygons around the barrows. This is a natural interpretation of the function of the monuments if the starting point for their interpretation is that:

(a) the barrows were built by sedentary farmers, the need for territorial markers being linked to the need to appropriate land for cereal cultivation,

(b) the barrows directly correlate with one social tier (usually assumed to be family units) of the 'chiefdom' which consisted of a hierarchically arranged social formation, and

(c) the monuments all served the same function simultaneously.

None of these three assumptions of the interpretation can be reasonably sustained. Chapter 3 developed the argument that the Neolithic communities of the chalklands may have been far more mobile, and far less reliant on cultivation than this model requires while the assumption that the social organisation could be characterised in social-evolutionary terms, with small (possibly family) territories within a larger 'chiefdom' unit, was also seriously questioned. The assumption that the monuments can not only be treated as contemporaneous, but that they served the same function simultaneously deserves some closer examination.

Renfrew rightly observed that an important assumption of his model was that the barrows '*must have been functioning in some sense - although not necessarily for burial, at the same time - no doubt at the end of the early Neolithic period*' (p544), which is difficult to refute. Obviously the barrows did not vanish after their immediate currency had expired. Rather they remained, large and imposing, as a visual reminder of the period during which they were used and as a signal of the status of the location. This seems to support Renfrew's contention that the barrows may therefore be treated as a group within the model. However, the statement is something of a half-truth in that the model actually requires that the barrows must not only have served *a function* at the end of the earlier Neolithic period but that they in fact served *the same function* at the same time. In other words, the model assumes that, at the end of the earlier Neolithic, barrows which were still being used actively for burial served the same 'territorial marker' function as did barrows which may have been disused for over 1000 years.

It follows that the barrows may be treated as a group in that they represent a group of locations which were, over a long period of time, selected for monumentalisation in the form of long barrow construction. These locations must then have remained important places after the function of burial had ceased. However, any explanation of the location of the barrows must at the same time consider the temporal relationships between them.

5.3.2 Visibility and GIS

One possibly important feature of the barrows, suggested by the size and prominence of the monuments, is their visibility. The visibility of prehistoric monuments is a subject which has received some attention within the literature; however claims about the relationships of visibility between monuments and landscape features, or between monuments and other monuments, or between monuments and astronomical alignments, have always proved extremely difficult to substantiate. Nevertheless, there have been some notable observations which may be of relevance to the interpretation of the barrows.

Devereaux (1991), has recently drawn attention to the relationships of visibility between the profile of Silbury Hill and the visible horizon, suggesting that the visual appearance of the monument was a factor in its location and design. More interesting, perhaps, is the relationship observed by Barrett (Barrett *et al* 1991), between the Dorset Cursus and earlier long barrows. The alignment of the Cursus monument seems to have deliberately deviated to enclose a long barrow on Gussage Cow Down. This may have been done so that, from the original terminal of the Cursus at Bottlebush Down, this long barrow then appears to form a frame for the midwinter sunset. When the Cursus was subsequently extended to the Martin Down Terminal, the Cursus again took account of a long barrow by incorporating the Pentridge 19 barrow into the west bank of the new monument.

Although GIS is not particularly valuable in the analysis of this type of very specific relationship, visibility is something highly amenable to study with the aid of GIS technology and chapter 2 has already explained how binary line-of-sight images can be derived for specific locations from an elevation model. These line of sight images have already found some application within archaeology: Gaffney and Stancic, for example, investigating the Greek period of the Island of Hvar, generated a viewshed map for the watchtower at Maslonovik, Hvar. This then demonstrated that a similar watchtower at Tor would have been visible from Maslinovik, which would in turn have been able to pass warnings to the town of Pharos and their result supported the assumption that such towers formed *`an integral system connected to the town and Pharos whereby watch was kept for any approaching danger'* (Gaffney and Stancic, 1991 p.78).

Viewshed maps were also suggested by Ruggles *et al* (1993) as part of a method for investigating the locations of the short standing stone rows of the island of Mull. This method suggested that viewshed maps for each of the standing stone row sites should be combined to create a binary *multiple viewshed map* consisting of the logical union of the individual maps. From this, prominent landscape features on which the stones may have been aligned could be identified, and viewshed maps generated from these locations. Finally a count could be made of the number of stone row sites falling within these landscape features, and *'the landscape features which best explain the observed placing of the stone rows are those for which this number is greatest'* (p127). An example of this method applied is not provided by the authors.

More interesting, perhaps, is the claim by Lock and Harris (forthcoming) that viewshed maps from long barrows of the Danebury region consistently fail to overlap: *'in almost every case the Barrows appear to have been located so as to exclude other Barrows from its viewshed'* and this is interpreted as support for the idea that the barrows are territorial markers, as proposed by Renfrew.

5.3.3 Long barrow visibility

In order to investigate the relationships within the long barrows of the study areas, a line-of-sight (LOS) map was generated for each barrow using the IDRISI viewshed program. These maps represent (depending on which way you consider it) either every cell in the landscape which could (theoretically) be seen from the barrow, or every cell in the landscape from which the barrow can be seen. An example of the result of this operation is given in figure 5.9.

The area of each of these maps was then calculated to provide a sample of long barrow viewshed areas (there are 26 barrows in the Avebury area, and 31 from the Stonehenge area). It was then necessary to generate a series of random 'barrows' with which to compare these samples. These random samples, each of thirty locations, were drawn from their respective study areas and treated identically to the barrows themselves. This procedure then provided two samples from each population (a population being defined as the population of all the possible viewsheds within a region), from which it should be possible to establish statistically whether or not the locations of the barrows significantly differ with respect to the areas of their viewsheds from the rest of the landscape.

Many statistical tests assume normally distributed data and this is rarely the case when dealing with spatial phenomena. Because normality of either the populations or the sample could not be assumed, a Mann-Whitney U test (Downie & Heath 1965) was selected to test the null hypothesis that the long barrow samples were drawn from the same population as the random samples with respect to their viewshed areas. Usefully, this test does not require the samples to be of the same size, although in this case the size of the random sample could have obviously been controlled to be the same as the barrow sample. These values are used to compute the U statistics.

For two samples of size N1 and N2, the rank order R of each value within the *combined* samples is obtained. Values for U_1 and U_2 can then be calculated from equations 5.1 and 5.2 respectively.

$$U_1 = N_1N_2 + \frac{N_1(N_1+1)}{2} - \Sigma R_x \qquad 5.1$$

$$U_2 = N_1N_2 + \frac{N_2(N_2+1)}{2} - \Sigma R_y \qquad 5.2$$

Barrows (x) m²	Random (y) m²	Rx	Ry
74656000	74016000	56	55
70656000	51468800	54	51
68793600	49030400	53	50
53606400	38694400	52	46
46784000	38233600	49	45
42764800	37740800	48	44
39628800	29260800	47	36
32518400	28825600	43	35
32352000	28294400	42	34
31993600	26144000	41	32
31532800	25676800	40	31
31110400	25484800	39	30
30201600	23673600	38	28
29683200	23097600	37	27
26528000	22406400	33	26
25017600	21766400	29	25
21120000	21350400	23	24
18009600	19148800	20	22
13849600	18233600	17	21
11942400	17792000	14	19
11520000	14560000	13	18
9728000	13734400	9	16
9689600	13395200	8	15
7078400	10419200	4	12
3116800	10252800	2	11
2169600	10118400	1	10
	9248000		7
	7961600		6
	7430400		5
	5862400		3
26	30	812	784

Table 5.1. Mann-Whitney U-test for Avebury sub-region long barrows. The areas (in m²) of the LOS maps for 26 long barrows (x) are compared to 30 random points (y) by generating the rank order within the entire series of each observation (Rx and Ry).

When either N_1 or N_2 is greater than 20, the U statistic is considered to be normally distributed. A z-score can then be calculated from equation 5.3 (either U can be used as only the sign of the z, and thus not its interpretation in terms of significance, is affected by the choice of U).

$$z = \frac{U_1 - (N_1 N_2/2)}{\sqrt{[N_1 N_2 (N_1 + N_2 + 1)/12}} \qquad 5.3$$

In a two-tailed test, for the confidence level α=0.05, z must be greater than 1.96 to reject H_o.

Barrows (x) m²	Random (y) m²	Rx	Ry
131424000	67366400	61	56
115577600	59590400	60	54
82188800	58828800	59	53
78380800	52416000	58	48
69568000	42841600	57	43
63532800	39808000	55	41
55059200	35859200	52	37
55040000	31974400	51	34
54720000	29420800	50	33
54105600	29004800	49	31
50841600	23750400	47	30
46342400	22707200	46	27
45779200	21152000	45	25
42892800	14867200	44	23
40755200	14835200	42	22
39200000	11840000	40	20
37676800	11475200	39	19
37107200	10771200	38	18
33657600	10067200	36	17
33228800	7424000	35	16
29049600	7353600	32	15
23571200	6272000	29	14
22854400	5836800	28	13
21734400	5324800	26	12
18054400	5126400	24	11
12697600	4108800	21	10
3244800	3980800	6	9
3238400	3923200	5	8
3020800	3488000	4	7
2617600	2675200	2	3
1190400		1	

Table 5.2. Mann-Whitney U-test for Stonehenge sub-region long barrows. The areas (in m²) of the LOS maps for 31 long barrows (x) are compared to those for 30 random points (y) by generating the rank order within the entire series, of each observation (R_x and R_y).

Table 5.1 shows the resulting table for the Avebury series. In this case U_1=319 and U_2=317 generating a z-score of 1.17. This is not significant at the 0.05 level so the test does not allow the rejection of H_o. In other words at the specified significance level, the areas of the LOS maps for the Avebury long barrows are not different from those of a random sample of points drawn from the same geographical area. For the Salisbury series (table 5.5), however, U_1=284 and U_2=646 generating a z-score of 2.61. This is significant at the 0.05 level so the test allows the rejection of H_o at the 0.05 level. This shows a significant difference (at the specified level) from the random series with respect to LOS areas. In fact, a visual inspection of the values reveals that the Stonehenge area long barrows tend to have larger viewsheds than the random sample.

At this point a note of caution should be inserted. Apart from the obvious scepticism which should apply to any unrepeated statistical test (each LOS map takes some ten to twenty minutes to generate, and time so constraints have so far prevented repetition of the experiment), *association* of long barrows with positions of high visibility (which is indicated by this experimental result) should be carefully distinguished from *causation:* this is discussed further below.

Lines of Sight	Population Area (km^2)	Area %	Cum. %	Sample Cases	Cases %	Cum %	D
0	37.19	37.19	58.286	7	26.9	26.9	10.3
1	23.73	60.92	37.187	2	7.7	34.6	26.3
2	12.51	73.42	19.601	4	15.4	50	23.4
3	8.9	82.32	13.948	4	15.4	65.4	16.9
4	6.33	88.65	9.914	3	11.5	76.9	11.7
5	4.06	92.71	6.369	1	3.8	80.8	11.9
6	3.01	95.73	4.723	3	11.5	92.3	3.4
7	1.8	97.53	2.824	1	3.8	96.2	1.4
8	0.92	98.45	1.437	0	0	96.2	2.3
9	0.65	99.09	1.012	0	0	96.2	2.9
10	0.38	99.47	0.591	0	0	96.2	3.3
11	0.3	99.77	0.476	0	0	96.2	3.6
12	0.16	99.93	0.247	1	3.8	100	0.07
13	0.05	99.98	0.084	0	0	100	0.02
14	0.01	100	0.02	0	0	100	0
15	0	100	0.005	0	0	100	0

Table 5.3. Kolmogorov-Smirnov Test for Avebury sub-region long barrow LOS. Dmax is 0.20, d is 0.26 at the 0.05 level.

5.3.4 Intervisibility of barrows: cumulative viewshed analysis

To further the investigation of the analysis of the visibility of long barrows, a new experimental method was designed to test the hypothesis that the barrows were sited in areas from which the number of other visible barrows was higher than would be expected through the operation of pure chance. In order to do this, it was necessary to obtain the number of barrows visible from each barrow, and to establish an appropriate significance test to test the hypothesis.

All of the LOS maps for the barrows in each region were summed using simple map algebra: each individual LOS map recorded 0 for 'not visible' and 1 for 'visible', therefore the arithmetic sum of the LOS maps resulted in a single overlay, describing the distribution of an ordinal level variable. The resulting maps (shown in figures 5.10 and 5.11) are a transformation of the elevation model into a surface in which every cell contains the number of barrows visible from it. The term *cumulative viewshed* will be used to describe this product, to distinguish it from the binary product used by Ruggles *et al* (1993), produced by a set union operation, and referred to by the authors as a *multiple viewshed.* Once these overlays were available, a point-select operation was performed using the long barrows as the point entities in order to extract the number of barrows visible from each long barrow in the region.

Lines	Population			Sample			
of sight	Area (km^2)	Area %	Cum. %	Cases	Cases %	Cum %	D
0	124.54	31.14	31.14	4	12.90	12.90	0.18
1	81.96	20.49	51.64	2	6.45	19.35	0.32
2	34.74	8.69	60.32	1	3.23	22.58	*0.38*
3	31.16	7.79	68.11	4	12.90	35.48	0.33
4	24.56	6.14	74.25	1	3.23	38.71	0.36
5	16.81	4.20	78.46	4	12.90	51.61	0.27
6	14.68	3.67	82.13	2	6.45	58.06	0.24
7	11.01	2.75	84.88	2	6.45	64.52	0.20
8	9.35	2.34	87.22	2	6.45	70.97	0.16
9	8.85	2.21	89.43	0	0.00	70.97	0.18
10	8.01	2.00	91.43	1	3.23	74.19	0.17
11	6.88	1.72	93.15	1	3.23	77.42	0.16
12	5.90	1.47	94.63	3	9.68	87.10	0.08
13	5.01	1.25	95.88	0	0.00	87.10	0.09
14	4.20	1.05	96.93	1	3.23	90.32	0.07
15	3.50	0.87	97.80	2	6.45	96.77	0.01
16	2.71	0.68	98.48	1	3.23	100.00	0.02
17	2.01	0.50	98.98	0	0.00	100.00	0.01
18	1.25	0.31	99.30	0	0.00	100.00	0.01
19	0.88	0.22	99.52	0	0.00	100.00	0.00
20	0.75	0.19	99.70	0	0.00	100.00	0.00
21	0.68	0.17	99.87	0	0.00	100.00	0.00
22	0.22	0.06	99.93	0	0.00	100.00	0.00
23	0.14	0.03	99.96	0	0.00	100.00	0.00
24	0.10	0.03	99.99	0	0.00	100.00	0.00
25	0.04	0.01	100.00	0	0.00	100.00	0.00
26	0.01	0.00	100.00	0.00	0.00	100.00	0.00

Table 5.4. Kolmogorov-Smirnov test for Stonehenge sub-region long barrows. Dmax is 0.38, d is 0.24 at the 0.05 level.

It should be noted that these values include the line of sight from each barrow to itself and therefore it is necessary to reduce the product of the operation by one in order to compensate for this. In this case one was subtracted from the value of the cumulative viewshed surface score so that the numbers refer to the number of other barrows visible from each barrow. This done, the cumulative viewshed surfaces can then be treated as a statistical population, from which the observations of the individual barrows are a sample. This fulfils the requirements for the construction of a significance test, in order to test the null hypothesis that the long barrows are sited with no regard for the number of other barrows with a line of sight from their location.

Figure 5.9. Example single line of sight image.

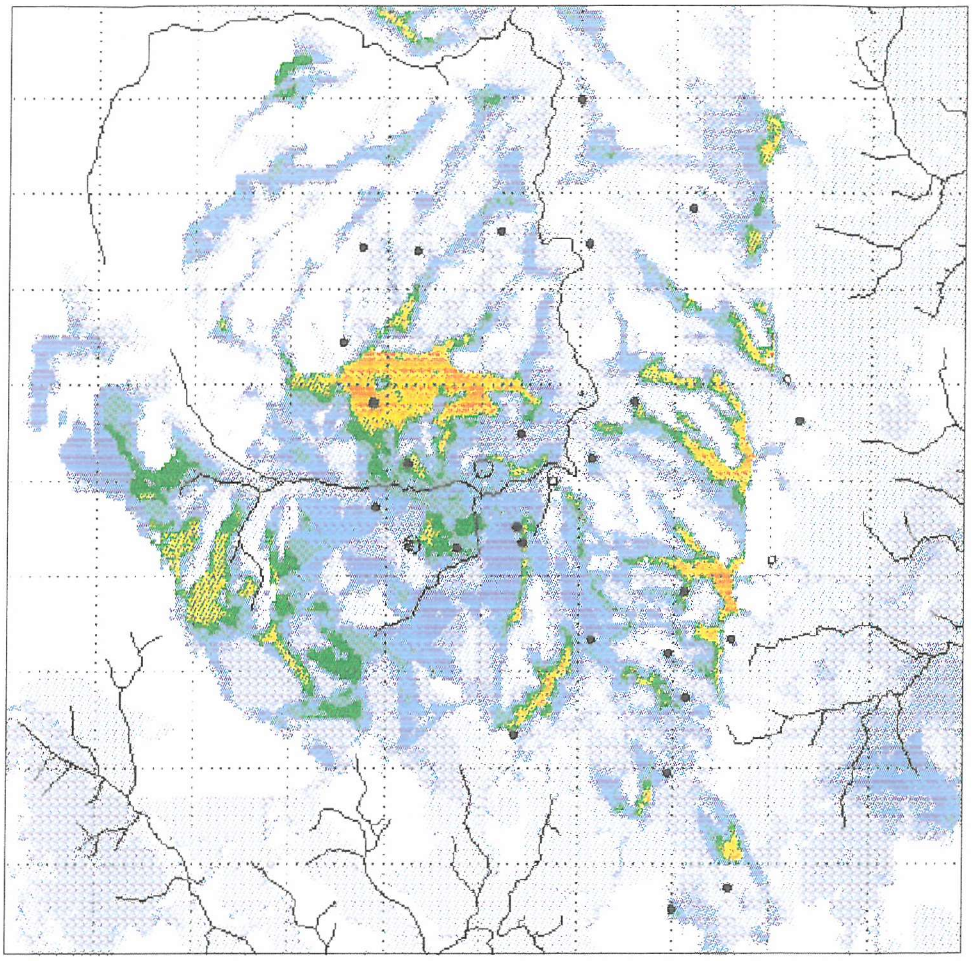


Figure 5.10. Cumulative viewshed image for the Avebury long barrows.

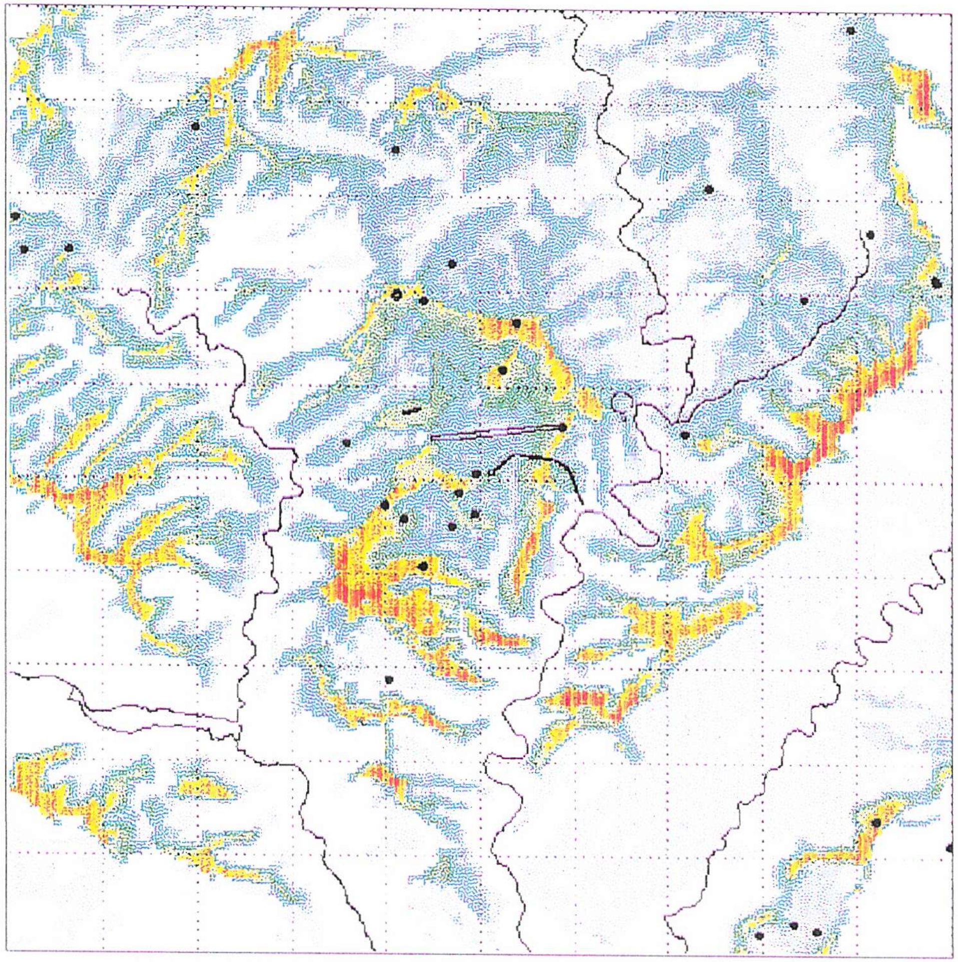


Figure 5.11. Cumulative viewshed image for the Stonehenge long barrows.

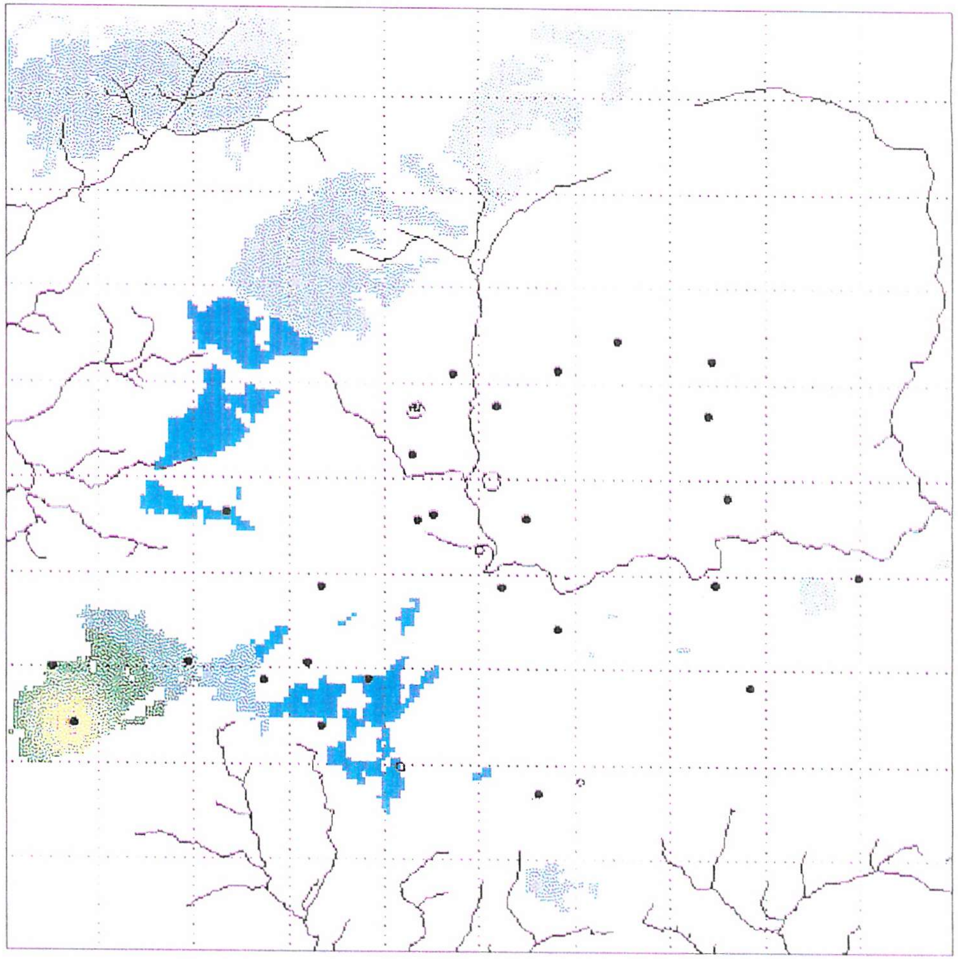


Figure 5.12. Example distance weighted viewshed map, colour coding is from 1 (red) to values approaching zero (blue).

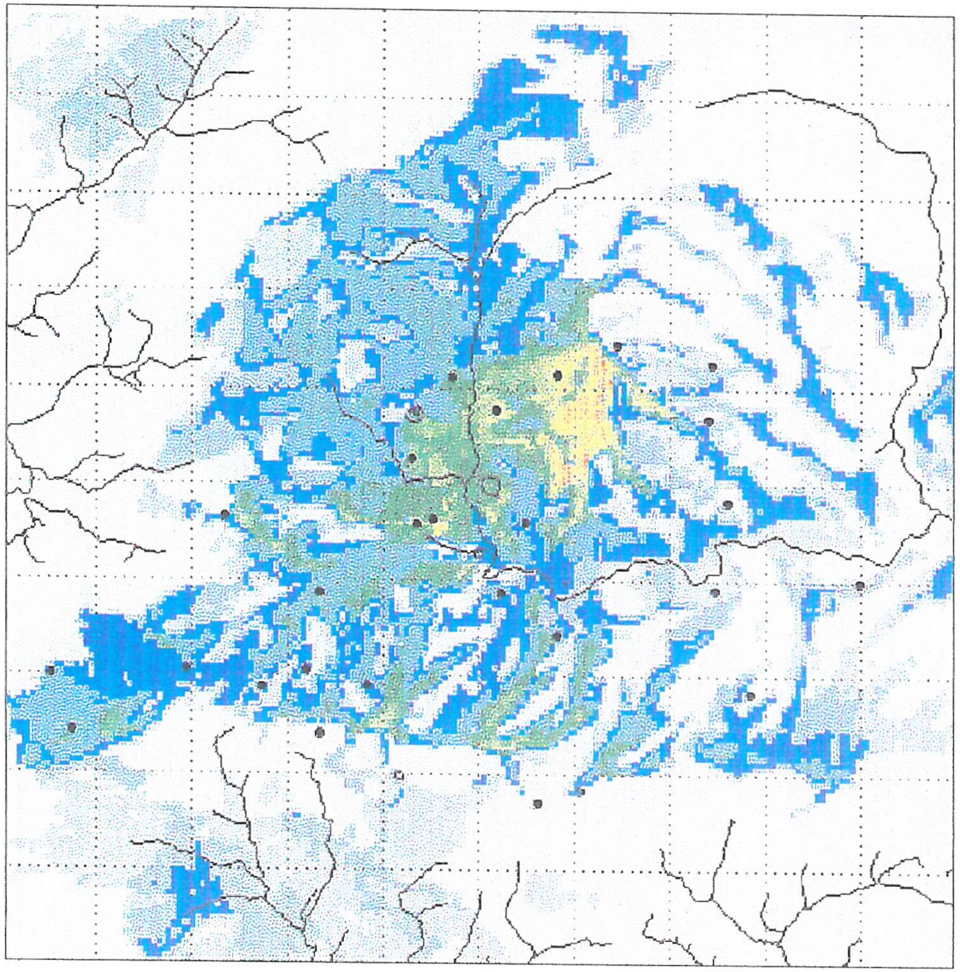


Figure 5.13. Visual impact of long barrows within the Avebury area. Values are essentially arbitary but range from around 3.5 (red) to approaching zero (blue)

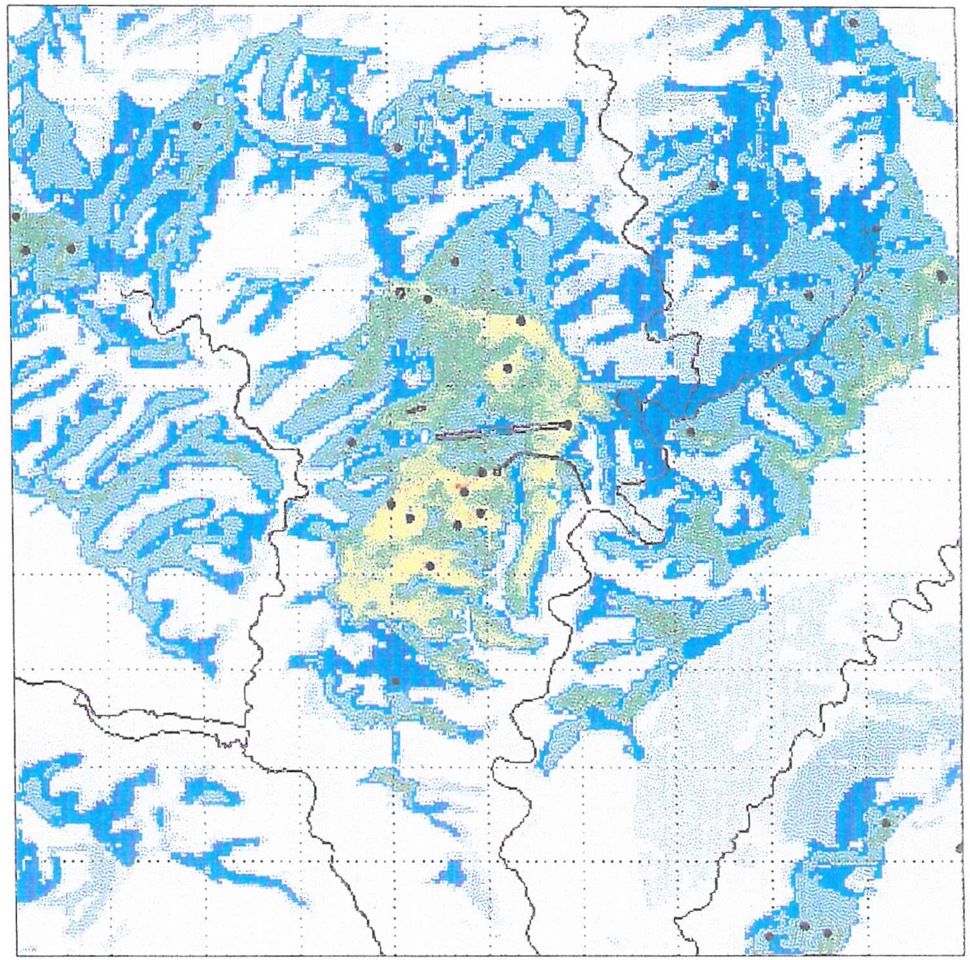


Figure 5.14. Visual impact of long barrows within the Stonehenge area. Values are as for figure 5.13.

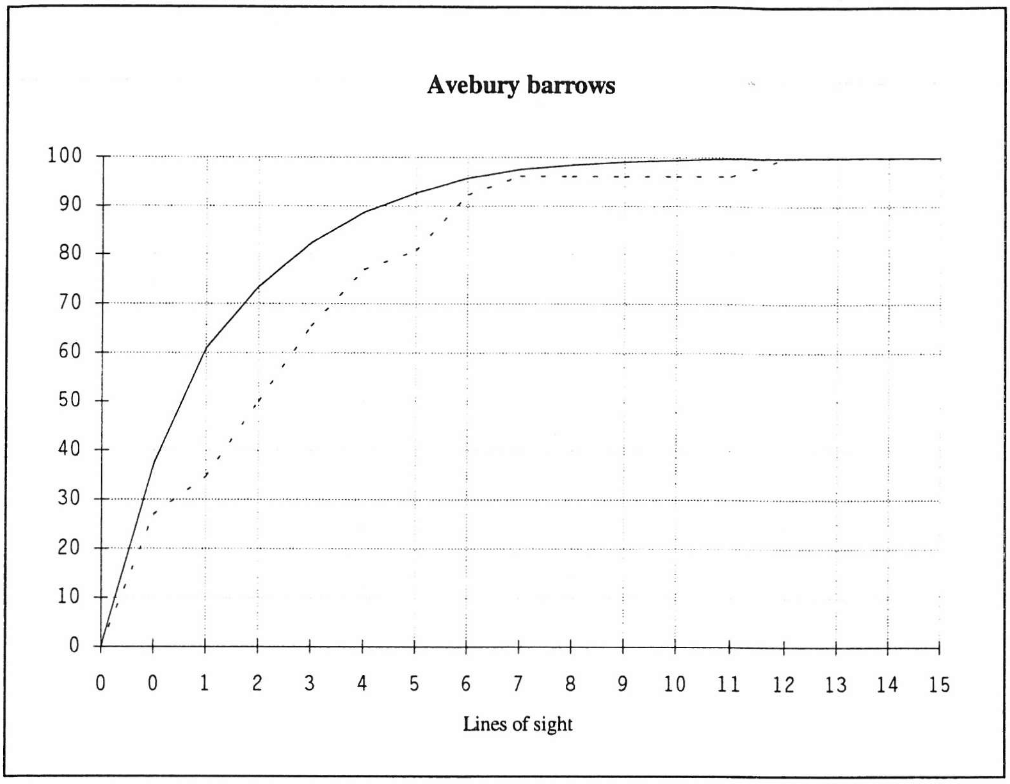


Figure 5.15. Kolmogorov-Smirnov test for Avebury region long barrow intervisibility. The population is shown as a solid line, the barrows appear as the dotted line (Dmax=0.26, for α =0.05, d=0.27).

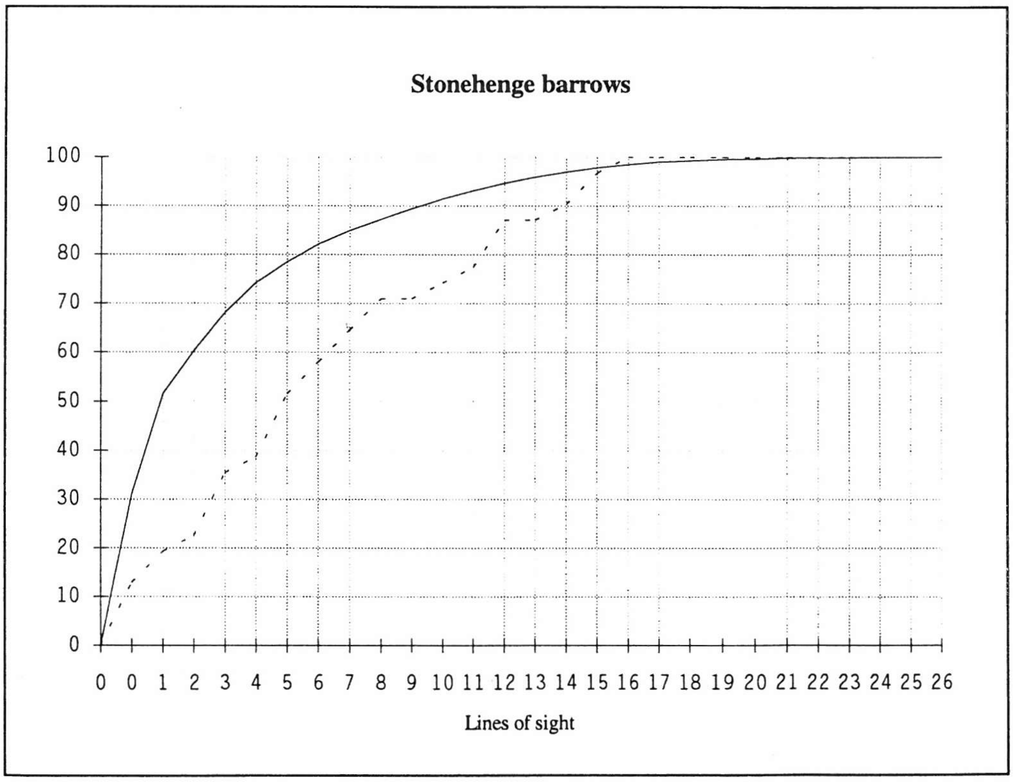


Figure 5.16. Kolmogorov-Smirnov test for Salisbury region long barrow intervisibility. The population is shown as a solid line, barrows appear as a dotted line (Dmax=0.38, for α=0.05, d=0.24).

Because the procedure above provides a population (the cumulative viewshed map) and a sample drawn from that population (the observations made with each barrow), a one-sample test is possible, an approach which is more powerful than the two-sample approach used for the area observations (Kvamme 1990b). For the same reasons as above, however, a non-parametric test should be used, and a Kolmogorov-Smirnov test is appropriate in this situation. The procedure (Shennan 1988) is to plot cumulative percentages for both the background (population) and for the barrows (sample). The maximum difference (expressed as a fraction) between these curves (*Dmax*) is then obtained. Large sample theory (Kvamme 1983b) dictates that the critical value, *d* is approximately $1.36/\sqrt{n}$ for the 0.05 significance level.

The results of this test mirror those of the preceding test: for the Avebury series (table 5.3 and figure 5.15), *Dmax* is 0.263, which falls just below the required value of 0.267 for rejection of H_0 at the 0.05 level. For the Salisbury series (table 5.5 and figure 5.16) *Dmax* is 0.38 which easily exceeds the required 0.24 to allow rejection of H_0 at the 0.05 level.

Although the same difficulties of association/causation are obviously a concern here, the sampling problem is less acute. While one form of sampling error is not eradicated by the use of a one sample test because the landscape has still been characterised by a set of 80m square cells, this test removes the uncertainty associated with sampling from that landscape and should therefore produce the same result when repeated. This is not necessarily true of two-sample approaches such as the one used above.

The procedures described above are very time consuming. Even with a fast computer, the number of discrete calculations which the GIS must perform is high for each line-of-sight map. This means that at this resolution, each line-of-sight map takes around 10-20 minutes to process. The line-of-sight files also occupy a substantial amount of storage space on the computer disk. Because of this, batch processing was used to automate the procedure and this is described in detail in appendix VII.

5.3.5 Critical assessment of method

Any new methodology must be critically assessed to try and identify systematic bias or weakness in the statistics or mechanics of its operation. Having conducted these two experiments, including the definition of cumulative viewshed analysis, it is therefore necessary to critically review the method and attempt to identify sources of and the likely effects of bias and error which are inherent in the methods themselves. Two possible sources of error and bias can be identified, the statistical effects of the sample sizes on the results, and the effects of the spatial distribution of the barrows, particularly the bias introduced by edge effects.

Sample sizes

It is tempting to regard the *difference* between the result for the two regions as significant. The Salisbury plain group of barrows produces a positive result, in that (at the 0.05 confidence level) we can reject the hypothesis that the distribution of barrows in areas of high intervisibility is due to chance. The result for the Avebury region is negative in that it does not allow the rejection of the hypothesis, therefore it is necessary to accept that the distribution of the barrows could be caused by chance. However, this is not to say that there is no relationship within the Avebury long barrows: observation and the cumulative frequency distribution demonstrate that there is, in fact, the same relationship between barrows and intervisibility in the Avebury series as in the Salisbury Plain series. *Not rejecting* the null hypothesis that the barrow distribution is caused by chance and that there is no relationship, must not be confused with *accepting* the null hypothesis. While the tests show that there may be a relationship within the Salisbury plain barrows, they do not show that there is no relationship within the Avebury data.

It is possible that the reason why the results are different for the two regions is that the sample sizes are different. There are 31 barrows (observations) in the Salisbury Plain series, while there are only 26 in the Avebury series. Statistically, the larger the sample, the greater confidence we may have that the difference between the distributions is not caused by chance and this intuitive relationship between sample size and significance is reflected in the Kolmogorov-Smirnov procedure in that the d value which must be exceeded for rejection of the null hypothesis is related to sample size. Consequently a stronger relationship must be apparent for the Avebury series ($d = 0.267$) than for the Salisbury Plain series ($d = 0.24$). Even then, the Avebury series barrows only just fails to pass the test at the selected confidence interval (*Dmax* is 0.263 against d of 0.267), and the result must be regarded as marginal.

Edge effects

A second possible cause of error within these calculations may be termed edge effects. These are an inevitable consequence of defining an edge to any study region, and might bias the results of the test in the following two ways. Firstly, in the viewshed areas analysis, the areas of the samples (both barrows and random) will be underestimated if they are close to the edge of the study area because the GIS only calculates the area which is visible within the raster. As this is equally true of the barrows and the random samples, it should only effect the analysis if the barrow samples show a tendency to be unusually spatially central or spatially peripheral. Secondly, the cumulative viewshed products may be prone to under-estimate towards the edge of the study area, because the calculation will not take account of barrows outside the study area which are in fact visible from within it.

Observation of the barrow distributions suggests that there may be some cause for concern in this regard: the Avebury barrows seem to be centrally clustered, while the Salisbury Plain series have a number of cases close to the edge of the study region. The distribution of the observations with respect to the edge of the study area can be quantified to some extent. The regions are divided into square bands which are 0-2, 2-4, 4-6, 6-8 and 8-10 km from the edge of the area, the barrows in each band are counted, and the density of barrows per km^2 for each band calculated. The result is shown in table 5.5 and seems to confirm the visual appraisal: the Avebury series have a low density for edge areas, steadily increasing to the centre of the region while the Stonehenge series show a high number of barrows at the edge of the region and potentially susceptible to under-estimate of visible area. Random selections of points can be expected to have a density which does not change through distance from the edge.

	Avebury Series		Stonehenge Series	
Distance from edge	Number	Density	Number	Density
0-2km	2	0.014	12	0.083
2-4km	3	0.027	4	0.038
4-8km	9	0.113	3	0.038
6-8km	7	0.146	5	0.104
8-10km	5	0.31	7	0.438

Table 5.5. Distribution of the barrows in bands of different distance from the edge of the study.

Although this seems to indicate that the results may be suspect because of edge effects, it should be noted that the effect on the Stonehenge series must be to (a) reduce the calculated view areas of some of the barrows in the area analysis and (b) to underestimate the number of barrows which can be seen from some barrows in the cumulative viewshed analysis. But the results of the area analysis shows the barrows to have significantly *greater* areas than would be expected, while the visibility study also shows the barrows to have a larger number of other barrows visible than would be expected in a random sample, therefore it seems likely that the bias introduced by the edge effect would, if anything, increase the significance of the result rather than decrease it. Those barrows which must be most prone to under-estimation are those which occur towards the corner of the study area. Most of these also fall to the south of the River Bourne and therefore an archaeological case might be made for their exclusion from the analysis, in that they might belong to a separate geographical group of barrows to the east. As above, their inclusion in the sample might have reduced the significance of the result by under-estimating the observed intervisibility and visible areas but seems extremely unlikely to have increased the significance of the result. In both tests, therefore, it seems that the result was probably unaffected by edge effects.

In the case of the Avebury series, the edge effect may have biased the result the other way: the fewer-than-expected edge observations will have tended to (a) over-estimate view areas in comparison with a random sample and (b) bias the cumulative viewshed result to over-estimating intervisibility. In both cases, the effect would be to reduce the significance of the result rather than increase it and as the Avebury series did not show a significant result in either experiment it seems reasonable to conclude that the result would not be different in the absence of edge effects.

5.3.6 Critical assessment of results

Causation verses association

If it is accepted that the results are methodologically plausible, then it is necessary to identify precisely what has been observed, and critically assess what the results means in statistical terms. Clearly the test on the Stonehenge barrows might be interpreted as revealing an intention on the constructors' behalf to site long barrows in areas of high visibility: in other words that the intervisibility of barrows is the cause of their location. However, it is equally plausible statistically that a third entirely different variable which is itself associated with the area of the viewshed may be the causing factor. The most obvious candidate, in this regard, is topographic height: it is possible that it was the elevation of the locations which influenced the constructors of the monuments to choose the locations they did, and that the relationship of intervisibility is simply a product of this choice.

It is difficult to refute this interpretation, although it can first be pointed out that even if this explanation of the relationship should be accepted, an archaeological explanation would have to be sought for the selection of high places within the landscape. Any such explanation would rely on the characteristics of high places, perhaps the most obvious of which is the increased visibility which is generally obtained at such places. It could be argued that in archaeological terms it is far easier to conceive of the elevation as a by-product of visibility, than of the visibility as a by-product of elevation. Unfortunately the relationship between elevation and visibility is such that it will never be possible to entirely ascribe the cause of the location to one or the other.

Nevertheless, it is possible to test the locations of the barrows with respect to elevation in exactly the same way as for intervisibility. The elevations for each barrow in each series was therefore extracted with a point-select operation within the GIS, and the areas of each value in the elevation model were generated in report format. This provided the sample and population for a Kolmogorov-Smirnov test as above, the results of which are shown in figures 5.17 and 5.18. The tests were then carried out to decide whether or not to reject the null hypothesis:

H_0: *that the long barrows are randomly distributed with respect to elevation*

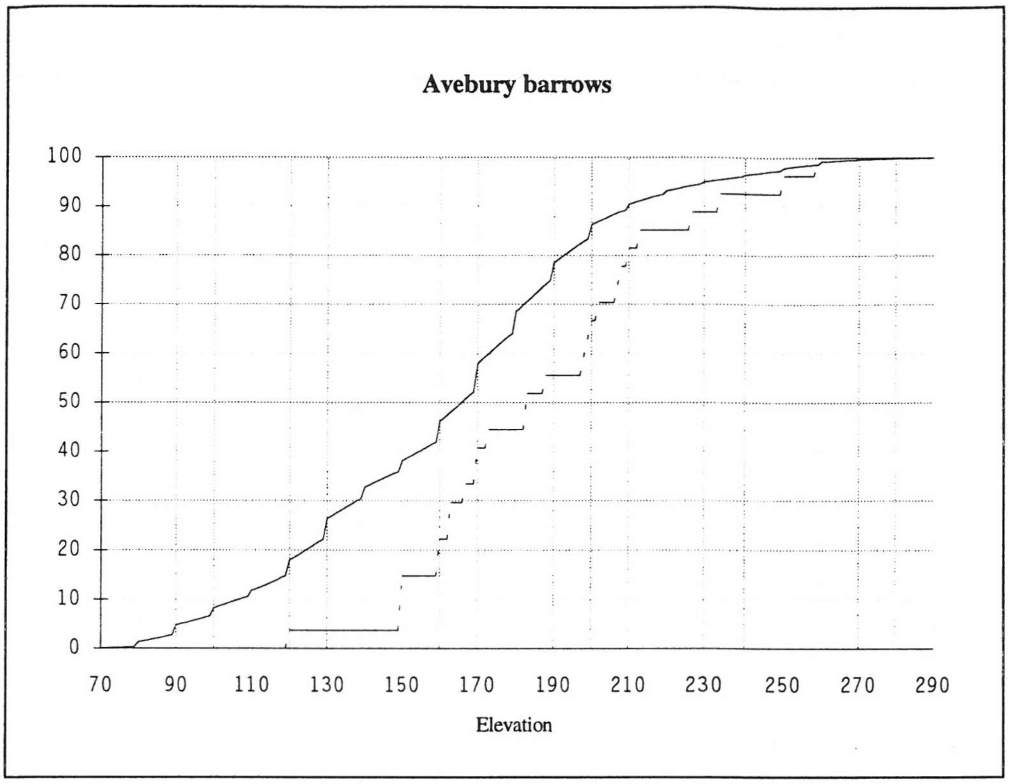


Figure 5.17. Kolmogorov-Smirnov Test for Avebury region long barrow elevations. The population is shown as a solid line, barrows appear as a dotted line (Dmax = 0.36, for α =0.05, d=0.26).

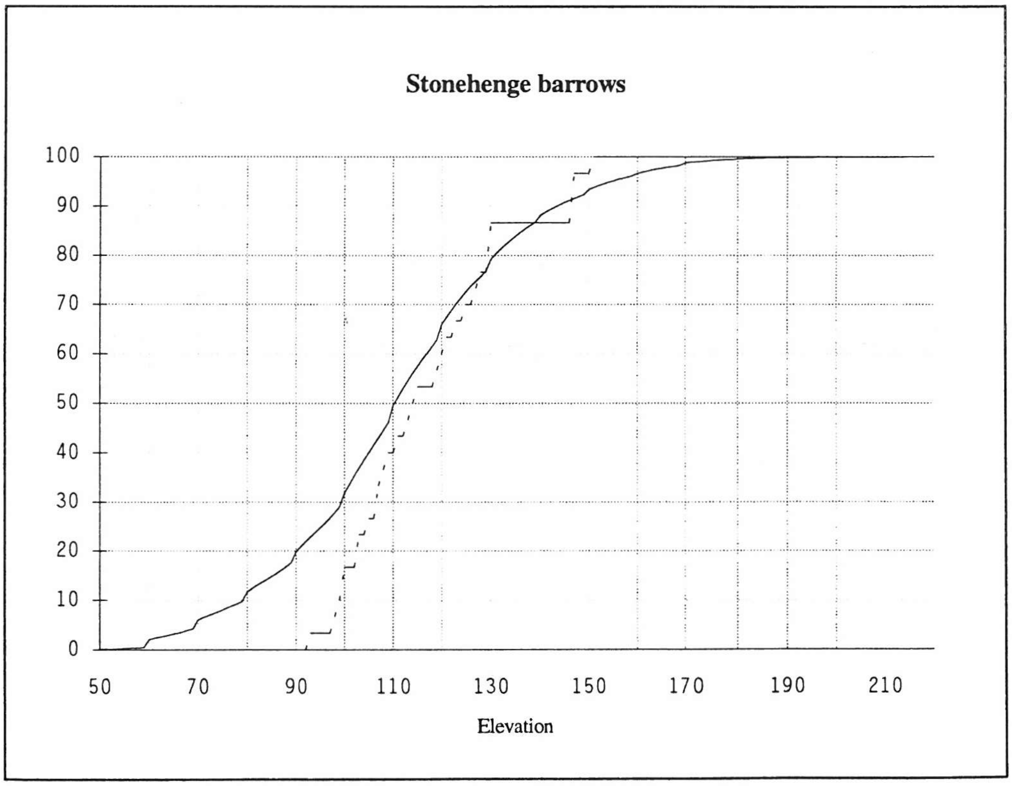


Figure 5.18. Kolmogorov-Smirnov Test for Salisbury region long barrow elevations. The population is shown as a solid line, barrows appear as a dotted line (Dmax=0.23, for α=0.05, d=0.24).

In contrast to the cumulative viewshed analysis, these two tests show entirely the opposite result: the Avebury test returns *Dmax* of 0.36, which exceeds the 0.27 threshold to allow rejection of the null hypothesis at α=0.05. At this confidence level, therefore, the Avebury barrows are not randomly distributed with respect to elevation. Observation of the distribution reveals that, as expected, they tend to occur in areas of higher elevation. For the Salisbury Plain series, however, the test returns *Dmax* of 0.23, against *d* of 0.24 which does not allow rejection of the null hypothesis, and at the 0.05 level, the relationship between barrows and elevation cannot be shown to be non-random.

Although these results are far from conclusive, and the same concern must be shown towards negative results as were discussed above, it is interesting that the difference between the cumulative visibility population, and the Salisbury Plain barrow intervisibility is clear enough to reject the hypothesis of no difference, whereas the difference between the elevation population and the barrow elevations is not. Entirely the reverse is true of the Avebury series, where the null hypothesis may not be rejected for the difference between the cumulative viewshed population and the barrows, whereas the difference between the distributions of barrow elevations and elevation is clear enough to reject it.

Calculated versus actual visibility

Along with the unresolved problems of causation and association discussed above, is the rather more pragmatic problem of how the line-of-sight maps actually relate to visibility in a prehistoric landscape. At a practical level, observations made at Windmill Hill on a clear day, drawing lines on a map from Windmill Hill to recognised points on the horizon, suggest that the equivalent line-of-sight map for Windmill Hill was a good approximation of the visible landscape from that location. Ruggles *et al* (1993) have suggested that correction should be made for earth curvature at larger scales, and this is clearly important in studies which rely on precise alignments. Practical observation, however, suggests that this may not be necessary in this type of study at this scale.

Because the landscape was considerably more wooded in prehistory than it is today, a line of sight does not necessarily mean visibility - it cannot be disputed that people cannot see through trees. One way of approaching this might be the creation of a deductive model of woodland according to ecological preference, followed by the simulated 'clearance' of some areas suggested by environmental work (e.g. Smith 1984). Allowance could then be made for the height of the woodland in the landscape in the calculation of the LOS maps. What degree of certainty could be assigned to such models is highly unclear. Certainly it would be interesting to add a value for the mean height of trees to the elevation raster and then repeat the study - intuition suggests that the LOS of the barrows should be reduced to a similar degree to that of the rest of the landscape and thus the differences between barrows and background populations may be unaffected.

However, there is considerable environmental evidence that the earlier Neolithic period is characterised by a marked opening of the landscape (see chapter 3). Although difficult to date, pollen sequences (e.g. Waton 1982) suggest a marked reduction in the forest cover during this period. Snail evidence, better dated and more localised in its relevance suggests that long barrows tended to be built in cleared areas: South Street, for example, was built over a cultivated field from which ard marks were preserved at the base of the ploughsoil (Ashbee et al 1979). Other barrows, such as Horslip, Beckhampton Road and West Kennet (Evans 1972) have been shown by molluscan analysis of buried soils to have been built in areas of open grassland. Although intervisibility cannot be reliably claimed for every single line of sight, therefore, it is likely that there was a significant element of intervisibility between monuments. Additionally, it may be noted in mitigation that while a calculated line of sight does not necessarily mean genuine visibility, visibility is not possible for areas without a line of sight. In general terms it therefore seems likely that the real visibility and the calculated visibility within a landscape are closely related in distribution if not in magnitude, whether or not more trees are present to reduce visibility. As the cumulative viewshed analysis tests the difference between the population and the barrow sample, the relationship would likely persist if the overall level of visibility was reduced.

It seems, therefore, that the relationship observed by the analysis is only likely to be altered if (a) the level of visibility between barrows was reduced to a greater degree than the overall level of visibility of barrows in the landscape or (b) the overall level of visibility within the landscape was so reduced as to preclude all intervisibility of barrows. On balance, neither of these conditions are likely.

Robustness and sensitivity

One group of approaches to the testing of this type of method might be grouped together as sensitivity studies. It should be possible, for example, to repeat the study with different values specified for the viewer's height above the ground surface, and then to observe how great the difference must be before the result is different. Time constraints have so far prevented this (the line-of-sight maps take some considerable time to generate) but work by Harris and Lock (Harris *pers. comm.*) suggests that the technique may be fairly robust in this regard. Other tests might include the random removal of barrows from the study, followed by repeating the test in an attempt to understand how robust the technique may be to missing (undetected) long barrows, or long barrows which may on excavation prove to be have been wrongly identified.

5.3.7 Distance and visual impact

Although it is easily possible to see 18km across the chalklands (personal observation), the visual impact of a monument at such a range is obviously considerably reduced. As an aid to understanding the visual impact of the monuments, and the relationship to centrality discussed above, a series of weighted viewshed maps were generated. These were calculated by generating a 'proximity surface' for each barrow, defined as shown in equation 5.4.

$$\frac{1}{(\text{distance from barrow})+1} \qquad 5.4$$

This results in an index which is intended to reflect the reduction in visual importance of a monument with increasing distance - the index varies from 1 at the barrow itself, declining to 0.5 at 1km, 0.3 at 2km, 0.25 at 3km, 0.2 at 4km and so on. A binary mask operation can then be used to apply the barrow viewshed to this surface, excluding all areas not within the line-of-sight map. An example of the result is shown in figure 5.12.

Having obtained a set of these distance-weighted viewshed maps for the barrows, it is then straightforward to sum these for each study area to produce a good estimation of the cumulative visual impact of each area on the long barrows (a similar batch procedure to that used for the unweighted viewshed maps may be used - see appendix VII). The results are shown in figures 5.13 and 5.14.

This procedure is essentially a combination of the linear centrality experiment above, and the cumulative viewshed experiment. The resulting images might therefore be interpreted in a similar way to the representations of centrality, although they clearly represent a rather more sophisticated concept of centrality than the linear distance model. The images essentially represent a relative measure of the extent to which each place in the landscape is visually dominated by the long barrow monuments. If the earthworks at the long barrows are intended to be a reminder of ritual or symbolic authority, visible to those who are engaged in their day-to-day activities, then these maps may be interpreted as at least an indication of the areas in which these activities may have taken place. Of course simply because a place is highly visually dominated by barrows does not impy that this was intentional, but in a landscape which is structured to the extent that these are by ritual and symbolic, monuments, these images may possibly suggest something of the intended 'visual audience' for the barrows.

The image of the Avebury area shows the extent to which the ridgeway path, to the northwest of the Kennet, is dominated by long barrows. Compared with the un-weighted visibility map for Avebury, the area between Windmill Hill and Avebury is shown to be more important. The positions of the enclosures with respect to the visual impact of the barrows, however, reflects the findings of the cost surface experiments: the marginal positions of Rybury and Knap Hill to the earlier Neolithic ritual complex is clearly revealed. Windmill Hill, although far less marginal, itself appears to be on the edge of the central zone of long barrow visual impact.

The visual impact image of the Stonehenge area is equally interesting. Compared with the unweighted image, the distance weighting reduces the importance of the southern sides of the Avon valley while emphasising the 'core' area between the two rivers. Edge effects may be partly to blame for this, but the character of the area in the centre of the image is unlikely to have been substantially affected. King barrow ridge appears as a prominent area of high value, while the notion that the ritual landscape might have been divided between north and south (discussed in chapters 6 and 7) is given credence by the apparent 'hole' which is traversed by the later Cursus monument. Again, the position of the enclosure site at Robin Hood's ball appears to be entirely marginal, standing right on the edge of the apparently core areas of barrow visual influence.

5.3.8 Archaeological interpretation of the result

Having accepted, with reservations, the conclusion that the Stonehenge barrows tend to be located in areas of high visibility, and that at least some barrows of the Stonehenge series were deliberately located in areas from which a high number of other barrows were visible, it is necessary to seek some archaeological explanation for this.

Before an explanation is advanced, however, it must be observed that the visibility and intervisibility of the monuments cannot have been the sole motivating factor in the choice of location. If the maximisation of visibility or intervisibility was the only factor in the location of the barrows, the monuments would all be located in one small area, probably in a circle. Therefore the selection of locations with high visibility and intervisibility, if it existed, must have been tempered by a somewhat conflicting desire to construct monuments within a wider area of the landscape. The observed distribution of long barrows must be interpreted as the product of both of these intentions.

With this in mind, one clue as to the social function of the choice of location has already been mentioned: the incorporation of existing monuments within later ones, as at the Dorset Cursus. Here Barrett *et al* (1991) have shown that the builders of the cursus included references to existing monuments in the design of cursus. This involved not only the incorporation of a long barrow into the cursus bank but also, most interestingly, revealed that the builders deliberately altered the course of the cursus to enclose a long barrow so that it `framed' the midwinter sunset viewed from the cursus terminal. Such physical and visual references to earlier monuments are perhaps an attempt to appropriate the status and associations of an older monument into the new structures, and the same mechanism could be at work in the positioning of the Stonehenge long barrows. In a society in which appeals to past traditions and practices are very apparent in other aspects of the funerary ritual it seems distinctly possible that visual references to other barrows may have constituted part of the mechanism by which social structures were reproduced and re-negotiated.

More specifically this could imply that those who directed the building of the monuments felt that the ability to see existing monuments added authority to the new structure through appeals to the historic authority of the existing monuments. Those in control of the new monuments, would then gain added legitimacy, and be in a better position to retain their own status and authority. This is not to imply that there was no counter to such strategies: other claims to social authority may be evidenced by the use of different locations and types of monument for ritual activities. What may have been revealed by this analysis is just one of many tactics employed by these people to negotiate and transform systems of authority and control.

This interpretation of the locations of the barrows differs in one other way from that of Renfrew (1973). As argued above, Renfrew's use of Theissen's polygons to generate the territories of the barrows assumes that they all functioned atemporally, in the same way, at the same time. Although, superficially, this analysis also treats the barrows as a group, the underlying argument and interpretation of the monuments is temporal in a number of ways. This interpretation recognises that there are several phases of activity at the long barrows: construction, use, re-use and then continued existence in the landscape. The link made here is between the construction of one monument and the existence of another. This interpretation also presupposes that the barrows were constructed as a chronological sequence, with the choice of location and construction of each new monument being influenced by the existence of earlier ones. Whereas Renfrew's choice of method, construction of Theissen's polygons to reveal the extents of 'family territories' is only a valid method if we assume that the barrows were all placed simultaneously into the landscape, the methodology devised and applied here is only valid if we assume the reverse: that the barrows form a chronological as well as geographical seqence.

It has been emphasised above that for reasons of sample size the different result of the Avebury and the Stonehenge series need not imply that there is a real difference between the series. However, even allowing for the effect of sample size, the statistics do leave a suspicion that if there is a similar relationship within the Avebury data, it may be that elevation itself is more important than intervisibility. Were this to be substantiated, then a number of possible reasons might be put forward for it.

Firstly the appearance of megalithic tombs in the Avebury sequence is different from Salisbury plain: the Avebury series contain examples of stone-chambered tombs while the Stonehenge series are exclusively earthen mounds. The creation of the impressive stone facade at West Kennet, seemingly unparalleled in the Stonehenge series, might be seen to add authority and importance to the tomb as well as focus attention on the activities of those performing at the entrance. In the Avebury sequence, therefore, the message of the monuments may have been focused less on the surrounding landscape and on the existence of other monuments than on the architecture and impact of the tombs themselves. If this is the case, then in the Avebury region those who held power were appealing more to monumentality in the structures themselves, and less to a relationship with the past as exemplified by existing monuments.

If this difference proves to be real, then it is an indication that there exist subtle regional variations within Neolithic communities. Seemingly very similar practices, such as monument construction, may in fact vary from location to location. That stone chambered tombs as well as earthen mounds were built in the Avebury region while only earthen mounds were built in the Stonehenge region might be explained in terms of the resources of the different locations themselves: those who build the monuments on Salisbury Plain may have had no access to large sarsen stones because they are found naturally in the Avebury region. On the other hand the difference may have been a deliberate assertion of cultural difference: two groups of people constituting themselves in opposition to one another and expressed through different practices.

In this context, a remark may be made concerning the result obtained by Lock and Harris (forthcoming). Although not tested within precisely the same framework which was used here, this seems to provide convincing evidence that the group of barrows investigated by these authors were situated for deliberate *non-intervisibility* from one another, a characteristic which, it is argued, supports the notion that these barrows were used as territorial markers. The simple fact that this result is different from that obtained here need not imply that it is wrong: it has consistently been emphasised here that different mechanisms are at work in the different areas. There must be a suspicion, therefore, that the people responsible for the construction of the barrows of the Danebury region had different motives again to those of the Avebury and Salisbury Plain region in their choice of location.

5.4. Conclusions

It was argued in the introduction to this thesis that, to date, GIS has been applied almost entirely to economic and settlement archaeology, frequently in a simplistic deterministic way. This chapter has suggested some ways in which existing GIS methods might be used to extend the understanding of ceremonial monuments rather than settlements, and ritual rather than economic archaeology. It was also argued that archaeologists who have used GIS methods have either imported wholesale the methods of economic geographers or, at best, used GIS to improve on existing archaeological methods. This chapter has presented a new method for the analysis of archaeological monuments, a method which stems from archaeological questions.

Through the generation of cost surfaces, which may be used to represent the time taken to travel between places in the landscape, the GIS can be used to view the ceremonial landscape in a new way. This, it is argued, is one way in which the human experience of the landscape may be incorporated into quantified studies of spatial phenomena: in this case the linear distance between monuments which has formed the starting point of most spatial analysis is modified by the experience of the time required to move acoss the landscape. The resulting transformations of cartesian space may then be a better representation of how the landscape was viewed by those who lived within it, and used to gain insight into the structure of the landscape.

The use of cost surfaces in a non-economic context can be seen to aid the interpretation of the earlier Neolithic monuments in both study areas, and is particularly revealing of the relationship between the three enclosures of the Avebury area. The study suggests that, if they share anything more than a construction method, the enclosures should be thought of as places which occur on the periphery of the geographic regions of the earlier Neolithic. This is in contrast to the current interpretation of these monuments as 'central places'.

The application of line-of-sight maps to the long barrows has similarly been shown to produce unexpected insights into the organisation of the earlier Neolithic ceremonial landscape. The study produces evidence (albeit of a not entirely conclusive nature) that the long barrows which surround the Stonehenge monument on Salisbury plain may have tended to be deliberately situated in locations which firstly had higher levels of visibility, and more interestingly exhibit higher than expected intervisibility with other barrows.

The archaeological interpretations of each of these studies may be open to question and debate. However, they are not intended to provide definitive answers to the archaeological problems of the Neolithic, rather to suggest new avenues which may not have been apparent without the methodological advances offered by the GIS. Together they demonstrate that the application of GIS to non-economic archaeological problems is both productive and worthy of further archaeological application.

Chapter 6
GIS and the interpretation of extensive survey data

6.1. Introduction

This chapter investigates the use of GIS in the interpretation and presentation of extensive survey (surface collection) data. Field survey of this type provides information of a fundamentally different nature to intensive investigations, and the resulting data is lacking in the spatial and chronological detail of, for example, excavation. Contrastingly, however, field survey data generally has the advantages over intensively collected data of greater spatial extent and greater chronological depth although these are often at the expense of both spatial and chronological resolution.

It is likely that data of this type will be of increasing importance within archaeology for a variety of practical and theoretical reasons. The first reason is theoretical: that many archaeologists have become dissatisfied with the notion of the archaeological 'site'. This was originally proposed by Foley (1981) in the context of Palaeolithic archaeology but is an idea which has found favour in the interpretation of later prehistoric archaeology as well. It is an approach which is more concerned with the activities of human populations through time within a wider spatial and temporal context and was termed 'offsite archaeology' by Foley (1981). According to Schofield, the aim of offsite archaeology has been to:

> *'reject the 'site' as a unit of analysis and to produce a viable alternative. This is based on the size of collection units, the nature and density of flint collections and on the distinction between types of activity and the locations within which they were predominant. In other words it is a means by which to distinguish between types of activity at a local scale.' (Schofield 1988 after Foley 1981).*

One characteristic of this approach has been its rejection of the use of type fossils to understand extensive lithic distributions. Schofield, for example, notes that these generally represent only around 1-2% of assemblages, and that a methodology for interpretation of lithic distributions should therefore be based on the majority of the material rather than the minority of artefacts. Unfortunately this is problematical because extensive surface collection produces a vast number of observations which make the task of observing patterns within the database difficult.

Offsite archaeology is also clearly related to the idea of 'landscape archaeology' (Roberts 1987), which developed through the 1970s. Although more difficult to define (see chapter 1), landscape archaeology draws on the increased availability of regional scale archaeological information, such as environmental sequences and the systematic use of aerial photography, to focus attention on the relationships between people and their surroundings at a regional scale.

Landscape archaeology frequently describes processes which take place over longer periods of time than archaeological narratives which depend on intensive archaeological practices.

The second reason why extensive survey data is of increasing importance within archaeology is far more prosaic. The last few years have seen substantial changes in the way in which the archaeological resource is managed. These changes, particularly the increased emphasis on the management of archaeology within the planning framework, have led to an increased use of 'non-intervention' techniques such as field survey and geophysical prospecting, rather than excavation. The full effect of these changes on the nature of archaeological research has yet to become fully apparent, but will certainly produce information of a wholly different quality but which nevertheless requires explanation. Such activity seems likely to encourage and facilitate the development of the 'offsite' and 'landscape' schools of archaeological theory as discussed above, and requires the development of appropriate methods for manipulating and storing information.

6.2. Aims

Within this context, it is the aim of this chapter to investigate the utility of GIS in the interpretation of extensive survey (surface collection) data. The data used for the investigation is a subset of that collected by Richards and other workers for the Historic Buildings and Monuments Commission for England (HBMC or English Heritage) between 1980 and 1986, generally referred to as the Stonehenge Environs Project (Richards 1990). This data was kindly made available in digital form by Wiltshire County Council Museums Service. Only data concerning the collection of flint artefacts was made available in this form. The pre-processing and import of the data itself is discussed in chapter 4.

Several principal tasks were approached with the flint data. Firstly, the visualisation capabilities of GIS may be employed to some advantage. Systematic field survey databases are, by nature, very large and this makes it extremely difficult to appreciate any patterns which may emerge within the data, particularly spatial patterns. The only attempt to visualise the Stonehenge Environs project data is that by Richards (1990) where different sized dots were plotted to represent different flint densities (see figure 6.1). This is clearly successful to the extent that gross patterns may be perceived in the flint densities, but the level of generalisation required to ensure that the dots remain of a small enough size may well be obscuring some of the more subtle changes within the data. The first aim of this part of the study is therefore to devise alternative methods of presenting the data which may overcome such limitations, and to explore whether it is possible to use GIS to reveal any patterns within the data which would be difficult to present meaningfully in other formats.

The second aim of this section is to integrate and compare the field survey data with other types of information, particularly existing archaeological data in the form of sites and monuments records and geomorphological data such as soil, elevation, aspect and drainage. In this way it is hoped that any patterned relationships between the field working data and other data may be made apparent.

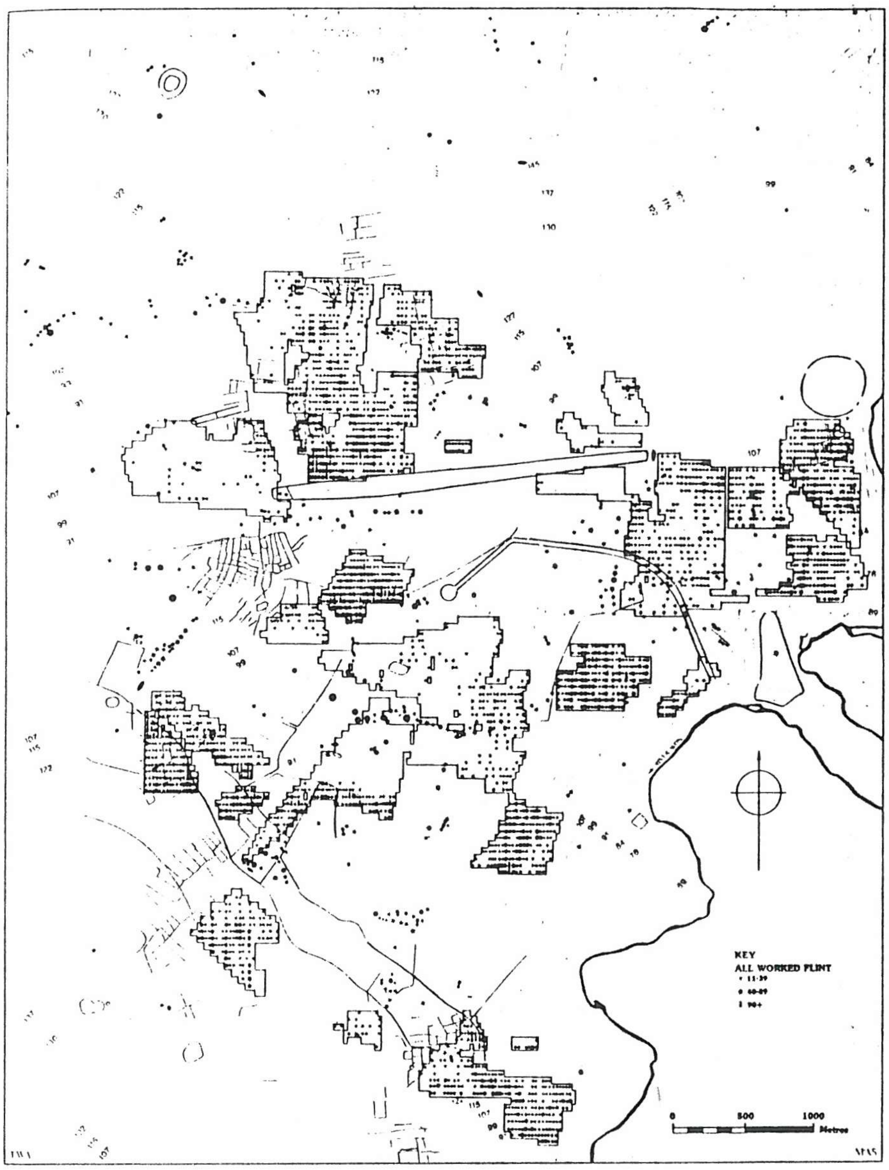


Figure 6.1. Example of manually prepared presentation of flint density data, from Richards 1990.

Lastly this section will explore the application of GIS to the problem of using field survey information from the surveyed areas to generate information concerning unsurveyed areas. It is in the nature of field survey databases that they are incomplete: only some areas of the landscape are sampled. This may be due to restrictions of cost or access, for example, or because only those areas which are under cultivation are appropriate for systematic field collections. This section will investigate the extent to which GIS may be used to interpolate or to model flint density values in unsurveyed areas from those in the surveyed areas. This is an important goal for both the understanding of the distributions revealed by surface collection, and the management of the archaeological landscape.

6.3. Previous work

The Stonehenge Environs surface collection database covers a series of irregular shaped areas to the east of the river Avon and immediately surrounding the Stonehenge monument itself (see figure 6.1). The total surveyed area is approximately 7.1 square kilometres, consisting of around 6500 50m walked transects taken at 25m intervals. In the course of the project a total of 102,175 pieces of worked flint were collected (Richards 1990 p15).

As discussed in chapter 4, five categories are uniformly present in the database for all walked areas and they provide sufficient variation for the purposes of an analysis within the terms set out above. No details of 'type fossils' were present or indeed sought, as the aim of the study is to extract as much information as possible from the majority of the flint data, rather than the minority. Coincidentally these categories are quite similar to those categories chosen by Richards (1990) for presentation in the Stonehenge Environs Project report.

No computer facilities were available for analysis as part of the Stonehenge Environs Project, and the original analysis took the form of plotting the densities of flint categories in categories derived from a frequency histogram of the flint densities (p16). Using the procedure advocated by Hodder and Orton (1976), the inflexions of this distribution were used to estimate classes for plotting as distribution maps using different sized symbols for the different classes. These distribution maps then formed the basis of the interpretation of the lithic data and the main findings were detailed by Richards (1990). In summary, the general pattern observed was as follows:

- Low densities of flint in several 'peripheral' areas to the north-west of the area (areas 80, 90, 62)
- Equally low densities of flint in the apparently 'core' area surrounding Normanton Down, south of Stonehenge (areas 54, 61, 79, 55, 56, 88 and 84)
- High densities in all areas south-west Normanton Down, in the dry valley running from Winterbourne Stoke Crossroads to Normanton Bottom (areas 50, 59, 77, 75 and 67), with the north side of Normanton Bottom apparently forming an abrupt discontinuity
- Very high values in the area immediately north of the Cursus (area 52), and south of the Cursus, to the west of Stonehenge (area 54)
- Moderately high values, although fragmented, in the areas south of Durrington Walls (areas 60, 69, 71 and 72)

Some analysis of the flint density data was also undertaken by Maskell (1993), who entered a subset of the data from the paper record to a database, and then undertook some analyses with the IDRISI GIS. Maskell's main conclusions were as follows. Firstly, using the Moran's I statistic, Maskell showed that the flint density data exhibited spatial autocorrelation: in other words that the flint densities were not random in their spatial distribution (high values in one location made it more likely that high values would occur in neighbouring areas). This is a useful first step in the analysis and Maskell's finding that there is spatial autocorrelation within the flint data will be confirmed below prior to the use of spatial interpolation procedures.

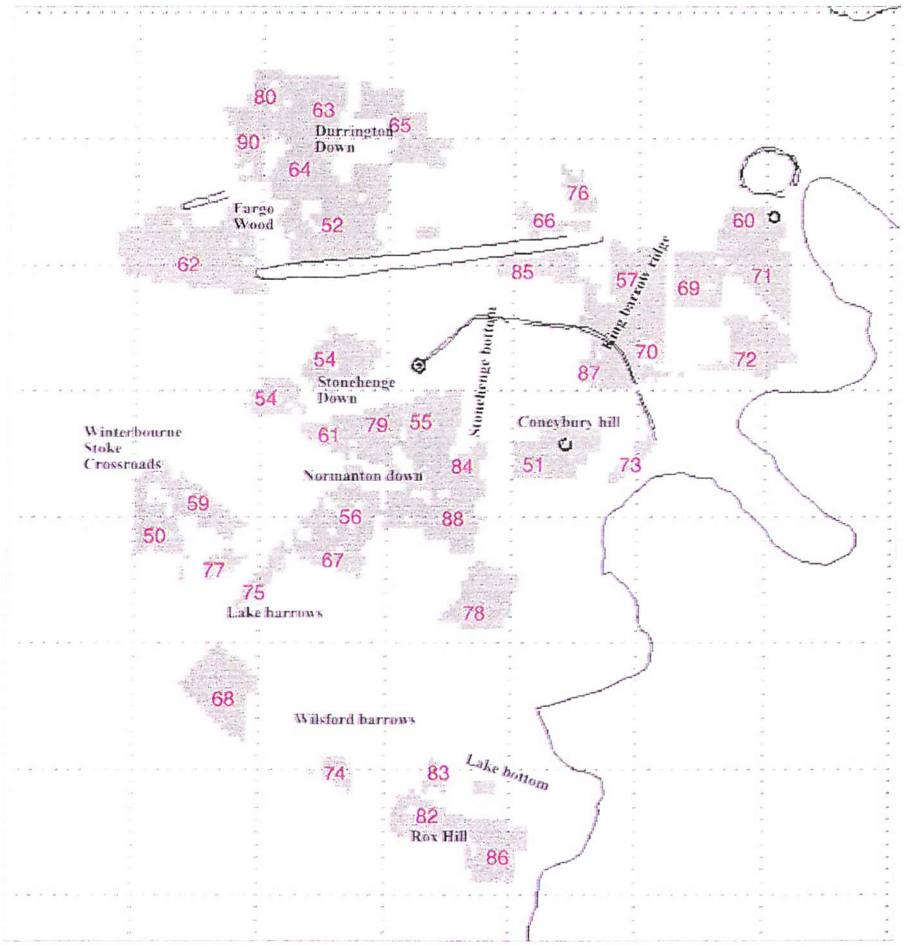


Figure 6.2. Area represented in the Stonehenge environs database, showing most of the original fieldwalking area numbers in red (from Richards 1990) and locations mentioned in the text. The grid on this and all subsequent images is the OS National grid, 1km intervals.

Secondly Maskell examined the relationship between presence/absence of these flint classes and three environmental variables: land use, slope and aspect, undertaking one-sample Kolmogorov-Smirnov tests as suggested by Kvamme (1990b). These suggested that there may a relationship between aspect and the presence of burnt pieces, scrapers and other tools. More specifically, the result suggested that burnt flint, scrapers and tools were more often present on north and north east slopes than would have been expected while less often on south and south-west aspects.

Curiously, there seems to be no such relationship for flint classed simply as 'retouched', which must include both scrapers and other tools.

Unfortunately, the simplification of the data into presence/absence of flint categories instead of flint counts to undertake the significance testing (necessary because of time constraints) discarded a sizeable proportion of the information content of the subset of the flint data which was analysed. Flint 'presence' may be caused by a variety of post-depositional processes leading to a possibly misleading results, whereas the comparative densities of flint should bear a far closer relationship to the original distributions. Statistics to analyse flint density as interval level data are to be preferred, and therefore correlation, covariation and regression techniques, as advocated by Schofield (1988), are to be used in the analysis.

6.4. Presentation and primary investigation

The import and primary arrangment of the data from 25m fieldwalking units into 50x50m resolution data themes was discussed in chapter 4. Once available within the GIS, each of the five selected attribute files was used to generate a SPANS quadtree overlay at 50m resolution, in which each 50m cell contained the count of flints of the particular class. One further overlay was generated by recording value 1 for quads which contained a point record and zero for those which did not. This was then used as a basemap in the subsequent analyses.

6.4.1. Presentation as density maps

The resulting flint density maps could then be displayed on screen in order to visually assess the distributions of the flint. Appropriate colour palettes were applied to the maps. For consistency the general principle was adopted that colours should run from blue for the lowest counts through green and yellow to red for high counts. For clarity, digitised outlines of the major monuments of the area and a projection grid at 1km intervals was overlaid on the maps before they were printed. The results are shown in figures 6.3, 6.5, 6.7, 6.9 and 6.11.

Inspection of this initial presentation of the flint counts suggests that the use of colour in this way is a valuable summary of the data variation. Compared with Richards (1990) plots, the areas of high flint densities are generally easier to observe in the output, and a greater range of the variation within the data is preserved in the summary. The generalisation of the units from transects to 50m square land parcels does not appear to have reduced the information content of the printed result, and the symmetry may indeed improve the interpretive value by removing the horizontal 'stripes' which appear in the original.

Unfortunately, the full range of colour and intensity variations available on the computer monitor is not well represented by the printed output. Nor can the printed maps adequately express the advantages of interactive visualisation of the lithic data: the ability to 'zoom' areas of the data, and to iteratively alter the palette to improve representation. The printed images show a small range of monuments and other features, whereas interactive use of the GIS allows any vector data to be overlain at will.

However the printed coloured output does seem to provide better differentiation of the variation within the flint density and between areas of high density. For example, in figure 6.1, the clusters of lithic density are apparent in a number of areas. Comparision of the area south of Durrington Walls (60, 71) with the area North of the Cursus (52) suggests that they are similar in character. The two areas of nucleation are equally clearly represented by figure 6.3, but here the far higher densities of lithics in the area north of the Cursus as compared with the Durrington Zone are immediately apparent.

6.4.2. Mean filtered maps

Close inspection of figures 6.3, 6.5, 6.7, 6.9 and 6.11 shows that there seems to be a level of random 'noise' in each of the images. This is to be expected within a sample of this type, and represents the chance effects of obtaining unrepresentatively high or low values in the samples for some land units. Were the fieldwalking repeated, identical values would not be expected within each of the sampled transects, and if it were possible to sample the same areas a large number of times these effects would average out to produce a less 'noisy' image.

To reduce the effects of the noise within the images, a second series of quadtree data themes was generated by applying a 3x3 mean filter to the images in figures 6.3 to 6.11. This is a standard image processing technique for reducing noise in images. In order to avoid the distortion which would arise due to edge effects in this case, a basemap of only those cells which had been sampled was used to force the filter to regard cells outside the basemap as 'no data' rather than zero values. The filtered images are presented in the same format as above in figures 6.4, 6.6, 6.8, 6.10 and 6.12.

The effects of the filter in removing the element of noise are readily apparent in figures 6.4, 6.6, 6.8, 6.10 and 6.12. A little of the original pattern within the data has been discarded, yet the clusters of flint visible within the original images are made easier to distinguish by eye. While these filtered images may not be as immediately accurate a representation of the data, it seems likely that they provide a better representation of the relative variation between the units. For example the 'striped' effect visible in area 68 of the original total flint and total flake images (figures 6.3 and 6.5) is likely to be an artefact of the collection rather than a genuine characteristic of the population of flints. This effect is removed by the mean filter (at the expense of a little detail) and a general scatter, with higher values towards the centre of the area is produced in the filtered images of area 68.

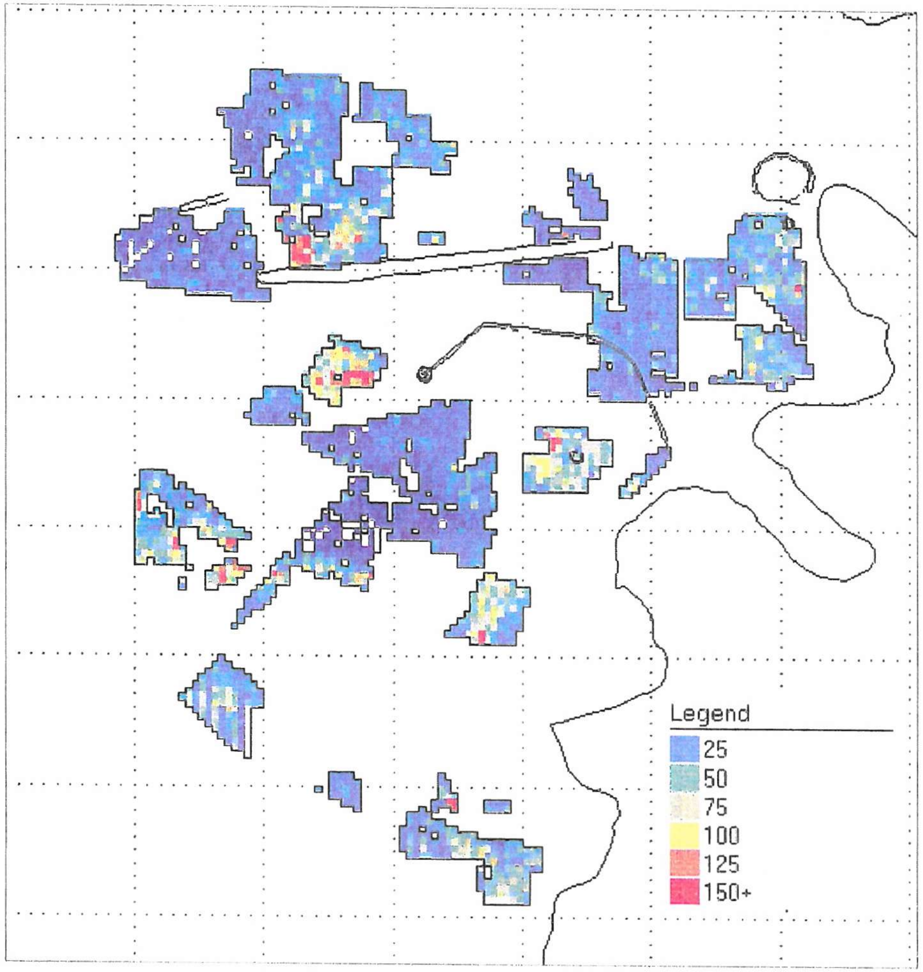


Figure 6.3. Total flint densities

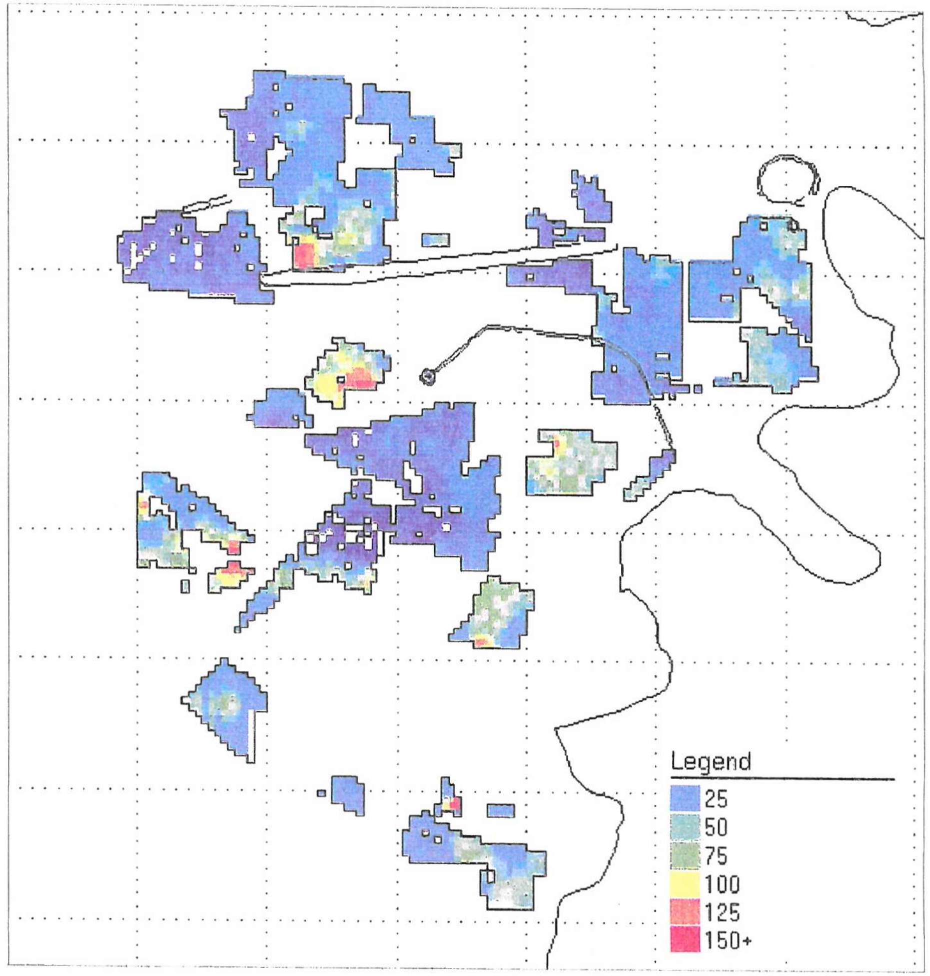


Figure 6.4. Total flint densities, mean filtered

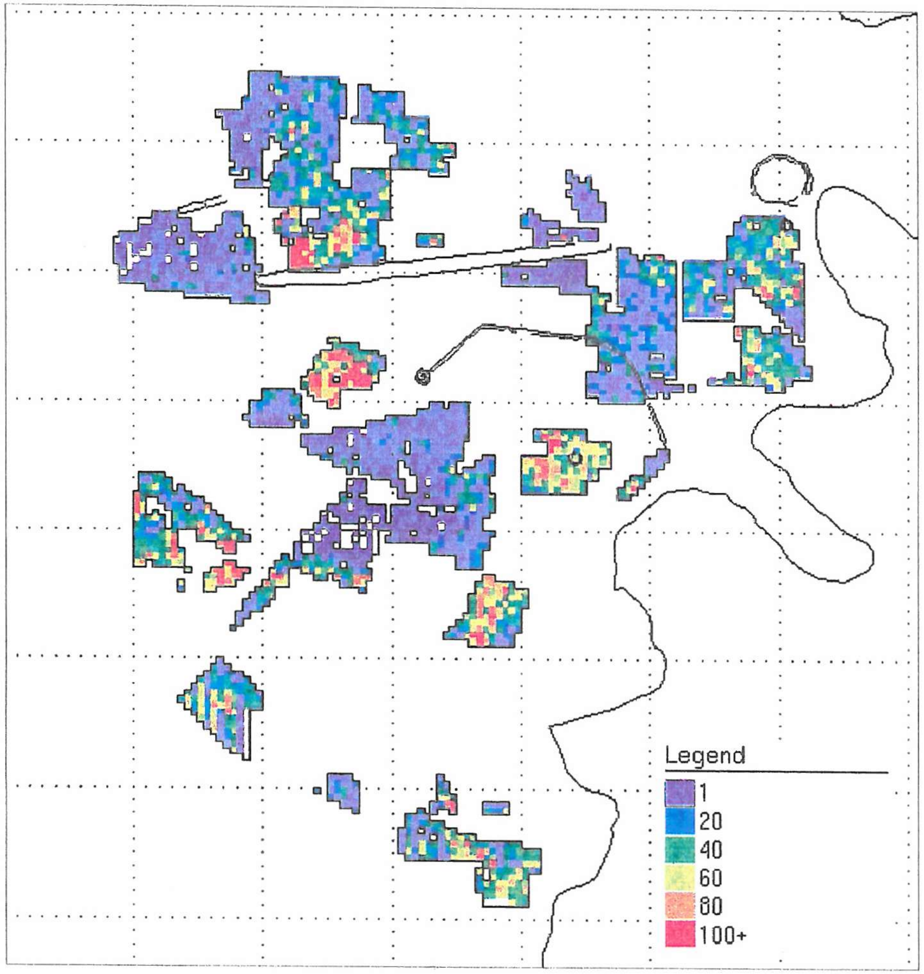


Figure 6.5. Flake densities

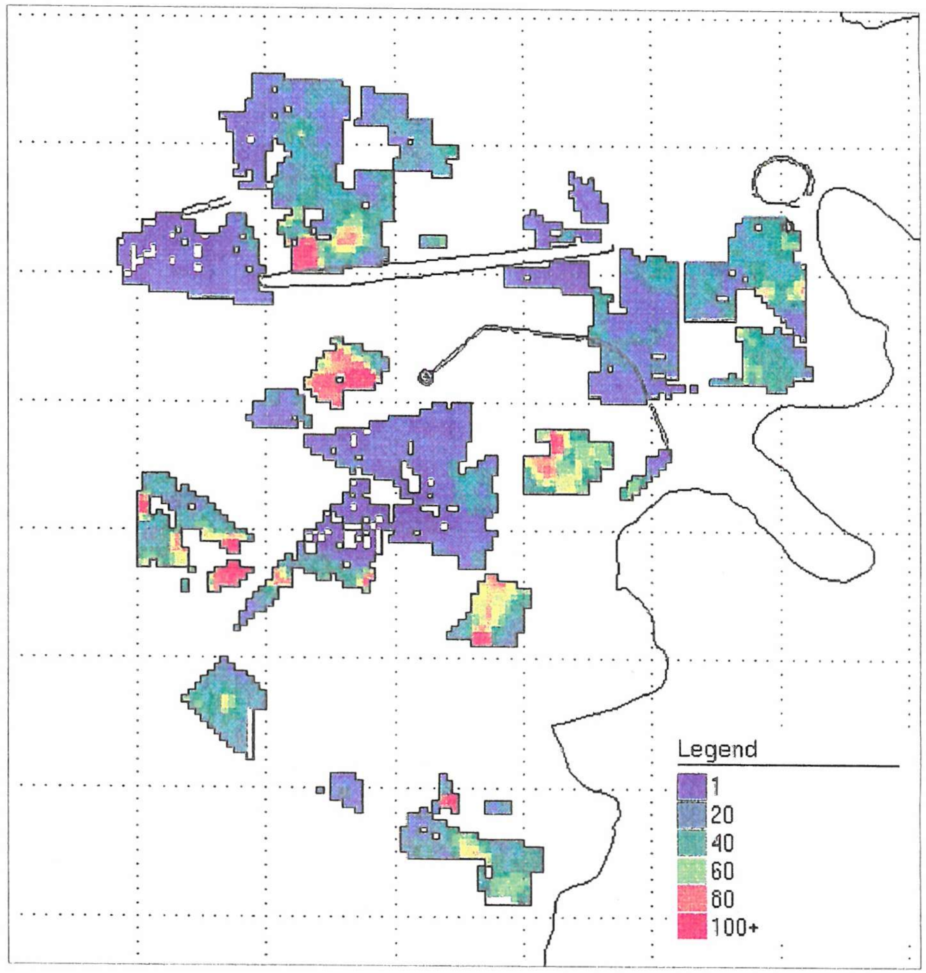


Figure 6.6 Flake densities, mean filtered

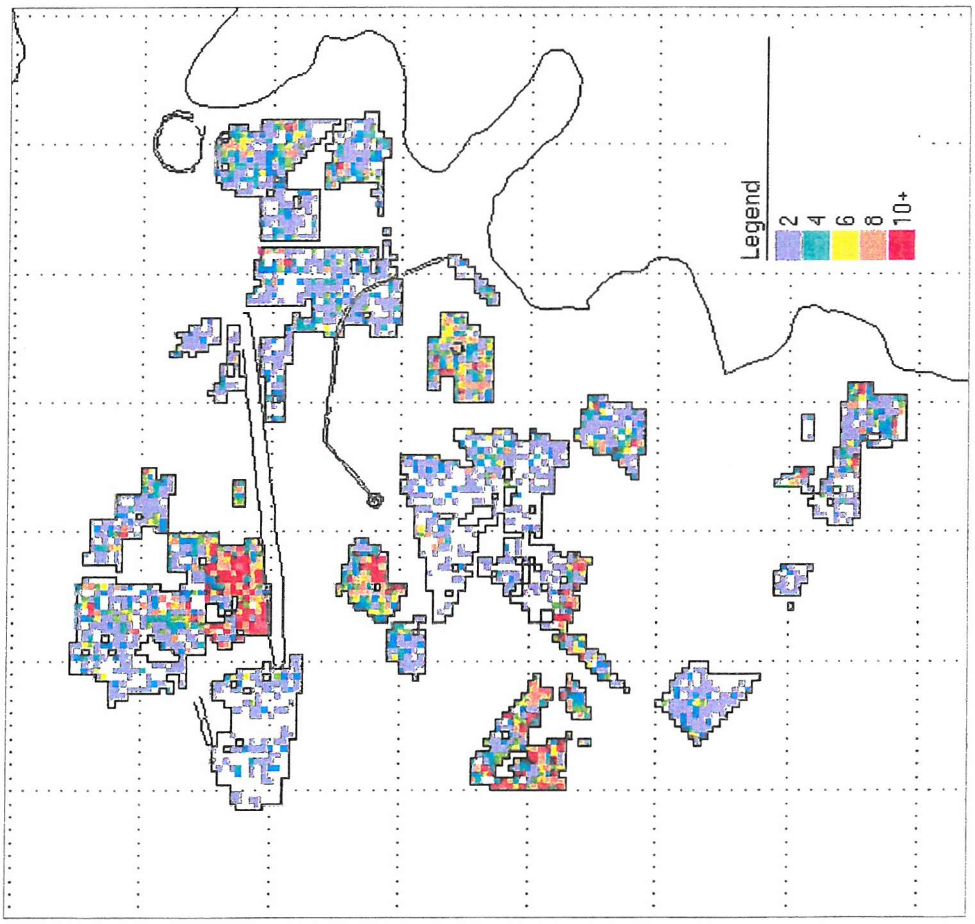


Figure 6.7. Core densities

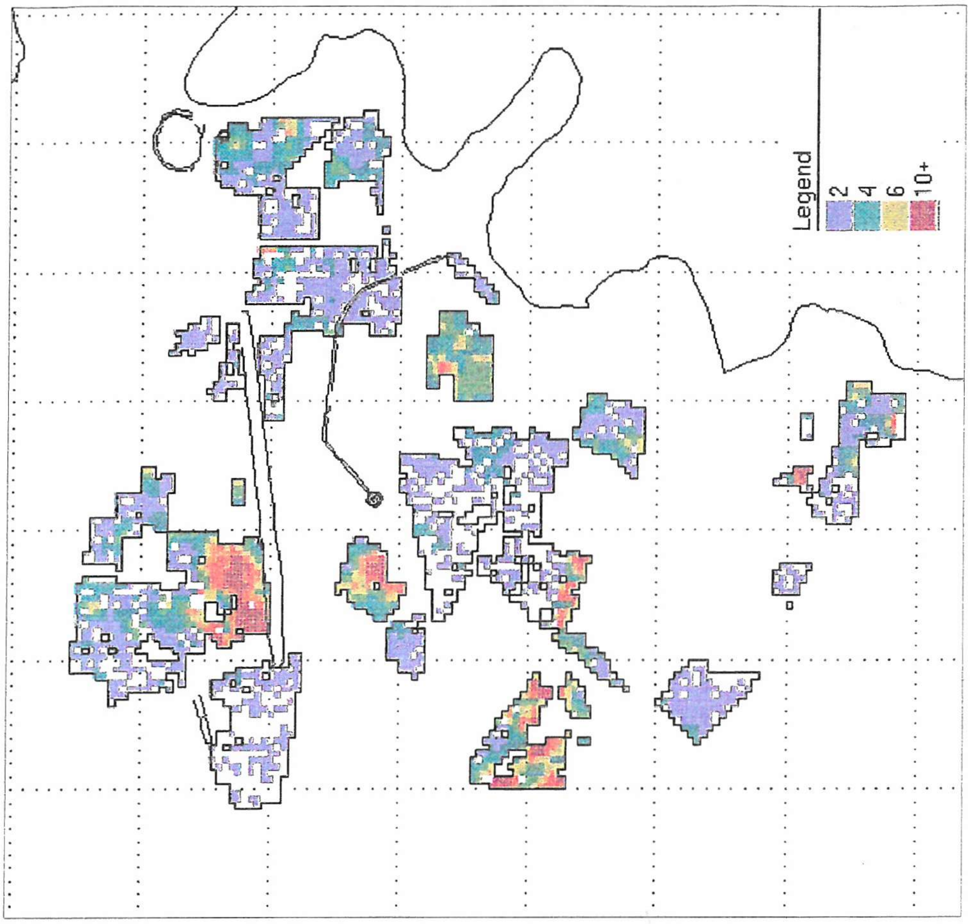


Figure 6.8 Core densities, mean filtered

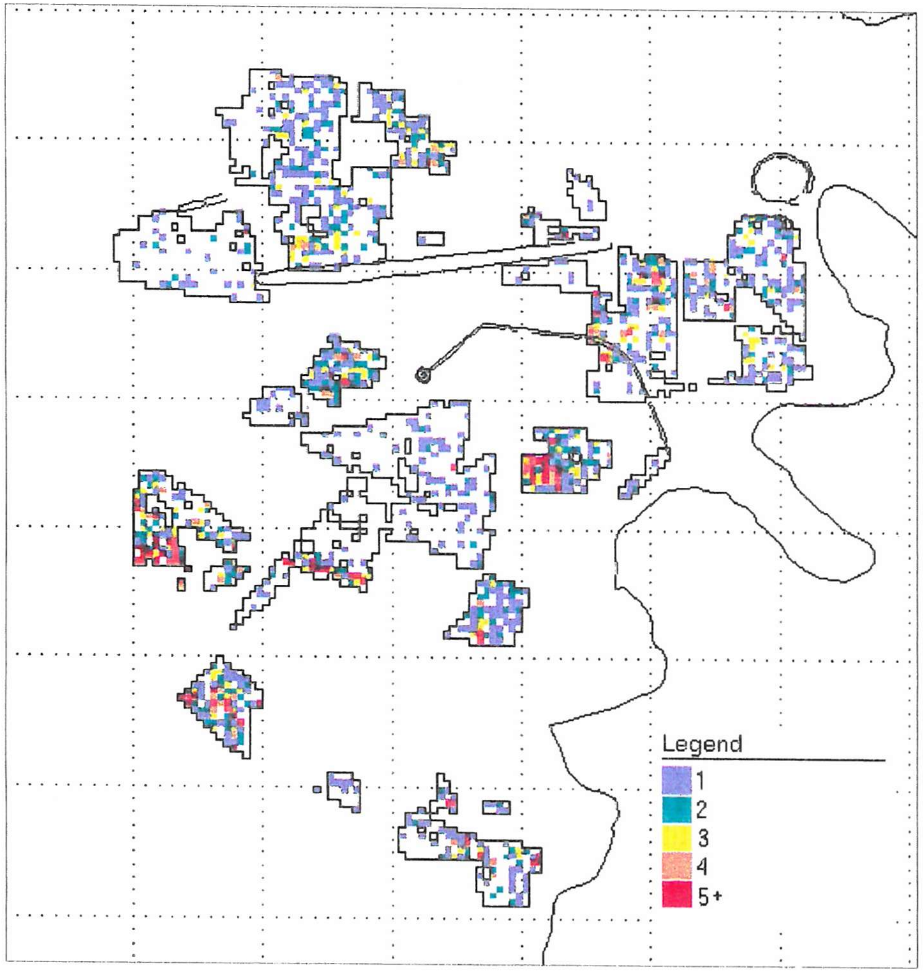


Figure 6.9. Retouched flint densities

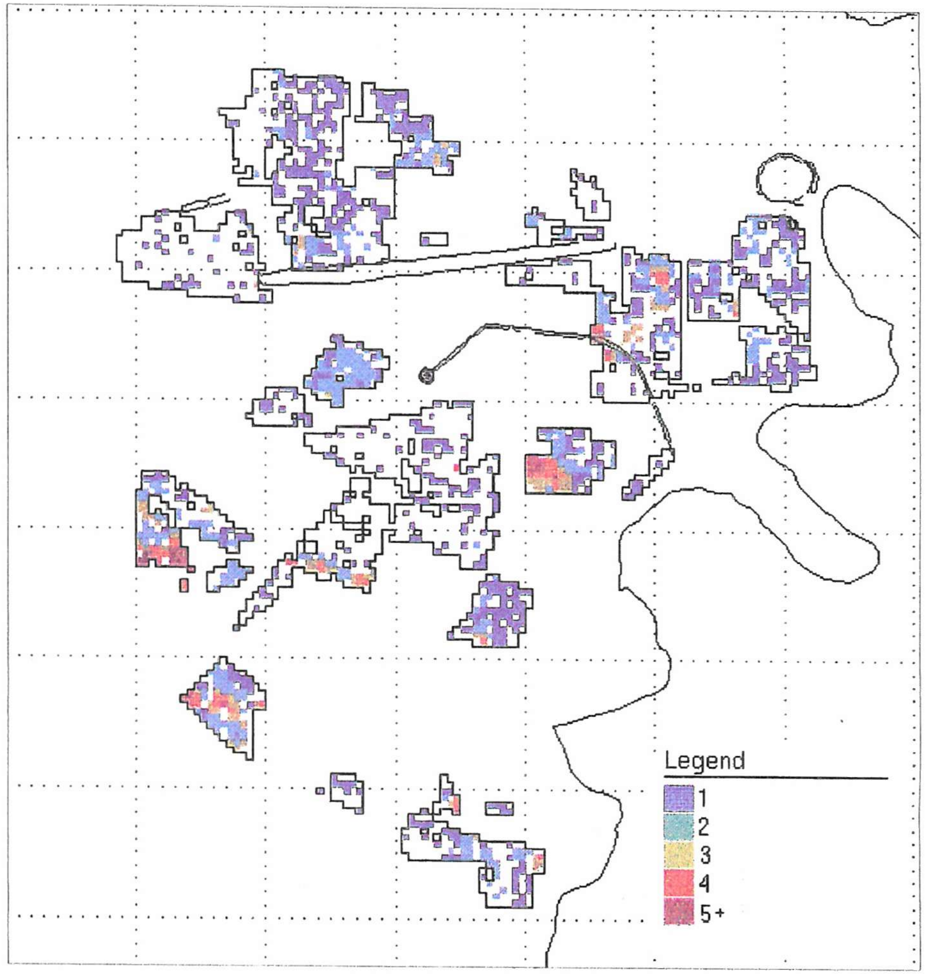


Figure 6.10. Retouched flint densities, mean filtered

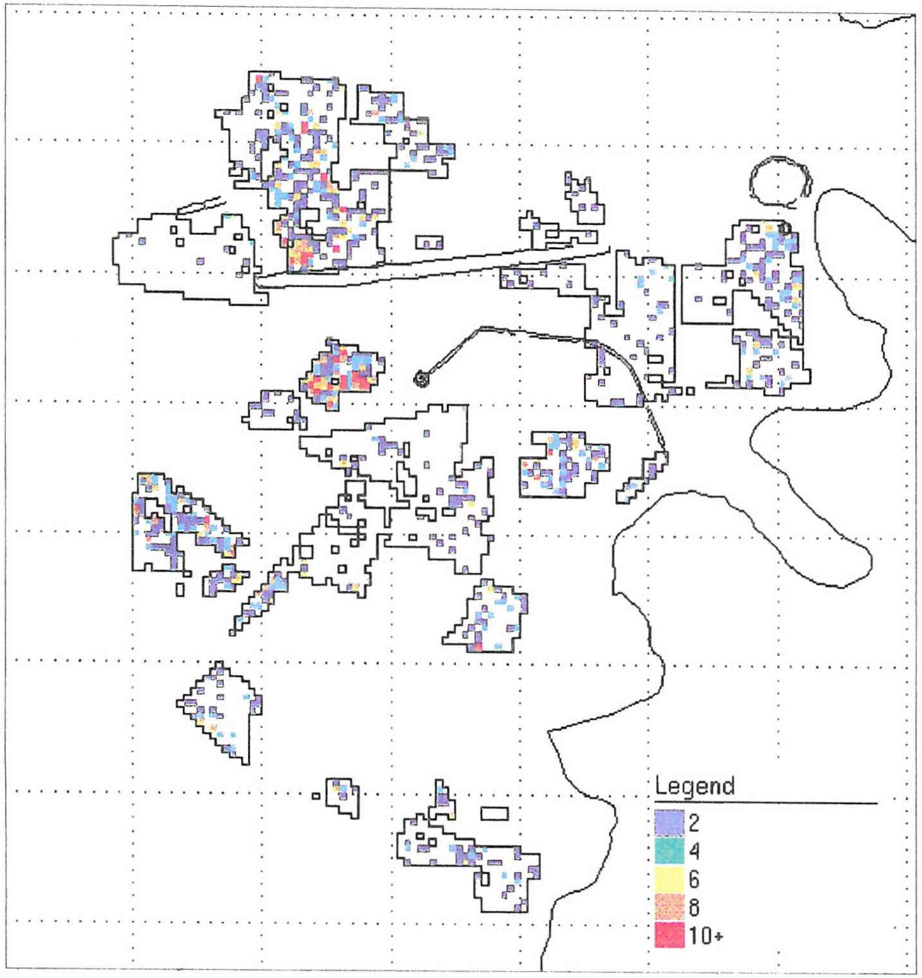


Figure 6.11. Burnt flint densities

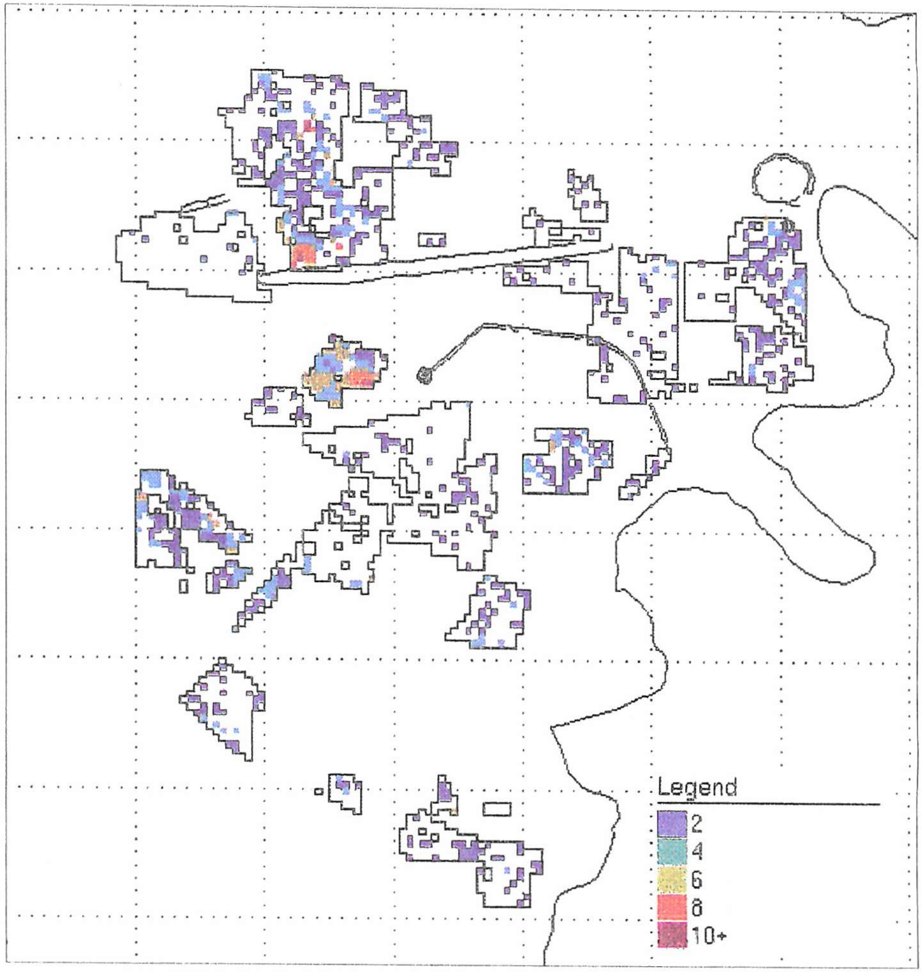


Figure 6.12. Burnt flint densities, mean filtered

While the particular values for flint density within the filtered images may therefore be less accurate in the filtered images, the general *characteristics* of the filtered images are probably a rather better representation of the characteristics of the flint density population for this area. Because of this, the mean filtered values for flint density are used in some of the following experiments instead of the raw counts. While there is no intention to mislead through the use of mean filtering of the data, it is recognised that use of such techniques to process flint densities may be problematical and controversial. Where possible, therefore, the filtered images are not presented alone and unfiltered images are presented physically side-by-side with them so that the effects may be made as clear as possible.

6.5. Spatial pattern within the flint density data

6.5.1. Autocorrelation

The first step in the spatial analysis of a data series such as this must be to determine that there is some spatial pattern to the observations: that their organisation is not the result of random processes or that the patterns present are not obscured by noise. From visual appraisal of the density maps this certainly seems to be the case, however such visual appearances have been known to be misleading. Consequently, the five data themes were first tested for autocorrelation, as measured by the Moran's I statistic using the 'kings case' procedure. The results were as shown in table 6.1:

		Normality assumption		Randomization assumption	
	Moran's I	Variance of I	z-statistic	Variance of I	z-statistic
All flint	0.6199	0.0001	60.8235	0.0001	60.9402
Cores	0.4439	0.0001	43.5634	0.0001	43.6355
Flakes	0.6207	0.0001	60.9025	0.0001	61.0203
Burnt	0.3348	0.0001	32.8608	0.0001	32.9714
Retouched	0.3572	0.0001	35.0560	0.0001	35.1356

Table 6.1. Autocorrelation within the five flint categories. In each case the expected value of I if the data were not autocorrelated is -0.003.

Autocorrelation is a measure of how good an indicator each value is of values in neighbouring units. If a high value in each unit makes it more likely that a high value will occur in neighbouring units, this is referred to as *positive autocorrelation* while if a high value in a unit makes a low value in neighbouring units likely, this is referred to as *negative autocorrelation*. The Moran's I statistic varies between 1 (for positive autocorrelation) and -1 (negative autocorrelation), although in practice never reaches either.

The IDRISI 'autocorr' program was used to obtain the level of autocorrelation within the flint walking data, using the king's case option. This tests the relationship between all eight neighbouring units, assigning a weighting of 0.7071 to the diagonals to account for their greater distance from the observation. High positive values imply a more clustered pattern than would occur if the values were random, whereas high negative values imply a more dispersed pattern than would randomly occur. The significance of the departure from the expected value of I may be tested using either of two assumptions: that the units represent independent drawings from a normal population, or that the units are randomly chosen from all possible arrangements for the specific set of cell values (Hodder & Orton 1976 p178).

In each of the flint classes, the Moran's I statistic is positive, indicating some degree of clustering within the flint density data. Under each of the assumptions (it is not clear which is the more appropriate for lithic density data of this type, although Hodder and Orton 1976 recommend the randomization assumption for most archaeological circumstances) the result is significant at the 0.001 level. This result supports the visual appraisal of the flint densities, and the result of Maskell, that there is a degree of spatial clustering in all of the flint categories. Having established this, it is possible to progress to try to explain the clustering.

6.5.2. Use and movement of flint

As an initial exploration of the data, some classifications and map algebra interactions may be used to reveal some of the spatial pattern in the flint density data itself (in section 6.7 the relationships between the flint density and other data themes are investigated). Having established that there is spatial pattern within the densities for individual categories, the central aim of this section was to investigate whether there are differences of spatial distribution between different types of assemblage.

Richards (1990 p18) defines three stages in the history of the exploitation of flint within the landscape. These can be summarised as follows:

1. Procurement and initial reduction at the procurement site
2. Reduction at site of manufacture
3. Use and discard of the artefacts

However, each stage did not necessarily take place in a different part of the landscape. For example, it is possible that in some instances, where surface nodules of flint were being exploited opportunistically, all stages from manufacture to discard could have taken place at the site of procurement. Similarly it is possible that unmodified nodules were transported, then manufactured, used and discarded in the same place. Alternatively, the artefacts may have been manufactured in one location, used in another and retained only to be discarded somewhere entirely different. There is also the possibility that, in a complex, ritually structured landscape such as that of Salisbury Plain, the discard of worked flint artefacts was neither random in nature nor driven by functional considerations. In addition, stages 1 and 3 may have activities at different places represented within them. For these reasons, it is perhaps more useful to divide these activities in slightly more detail, such that no stage could take place in more than one location:

1a. Procurement
1b. Initial reduction, without producing a finished product
2. Further reduction, i.e. manufacture
3a. Use
3b. Discard

Each of these stages may have taken place in a different location, or at the same location as the previous stage. Examination of the flint distributions makes it is clear that the distributions of flint artefacts are clustered, and this was demonstrated above through the use of first-lag autocorrelation statistics on the density values, a finding which confirms that of Maskell (1993), who tested for and found quite high levels of positve spatial autocorrelation within flint presence/absence data. In order to gain some further insight into the movement of flint during its use it is possible to postulate, as a broad generalisation:

1. That those areas with high levels of cores generally represent areas in which manufacturing-related activity dominated, in other words these represent areas in which stages 1b or 2 (reduction and manufacture) were more often undertaken than stage 3 (use and discard).
2. Conversely, that areas with high levels of retouched artefacts represent areas in which stage 3b (discard) was more common than stages 1 to 3a (procurement to use). These might be termed 'discard areas' without prejudicing any interpretation of such activity.

It follows that if there are differences between the distributions of cores and retouched pieces, it seems reasonable that this may be taken as evidence that the different stages defined above were undertaken in different parts of the landscape. If, on the other hand, there is no evidence of differentiation between the distributions of cores and retouched pieces then it may be that flint was manufactured, used and discarded generally in the same places. Preliminary inspection of the densities for cores and retouched pieces suggest that there may be a difference in the distributions. The high densities of cores visible in area 52 (north of the Cursus) do not seem to be present in the retouched data. Similarly, the apparent cluster of retouched flint in area 68 (south west of the Lake barrows) seems to have no corresponding cluster of cores.

6.5.3. Core/retouch ratios

From the above, it seems plausible that the ratio of cores to retouched flint pieces may be a useful measure of the type of flint working activity which tended to take place in a particular part of the landscape. If the ratio of cores to retouched pieces is high, it can be reasonably low, then the location was probably more often a site of discard than manufacture. The ratio between the number of cores and the number of retouched pieces can easily be calculated and mapped in the GIS using trivial map algebra techniques. The following equation can be used to generate the map:

Ratio = core density / retouch density

When applied to the SEP data, this produces an index which varies from 0.3 for areas with high retouch and low core counts through 1 for areas with equal core and retouch counts to a maximum of 28 where there are more cores than retouched pieces. Figure 6.13 represents this ratio of the (unfiltered) number of cores to the (unfiltered) number of retouched pieces. On the assumption that such a measure is equally prone to disruption from random noise as the raw counts, the mean filter operation described above was applied to figure 6.13 with the result shown in figure 6.14.

These core to retouch ratio maps seem to confirm the area of high value immediately north of the cursus (area 52), and less so towards the east of area 59. However, as has already been discussed, the cluster in area 59 correlates with unusually high values of cores and the counts of cores are in general far higher than those for retouched pieces. The effect of this may be that these ratio images are representative of the variation in core density far more than of retouch density: the variation in the measure, for example, is far greater where there are more cores than retouched pieces than the other way around. This effect can be compensated for by normalising the counts for cores and retouched pieces prior to generating the ratio maps. The maximum count for cores is 39, while the maximum count for retouched pieces is 13.

Dividing the counts for each variable by these maxima generates a pair of variables which vary between 0 and 1, and these normalised variables may then be used to generate a ratio map as above. One is added to the data values (a simple linear scaling operation to prevent the denominator or numerator being zero: division by zero is infinite), and one is subtracted from the result to compensate for this. The result is an index which varies around 0 for areas with identical 'normalised' core and flake densities, is positive for areas with higher core values than retouch, and negative for areas with higher values for retouch than core. Positive values can roughly be equated with 'manufacturing areas' and negative ones with 'discard areas'. The following equation describes this index:

Normalised ratio = [(core density/max core density)+1 / (retouch density/max retouch density)+1]-1

The mapped results of this operation are shown in figures 6.15 (unfiltered) and 6.16 (mean filtered as above), showing zero as white, positive values as red and negative values as blue. In practice the index varies between about -0.5 for 'discard areas' and 0.5 for 'manufacture areas'.

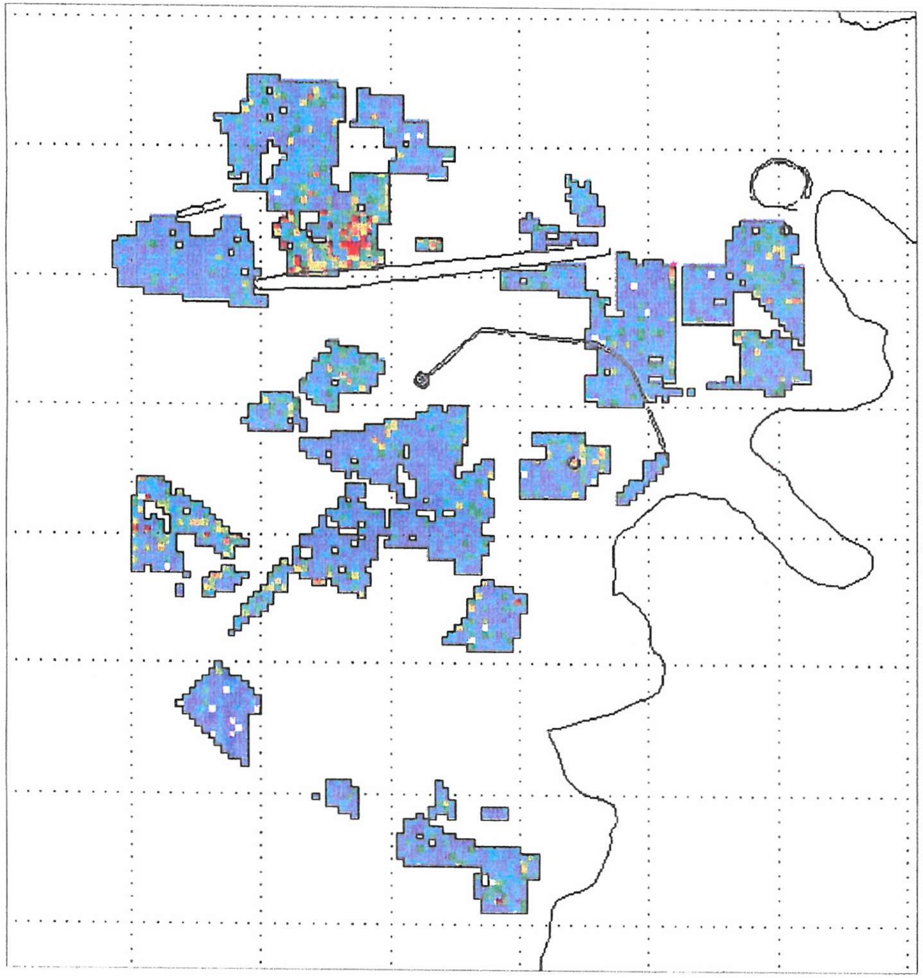


Figure 6.13. Core : retouch ratio

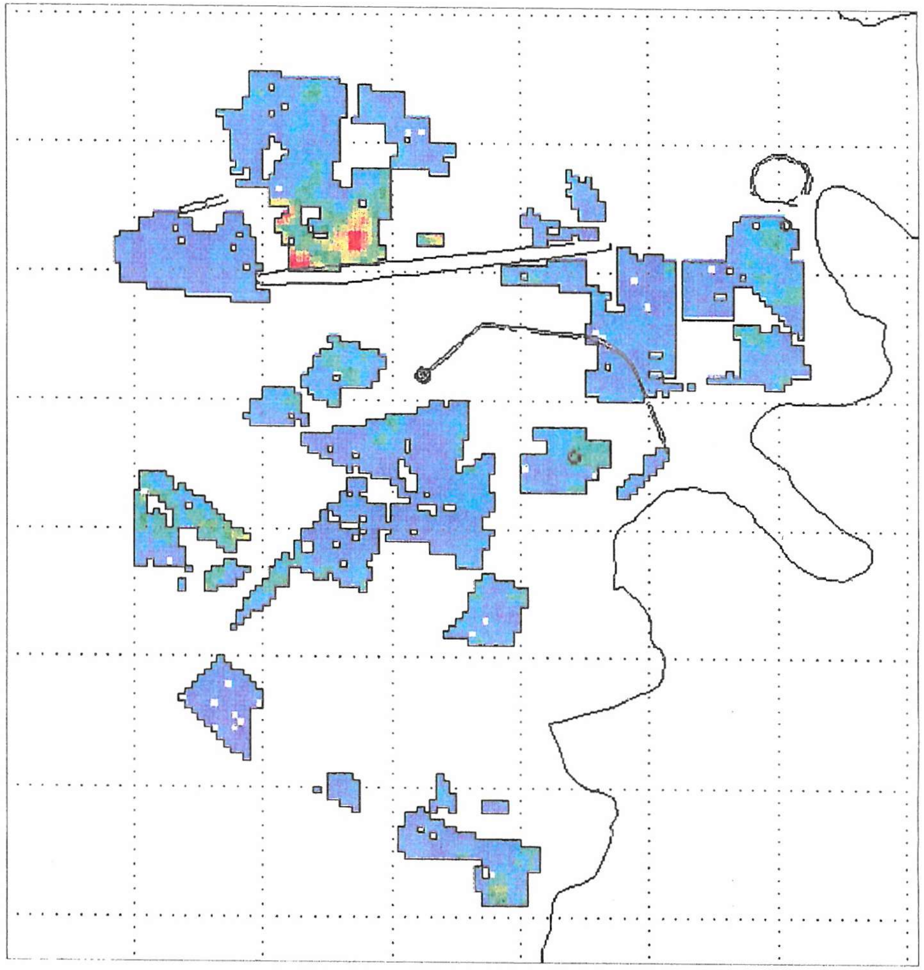


Figure 6.14. Core : retouch ratio, mean filtered

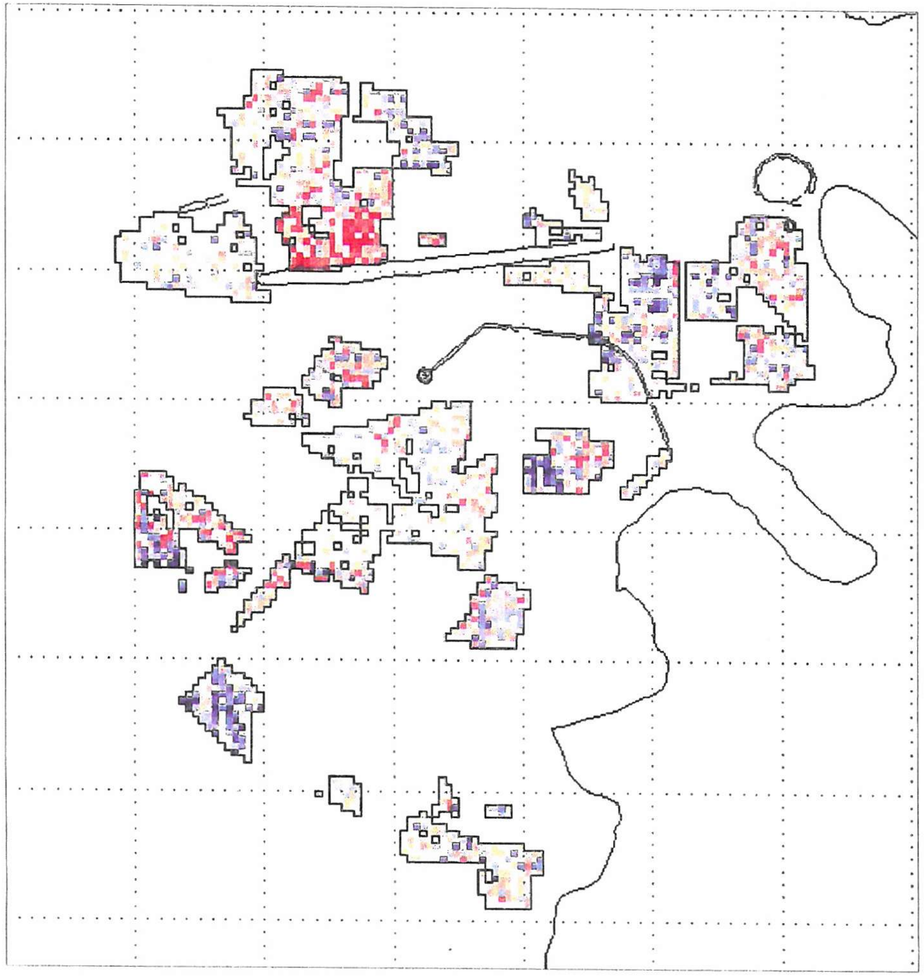


Figure 6.15. Normalised core : retouch ratio

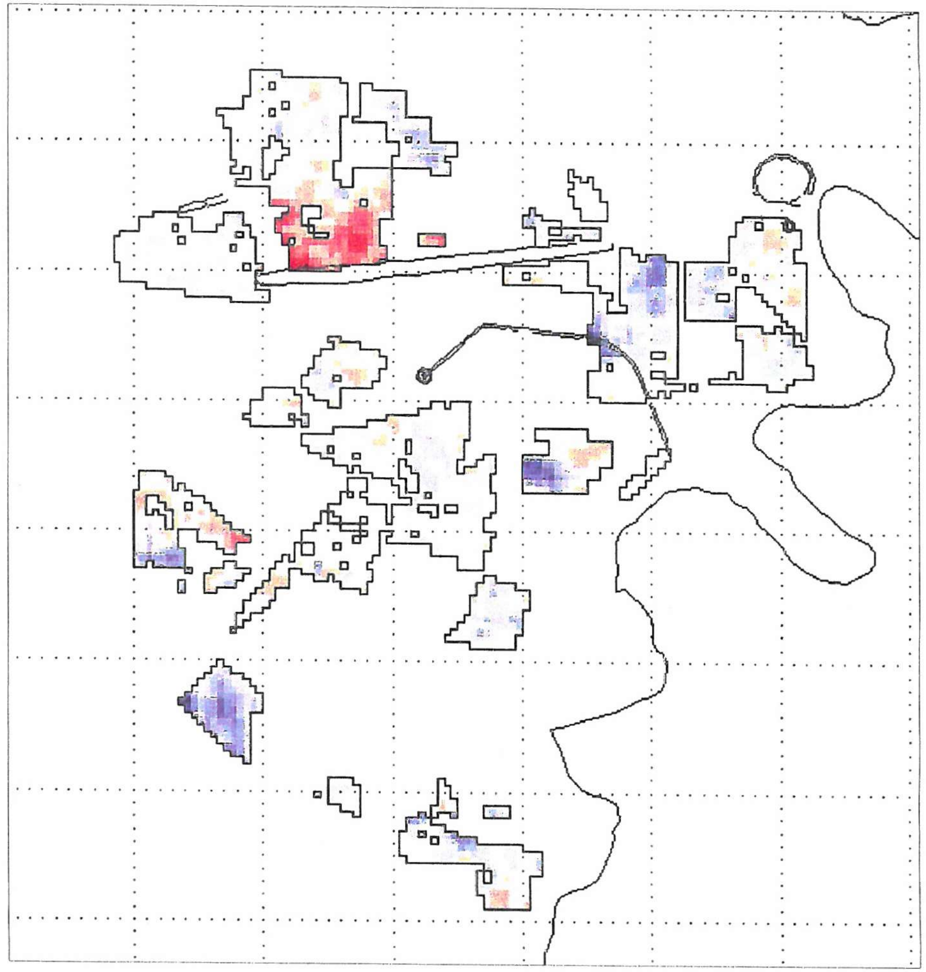


Figure 6.16. Normalised core : retouched ratio, mean filtered

These maps clearly confirm that the area of high core/retouch ratio north of the cursus (area 52), which is visible in the un-normalised ratio maps, almost certainly reflects more than simply the high core counts in that region. It is difficult not to infer that this area was used as a location for manufacturing activity in preference to discard either intensively at some time, or less intensively over a long period.

As well as this area of positive value, however, the distribution of this normalised measure reveals several clusters of negative value which are not revealed by the un-normalised ratio. These occur south west of the Lake barrows (area 68), adjacent to Coneybury henge (area 51), around the avenue just west of the second 'elbow' (areas 70 and 87) and again just northeast of there (area 57). It is obviously tempting to interpret these as 'domestic' sites, but it should be remembered that what is actually revealed are simply areas in which, over a period of time, more discard activity took place than manufacturing. They may indeed be domestic sites of some type, but equally they may represent deliberate discard of artefacts at these places as part of ritual or ceremonial activities, rather than domestic. The high 'discard values' around the Avenue are particularly suggestive in this respect.

Interestingly, the area of consistently high absolute counts of flint in the dry valley from Winterbourne Stoke Crossroads to Normanton Bottom (areas 50, 59, 77, 75 and 67) does not show very high values on this index. Whereas the low values for the index on Normanton Down obviously reflect genuinely low levels of activity, the indeterminacy in the dry valley where high levels of deposition did occur does not and might be interpreted in either of two ways.

Bearing in mind that this area is a dry valley, and that there may have been a considerable amount of colluviation since the neolithic and bronze ages (Bell 1982, 1983, 1985, Allen 1988), a taphonomic cause cannot be dismissed. It is possible that distinct depositional activities on either side of the dry valley have become obscured by movement of flint within the ploughsoil resulting in the observed less-than-clear pattern. Alternatively, however, it is possible that the observed 'mixed' assemblage is representative of the activities which took place within the valley, and that the location was used for both manufacturing and discard activity.

6.6. Interpolation of values at unsampled locations

6.6.1. Spatial interpolation techniques

So far the presentation and discussion of the flint data has been entirely confined to the areas which were fieldwalked. No attempt has been made to extend knowledge of the density of flint into areas which were not surveyed. This section will attempt to explore the utility of GIS in this area, and experiment with different methods of 'guessing' values of flint density at unsampled locations. As discussed in detail in chapter 2, there are two main approaches to the problem of interpolating unknown values from known ones: spatial interpolation techniques and predictive modelling. This section will explore the first of these options and attempt to use spatial interpolators to extend the observed distribution into unsampled areas of the landscape. Inherent in this approach is the assumption that the sampled locations exhibit coherent spatial patterning of some type: either clustered or dispersed. An approach based on interpolation is therefore only valid if it can be shown that the data exhibit autocorrelation. This was shown above in section 6.5, where the existence of positive autocorrelation (clustering) within the flint density data was demonstrated.

It should be noted that the aim of spatial interpolation is not the prediction of patterns (such as flint clusters) which have not been identified from the field survey, although this can be attempted using predictive modelling (see section 6.7). Instead interpolation assumes that the pattern which is observed within the survey data persists outside the surveyed areas, and the aim is to try and extend the observed pattern into unsurveyed areas. A variety of interpolation procedures are possible, some are available within particular GIS while others have not yet been incorporated into commercial GIS packages. Techniques for interpolating surfaces from point data include polynomial trend surface analysis, linear and non-linear contouring, topographic interpolation with splines, inverse distance weighting and Kriging. Ideally, the evaluation of each result would be through comparison of the interpolated values with new observations, taken by further fieldwalking. Unfortunately this was not possible in this study. However, in attempting to evaluate these techniques for the particular purpose of interpolating flint density a number of criteria can be assessed.

- *Plausibility*: the general appearance of the result may be evaluated in order to assess whether the result is likely or unlikely. This is inevitably subjective in nature, and presupposes that the observer has a 'gut feeling' for how the densities should be interpolated.

- *Fit*: for interpolators which do not constrain the result to pass through data points, it is possible to observe how far the interpolated result deviates from the known values at the sampled location. This is analogous with the use of residuals in regression analysis.

- *Portability*: all interpolators, however good, will provide worse estimates of unknown values as the unknown location becomes further from sampled ones. Interpolations and models may therefore be assessed in terms of how rapidly the results become unusable as this distance increases.

Polynomial trend surface

One possible approach to this problem is to fit a polynomial equation to the existing data points, and to use this equation to predict the values of the flint at the unsampled locations. This technique is generally referred to as 'trend surface analysis' (Hodder & Orton 1976) and has been applied to flint density data by, for example, Feder (1979). Trend surfaces are three-dimensional regressions: where a linear, two dimensional regression generates an equation to relate the value of a dependent variable y with the variation in x, a trend surface attempts to relate values of an dependent variable z with variation in both x and y.

Unfortunately it is a characteristic of trend surfaces that, in order to model a surface of any complexity, high orders of polynomials are required so that the calculation of a good fit rapidly becomes very time consuming. Polynomial trend surfaces are also wholly unconstrained, and in order to obtain the best fit of the surface at the data points, they may deviate dramatically from the expected result in the unsampled locations. If low order polynomials are used, then the deviation to maintain the fit is rarely vast, but where the pattern within the data is of any significant complexity then the surface is a very poor representation of the real pattern. Conversely where high order polynomials are used, necessary to obtain a reasonable fit to data of reasonable complexity, the deviation from the data range is frequently extreme. For these reasons it was not felt that further exploration of this approach would be productive.

Contouring

Although widely applied to a variety of exact and approximate interpolation procedures, the term contouring will be used here to refer to a specific approach to interpolation. In this context, contouring involves generating a triangulation from all the known data values to form a triangulated irregular network (TIN) which may then be used to describe a continuous surface. Where locations lie outside data areas but within the triangulation then their values may be obtained from the value of the surface at that location. When unknown values are assumed to lie on flat triangular facets, then the contouring operation is linear. Linear contouring behaves in a predictable manner, because values between the data points must lie within the range of the three nearest known values. It produces a 'faceted' model which is continuous, the derivatives of which, however, are likely to be non-continuous.

If small polynomial 'patches' are used instead of linear calculations based on triangles to maintain the derivatives of the surface at the junctions between the facets, then the contouring is referred to as non-linear. Non-linear contouring provides a surface whose derivatives are continuous (the higher the order of the patch, the higher the order of derivatives will be continuous) but removes the constraint that the unknown values must lie between the three defining data points.

In either case, unlike the fitting of a polynomial trend surface, the result is constrained to pass through the existing data points, and therefore the fit of the surface to the data is inherently perfect. Unknown values are also wholly defined by the closest three known values and are unaffected by other data points. In the case of the SEP data, this means that the values between the sampled locations will be entirely defined by the values at the edges of the sampled areas. Values not at the edge of the sampled areas will not affect the interpolated values.

Figure 6.17 shows the result of applying a simple linear contouring to the values for total flint (unfiltered). One characteristic of this approach is that the values interpolated outside the convex hull defined by the edges of the sampled locations are wholly unreliable, and these have been excluded. Such values are extrapolated values, rather than interpolated, being calculated by extending the slope of the last facet outwards. These values are consequently defined entirely by the edges of the sampled regions, with no further variation possible and the unreliability of the result is therefore to be expected.

Values inside the convex hull, but outside the sampled areas generally seem plausible. The area of high flint density north of the cursus (area 52) is interpolated as continuing south towards the area north of Stonehenge Down (area 54) as would intuitively be expected although the effects of the triangulation are clearly visible in the form of the flint cluster produced. Figure 6.17 (B) shows the characteristic triangular cluster produced in this way. Occasionally, however, the interpolated values seem to be seriously affected by anomalous values on the edges of the sample areas. Figure 6.17 (A) shows an area of very high interpolated value. Comparison with the original density map reveals that this is the product of a single high density value on the edge of the sampled area (area 67). Such reliance on edge values is manifestly undesirable in this context, particularly as the edge values themselves may be the least reliable field samples as a result of reduced line length at the edge of a field, trampling, shade, variable ground condition and other factors which may exhibit greater effects for edge samples than central ones. In this case the edge values are also likely to be more prone to 'noisy' spikes and troughs because the procedure used to normalise the values for 50m cells (see above) derived many of the edge cells from only one 50m transect, while the central ones were mostly derived from two.

One method of countering this reliance on edge values in linear contouring is to contour the mean filtered values rather than the original counts. This accepts that the original flint observations are 'noisy', and should reduce the impact of unrepresentatively high densities at the edges of regions. The result of this is shown in figure 6.18. This clearly reduces the impact of the single high edge value (A) without having significant impact on the interpolation of the values between area 52 and 54 (B). The triangulation effects inevitable from a linear faceted model area still readily apparent (B).

Non linear contouring was also attempted for the same two sets of data (raw total flint counts and mean filtered total flint counts). In both cases the unconstrained nature of the result was problematical. When allowed to extrapolate beyond the convex hull of the sampled areas, the non-linear contouring of the raw counts produced values of up to 3800, and frequently produced negative interpolated values of up to -2000. The mean filtered flint counts produced less dramatic results, variation was from around -1000 to around 1000. In neither case, therefore, can the interpolated values be regarded as reliable. Images of the result within the convex hull in each case are shown in figure 6.20 and show the large areas of predicted negative and very high density caused by 'folds' in the interpolated surface.

Numerical approximation with distance weighted average

An alternative strategy for interpolating values between locations is to use approximation procedures instead of exact interpolators such as contouring. These do not constrain the result to pass through the data points, but instead use the points to approximate the values within both sampled and unsampled locations. The most widely applied approach is that adopted by Shapiro (US Corps of Engineers 1993) within the GRASS GIS, referred to as *inverse distance weighting*, and is similar to transformations available within a number of other systems.

This approach involves calculating values for all locations within the landscape based on a fixed number of nearby points: usually between 10 and 20. On the principle that unknown values are more likely to resemble near values than distant ones, the values of these points are weighted in inverse proportion to their distance from the unknown value and these weighted values are then averaged to produce an estimate of the value at the unknown location.

The effect is a simple, generally robust approximation procedure for interpolation of surfaces from a wide variety of point data. Shapiro (US Corps of Engineers 1993) considers that :

'In comparison with other methods, numerical approximation allows representation of more complex surfaces (particularly those with anomalous features), restricts the spatial influence of any errors and generates the interpolated surface from the data points. It is the most appropriate method to apply for most spatial data'. Using 12 control points as recommended by Shapiro (US Corps of Engineers 1993 p427), the result applied to unfiltered total flint density (obtained using the GRASS s.surf.idw routine) is shown in figure 6.21.

Interpolation by this method of mean filtered values was not appropriate as the method itself includes an element of mean smoothing, the degree of which is controlled by the number of neighbouring points used in the calculation. This therefore inherently performs the same noise-reduction function as the mean filter. The result of applying the inverse distance method to the total flint densities may be gauged from figure 6.21.

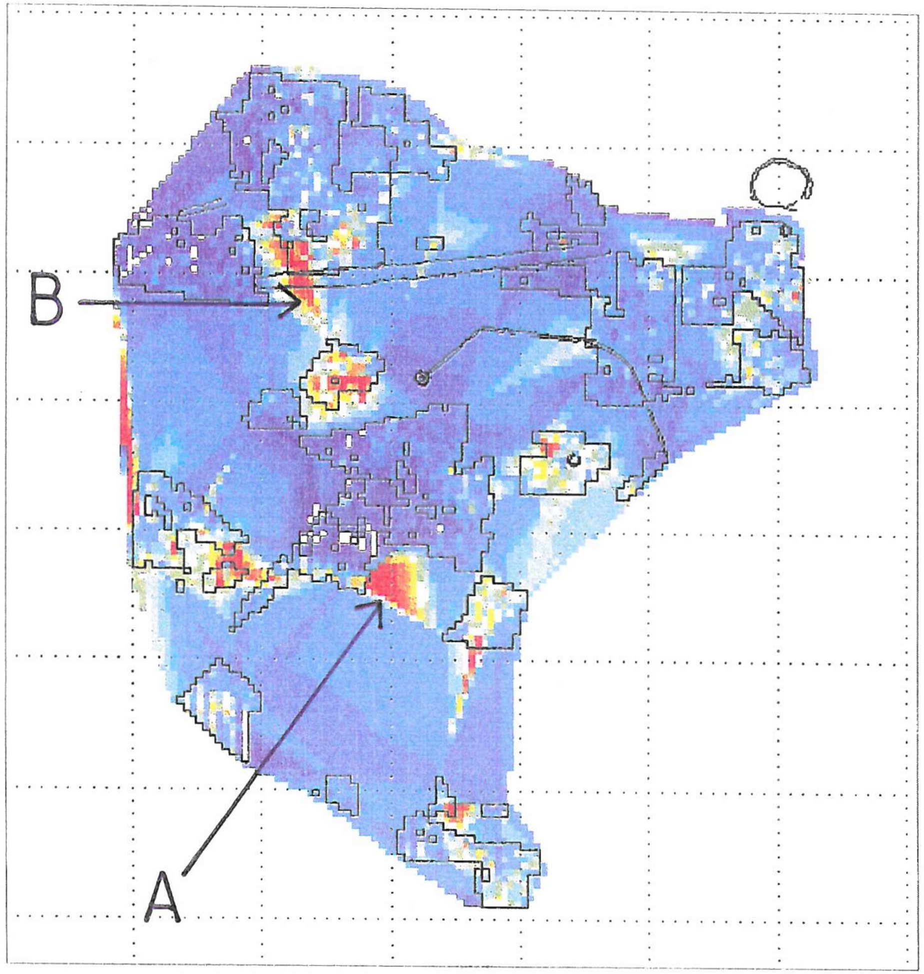


Figure 6.17. Linear contouring of total flint densities

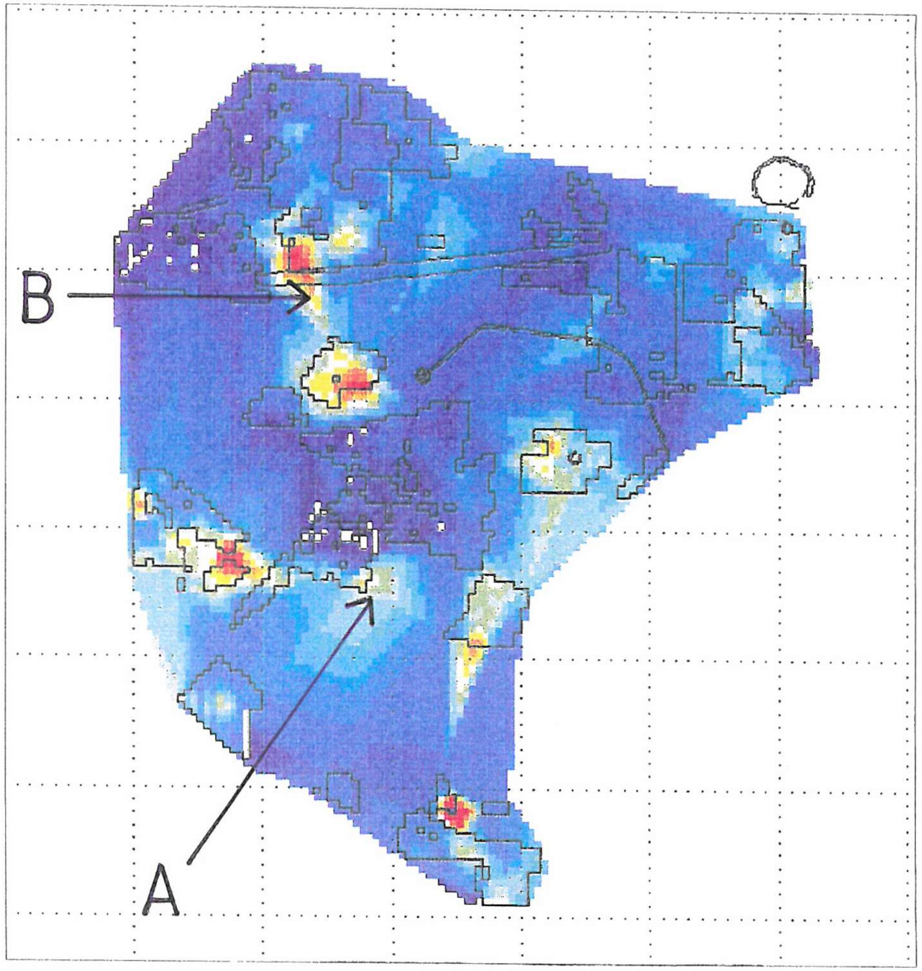


Figure 6.18. Linear contouring of mean filtered total flint densities

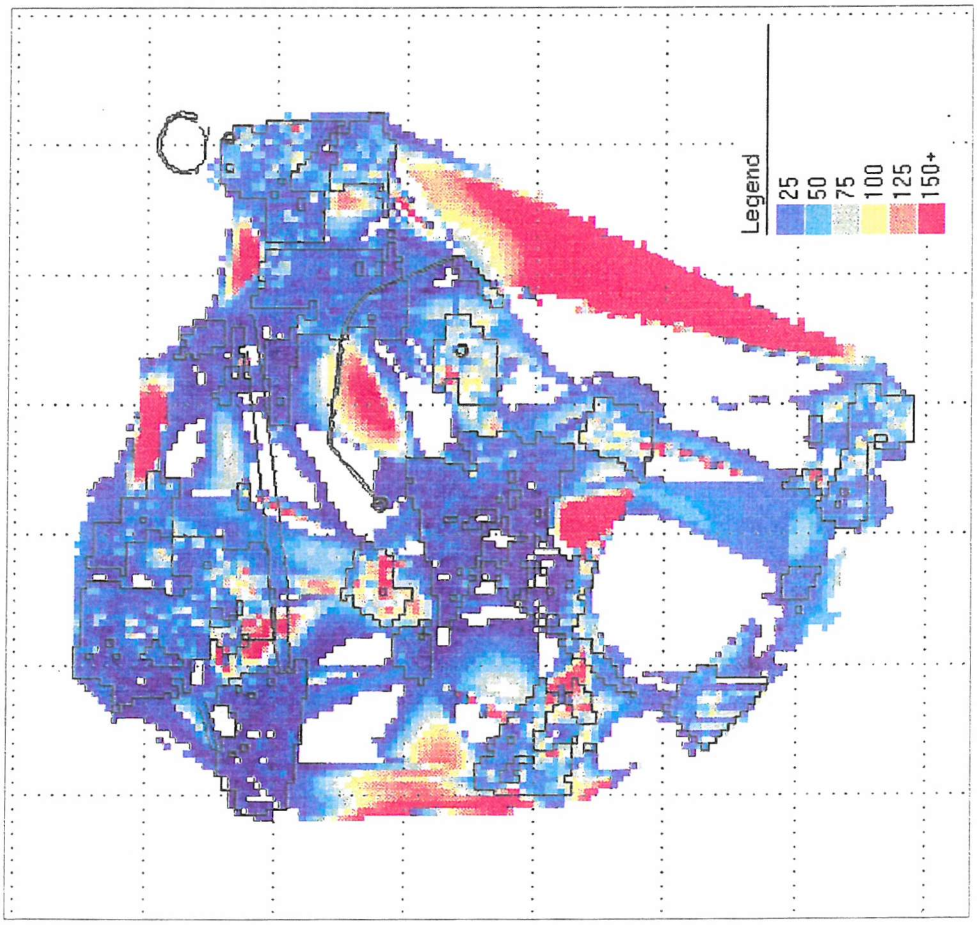


Figure 6.19. Non-linear contouring of total flint densities

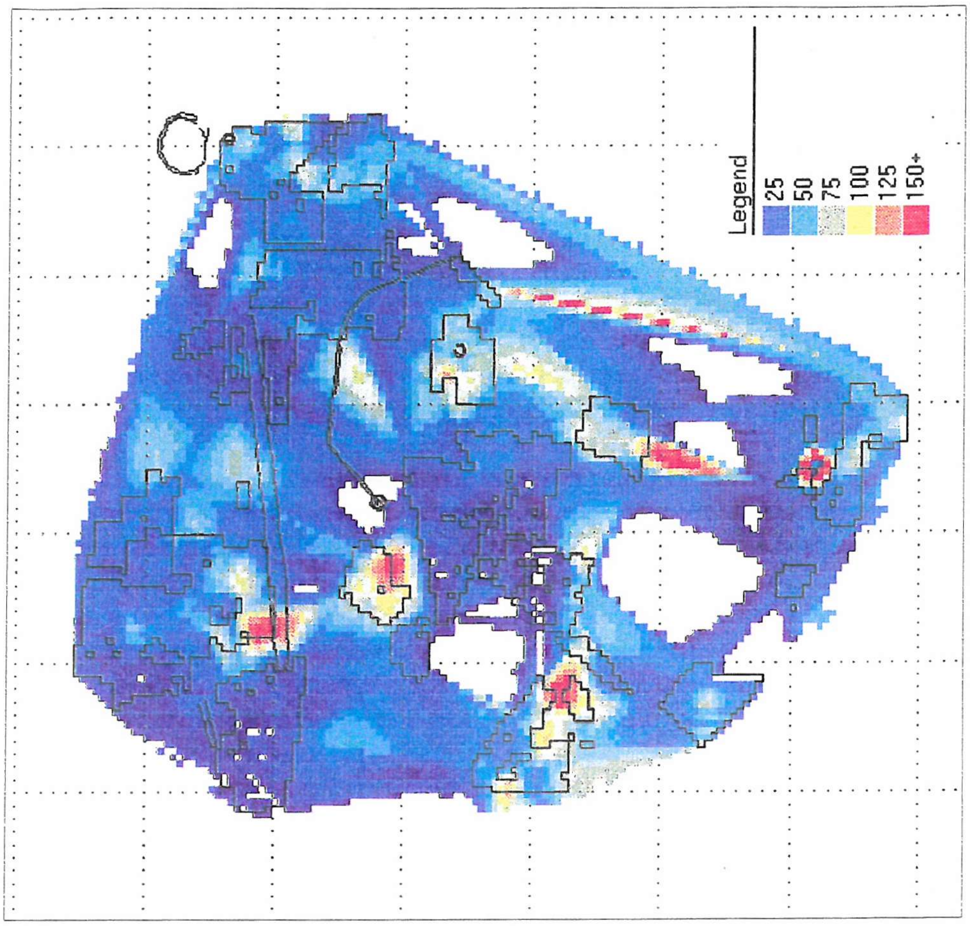


Figure 6.20. Non-linear contouring of mean filtered flint densities

It is not necessary, however, to restrict this form of interpolation to simple measures of lithic density. It has already been seen above that the normalised core/retouch ratio exhibits clustering. It is therefore appropriate to use an interpolation - if the ratio is interesting, then the extents of the areas indicated by it must be worth exploring. Figure 6.22 shows the result of using an inverse-distance interpolation on the normalised core/retouch ratio. The image is interesting in that it suggests a tripartite division of the landscape: the area of positive core/retouch ratio at the western end of the Cursus contrasts with several areas of negative value to the south east of that, seemingly distinguishing areas consistently used for manufacture from areas used for discard. The area to the southwest of the study region, however is rather more mixed in character and may represent either a mixture of activities, or alterations to the use of the area through time.

6.6.2. Contouring and weighted averaging compared

Comparing figure 6.21 with the figures 6.17 and 6.18, it is apparent that the artefactual triangular facets produced by linear contouring are not present in the distance weighted product, and that the character of the flint scatters is therefore more plausible. The (probably spurious) area of high density in the contoured figure 6.17 is not present in the inverse-distance product 6.21, reflecting that the inverse-distance method of interpolation is far less prone to dramatic fluctuation caused by edge effects. Inverse distance interpolation therefore seems to provide a more plausible estimate of the shape of lithic scatters in that the areas predicted appear to be more natural, the triangular facets of the contouring are clearly an artefact of the method, and do not reflect the likely form of the lithic scatters in unsurveyed areas.

Use of linear contouring provides the best fit of interpolated surface to data, although the necessity to pre-filter the observations to reduce the edge effects clearly reduces that as the sole benefit of that approach. Linear contouring of the data without filtering would therefore seem to be a poor technique in this regard.

It is clear that the fit of interpolated surface to data must be balanced against the desire to smooth to reduce noise and in this regard the inverse-distance method has clear advantages. The number of points used in the method may be varied to add smoothing at the expense of fit: increased smoothing is obtained by increasing the number of points considered, forcing the algorithm to provide greater generalisation. In the context of these lithic scatters, experiments with a variety of values (not shown) suggested that the recommended value of 12 seemed to provide an optimum balance.

Use of neither technique provides an estimate of lithic density which is portable, in the sense that it may be extended beyond the general area of the survey. The coherence of the products of

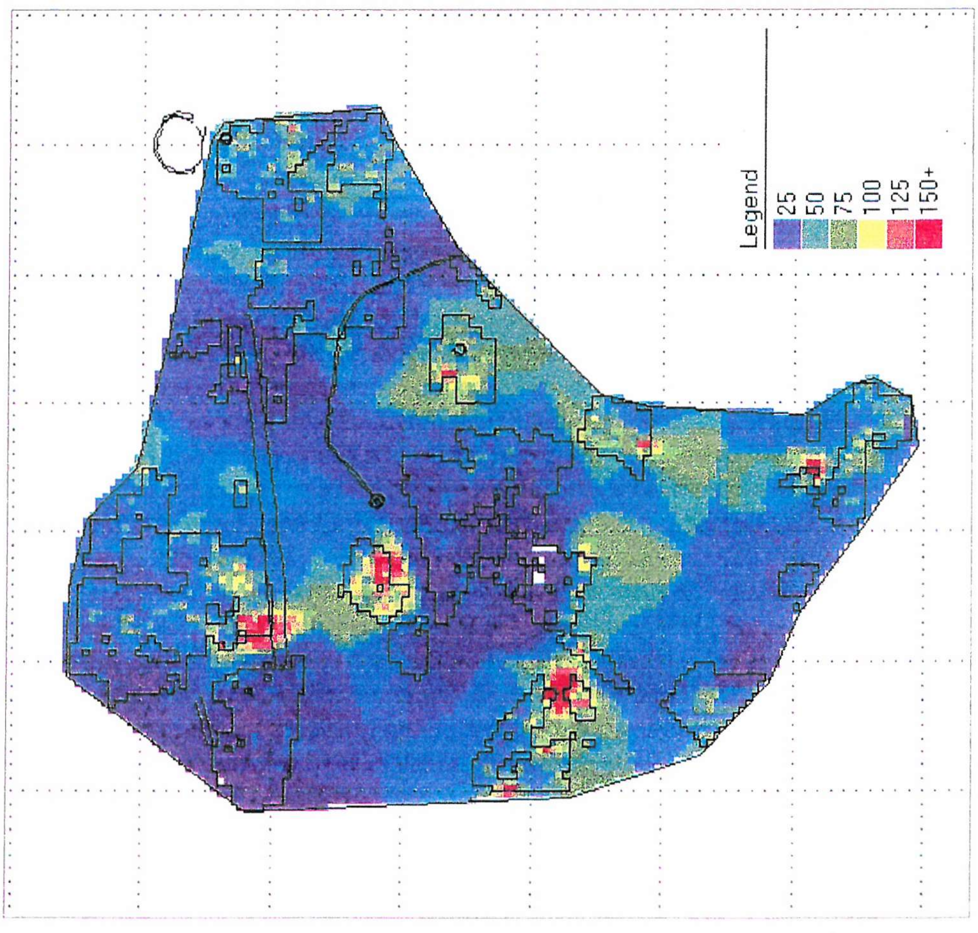


Figure 6.21. Inverse distance weighted averaging of flint density

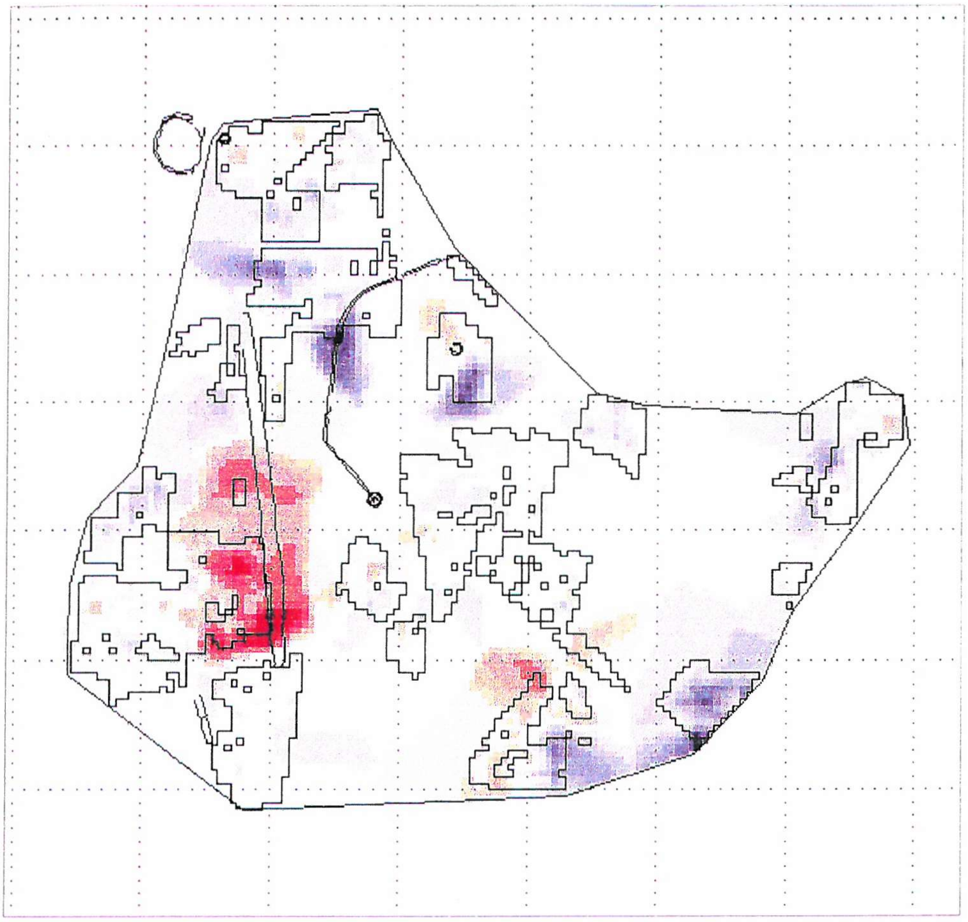


Figure 6.22. IDW averaging of normalised core : retouch ratio

both contouring and inverse-distance interpolation seems to be highly localised, and must be restricted to interpolation within the convex hull of the data points unless there are extremely good reasons to expect extrapolation beyond this area to be of any value.

6.7. Use of relationships with other data

6.7.1. Prediction from correlates

An alternative approach to estimating values at unknown locations is to use the correlations of the observed variable with other variables. In situations where the values of these correlates are known for unknown locations, then this may be used to predict the values of the observed variable outside sampled locations.

Approaches based on this principle are collectively known as predictive modelling methods and the history of these in archaeology, was discussed in detail in chapter 2. It is the purpose of this section to attempt the application of a predictive approach, which was primarily developed within the context of North American archaeology (e.g. Allen *et al* 1990) and for the understanding of dispersed mesolithic settlement sites (e.g. Kvamme & Jochim 1989) to the context of Salisbury plain. As an initial step, therefore, a model based on linear multiple regression was constructed.

One of the requirements of linear multiple regression analysis is that the dependent and independent variables are approximately normally distributed. If they are not, then appropriate transformations such as logarithmic or square root transforms must be applied to generate new, normally distributed variables (Shennan 1988). Another requirement of the process is that the variables are of an interval scale or above. In order to use categorical data within a regression, the use of dummy variables is necessary. The character of each of the variables, both dependent and independent, must therefore be carefully considered before they are used in the analysis.

6.7.2. What to predict?

An exercise such as this cannot hope to predict accurately the precise values of lithic density at unsampled locations. The aim is, instead, to try and observe the general pattern of variability of a dependent variable within the sampled areas, and to extend this general pattern into areas which have not been sampled. In order to use a multiple regression approach, it is also desirable that the dependent variable provides a reasonable range of variation.

For these reason, the filtered values of lithic density will be used as the dependent variable because it is held that these more closely represent the true form of the population from which the samples were drawn. Densities of retouch, burnt flint and cores provide rather low ranges of values which are therefore more prone to the effects of chance in sampling. The difference between the total flint densities and total flake densities is minimal, and therefore the experiment will attempt to predict the filtered total flint densities. An alternative approach to prediction will be described below, which may be more appropriate for variables with lower ranges and high numbers of zero observations such as retouched flint, burnt flint or cores.

Frequency distributions for all the possible dependent variables are shown in figure 6.23. Each of the histograms reveals the skewed nature of the flint densities, while the histogram to the right is of a natural logarithm transform of each variable (each variable was also incremented by one before the log transform to avoid generating any infinite values. This is a linear scaling which will therefore have no effect on the regression).

The histograms reveal the lack of range in cores, retouched and burnt flint densities which prevents the transformation from having any beneficial effect on the distribution. Most notably, there are extremely high values on the left tail of the distribution, reflecting the high number of samples with no values. The transformed distributions for total flint and flake densities, however, are markedly different and although both exhibit a minor skew to the right and a suggestion of bimodality the distributions are apparently quite close to normal.

This is reflected in the summary statistics for these variables which show the skewness statistic reduced from 1.76 to -0.8 for total flint, from 1.8 to -0.8 for flakes by the log transformation, but by noticeably less for burnt and retouched flint. The reduction in the skewness of cores by transformation from 1.8 to 0.1 seems high, although visual inspection of the distribution suggests that this may be misleading.

6.7.3. Selection and character of predictors

Landform indices

A variety of possible environmental variables may be used to predict archaeological variables. There are good reasons for expecting significant correlations between lithic densities and environmental variables. Schofield (1988) used analysis of variance and regression techniques to examine the correlation of flint densities with geology, soil, topography and proximity to river within the Meon and Avon valleys and demonstrated that such correlations were generally present. In the Meon valley, for example:

> *'Clear distinctions could be made between types of activity and intensity on the lower and upper chalk, while other relationships also appear to have existed in relationship to topography and soil type' (Schofield 1988 p355)*

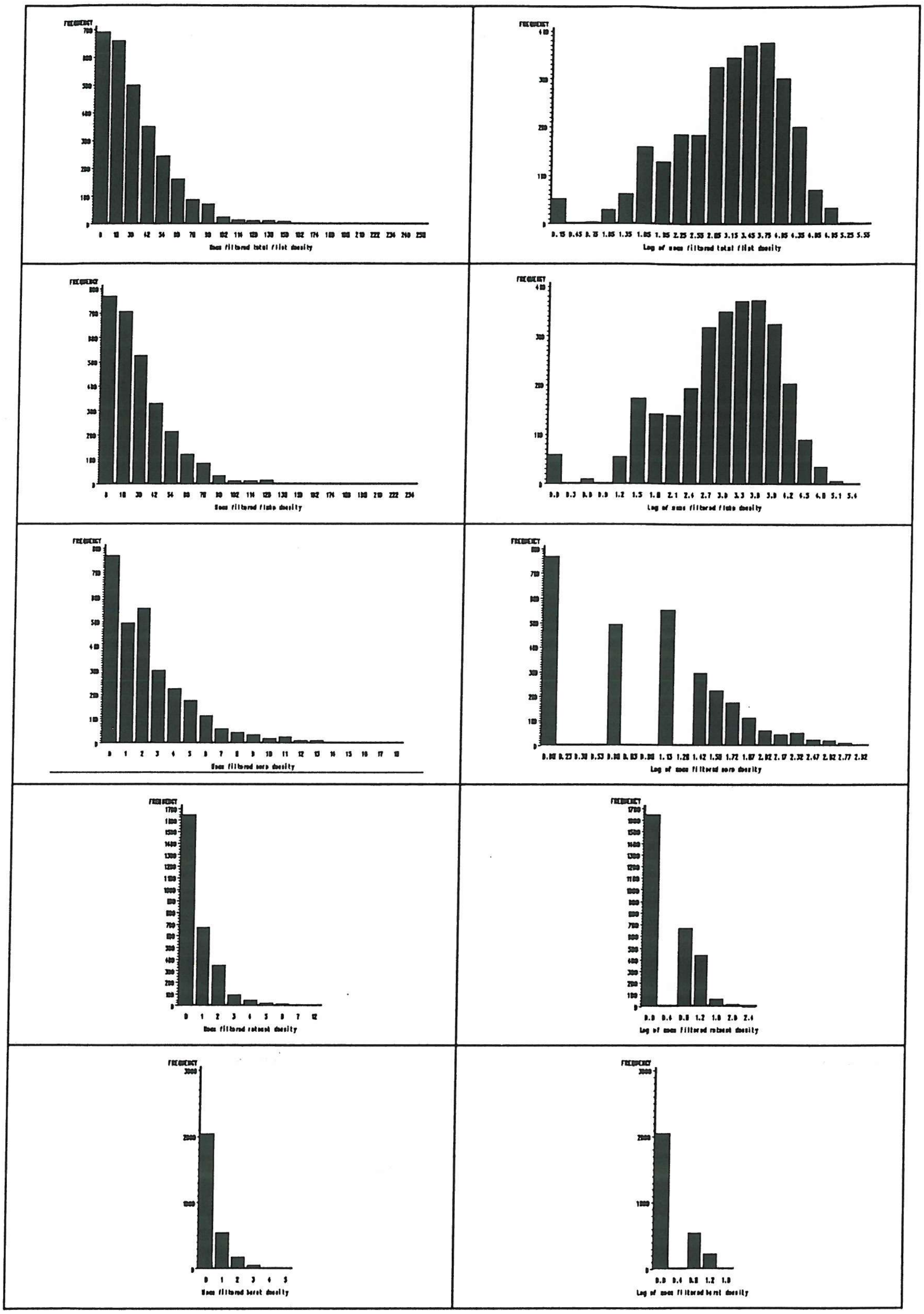


Figure 6.23. Frequency distributions of mean filtered flint classes (left) and the natural log of mean filtered flint classes (right). Top to bottom the classes are total flint, flakes, cores, retouched pieces and burnt flint.

Soil, geology and topographic measures were therefore selected for inclusion within a predictive modelling experiment. An indication of their spatial characteristics may be gauged from figure 6.25 (a) to (g), which shows a 'thumbnail sketch' of each independent variable. In more detail the indices used in the regression are as follows:

Topography

The derivation of elevation matrices was discussed in principle in chapter 2 and described in practicein chapter 4. Elevation, measured in metres above sea level, was introduced as the first predictor variable.

Figure 6.24, top left, shows the frequency distribution of the elevation values. Although not an ideal distribution, with some evidence of skewness, the histogram suggests that elevation approximates to a normal distribution and that a transformation would not be appropriate. The skewness measurement for this variable of -0.5 reflects this, and is not unacceptably high.

Aspect was derived from the elevation model, as discussed in chapter 2. This produces an index varying between 1 and 360 degrees of aspect, with values of zero representing areas for which no aspect could be calculated, in other words flat areas. The frequency histogram for aspect reflects the character of the variable, with an anomalously high value for zero, and then some variation of the number of land units over the aspect range.

Although seemingly close to a normally distributed variable in the histogram, this is therefore misleading. It may be possible to exclude all observations with aspect zero from a regression and to use aspect as an interval level measurement, but this would represent a systematic distortion of the sample. For this reason, aspect was divided into four dummy variables, NE (between 1 and 90), SE (91 to 180), NW (181 to 270) and SW (271 to 360), each of which recorded 1 for observations which fell into this category, and zero if not.

It should be noted that this approach leaves a fifth implicit category of 'no aspect' within the analysis, and consequently the representation of aspect in this variable is slightly compromised.

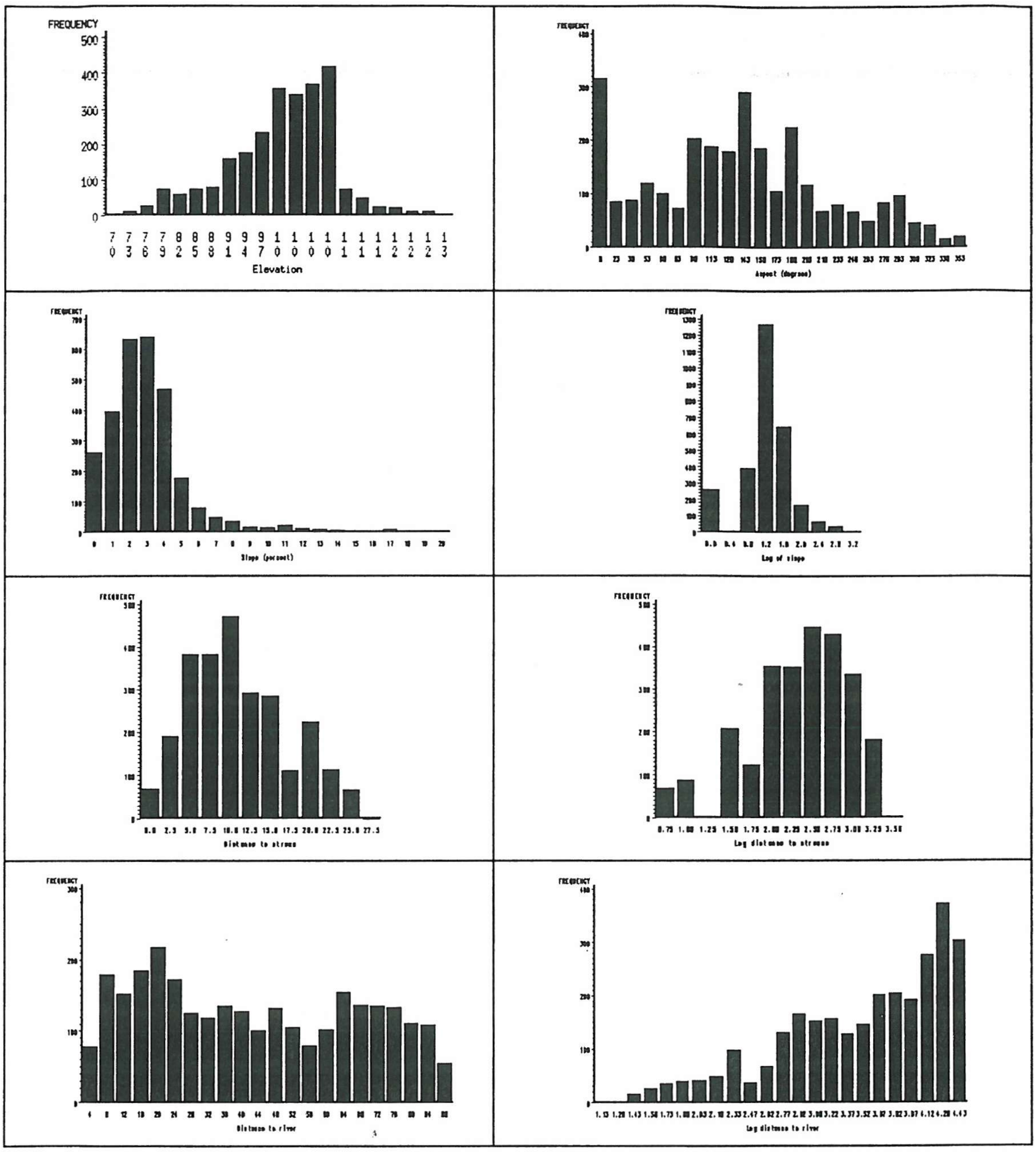


Figure 6.24. Frequency distributions of elevation (top left), aspect (top right), slope (centre left), log (slope) (second row right), distance from stream course (third row left), log (distance to stream courses) (third row right), distance to river (bottom left) and log (distance to river) bottom right.

As discussed in chapter 2, slope overlays were derived from the elevation model and slope, measured in percent, was included as a predictor. The resulting index varies between zero and 20% for the region in question and is presented as a frequency histogram in figure 6.24, second row left. It is clear that the variable is not normally distributed: the very large number of low values causing the frequency histogram to be heavily skewed to the left, reflected in a high value of measured skewness of 2.59. Application of a natural log transformation produces a new variable, for which the frequency histogram is shown in 6.24, second row right. This is a far better approximation of a normal distribution, reflected in the skewness statistic of 0.25.

Geology and Soil

Geology and soil variables were introduced through the creation of dummy variables as for aspect. Geology was divided into *upper chalk*, and *clay-with-flints*, with the remaining variation in geology (*valley gravel*) being accounted for implicitly by the absence of both of these. For soil, four dummy variables were created for *humic rendzina, grey rendzina, brown rendzina* and *typical brownearths* with the remaining category of *calcareous gleys* being implicit within these.

Distance from stream courses

One of the products of a water runoff model is an overlay of water accumulation. Application of an appropriate cut-off value results in a map which shows those parts of the landscape in which streams will naturally form. These areas are also those areas with valley-like characteristics and therefore the overlay is representative of those areas which offer a high degree of shelter. Application of a distance transformation in 100m units to this overlay results in an index which varies between 1 and 27 for the flint density observations. Although this index will be referred to as 'distance from stream courses' to reflect the derivation of the overlay, there is no implication that the areas identified actually represent flowing streams and the index therefore mainly reflects the distance of a location from the nearest wet or dry valley.

The frequency histogram for the index is shown in figure 6.24, third row left. This reveals an approximately normal distribution, although there is a suggestion of skewness similar to that for slope. The skewness value of 0.46 is not high. Attempting to correct the variable by applying a log transform, as for slope, results in a skewness to the right, and in fact increases the skewness statistic to -0.73. The index was therefore used without transformation.

Distance from river

Distance from the river Avon was calculated from the digitised river, and a distance map produced in 100m units. This produced an overlay with values for the observed locations between 2 and 89. The frequency distribution is shown in figure 6.24, bottom left. It is clear from this that, while there is no bias to either left or right (supported by the skewness statistic of 0.24) there is also no central tendency to the distribution, which in fact exhibits a suggestion of bimodality.

Application of a log transformation (6.24, bottom right) increases the central tendency, but at the expense of introducing a highly skewed distribution. This was therefore not adopted and uncorrected distance to the river was used.

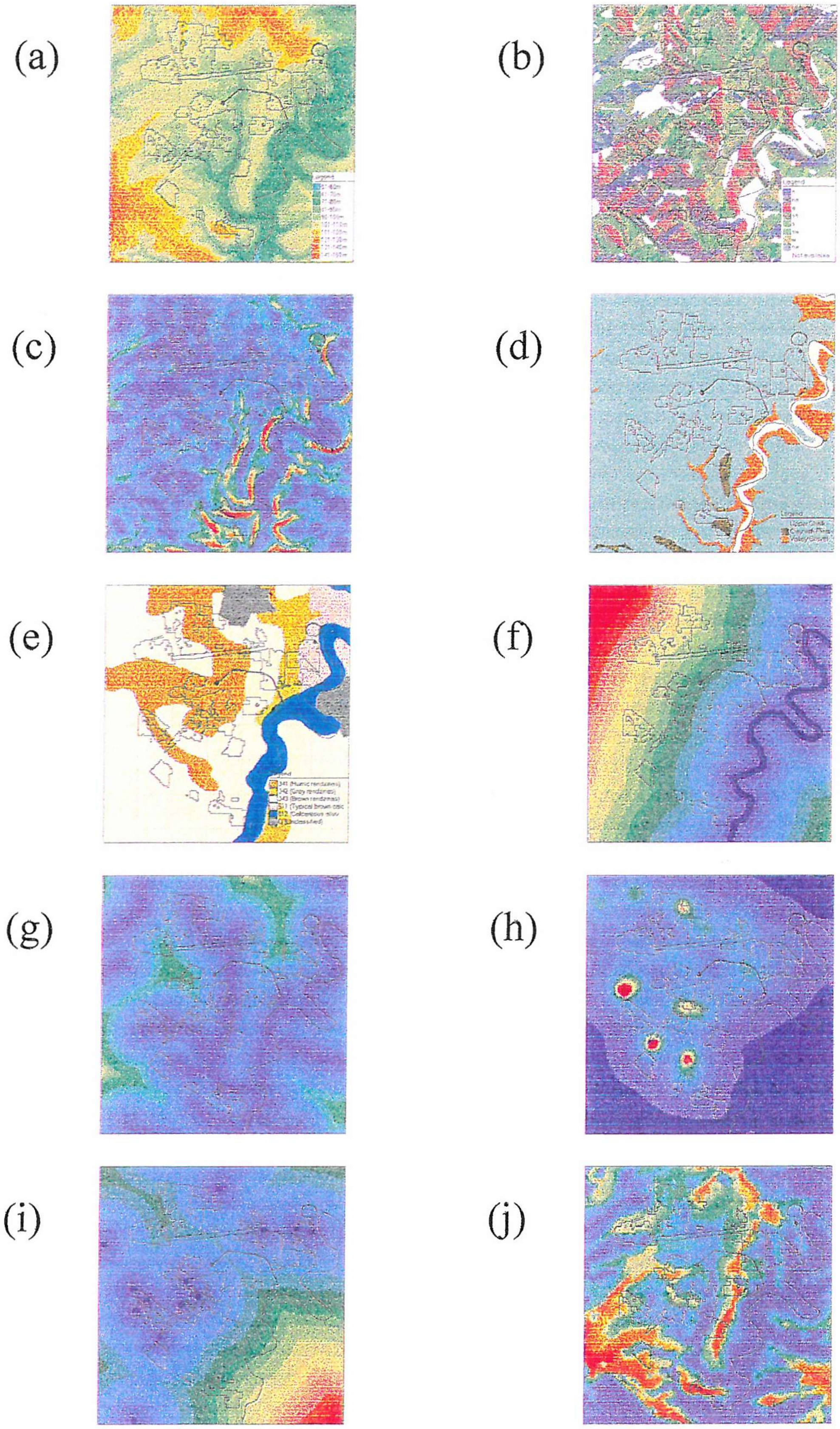


Figure 6.25. Independent variables. (a) elevation, (b) aspect, (c) slope, (d) geology, (e) soil, (f) distance from river, (g) distance from stream course, (h) density of round barrows, (i) distance to long barrows and (j) visibility of long barrows.

Cultural landscape indices

As discussed previously, one of the criticisms raised of predictive models, has been that of environmental determinism. It was suggested in chapter 1 that this may be more a product of expediency than sinister intent and that the inclusion of other types of variable within the analyses may be possible. To this end several variables intended to reflect cultural aspects of the landscape were introduced as possible predictors of lithic density. These were *density of round barrows*, *distance to long barrows* and *visibility of long barrows*. Map output is shown in figure 6.25 (h), (i) and (j).

Distance from long barrows

The importance of the locations selected for monumentalisation as long barrows has been one of the central themes of chapter 5, suggesting that differential use of the landscape for ritual activity and flint working is not unlikely. It seems possible that the proximity or otherwise of ritually important locations such as long barrows may be reflected as a correlation within the lithic density data.

The long barrow monuments of the area were therefore used to generate an index of distance in 100m units. This resulted in an index varying from 0 to 59 for the flint observations, the frequency distribution for which is given in figure 6.26, top left. This shows evidence both of skewness (the skewness statistic is 1.1) and some bimodality. Applying the log transform to the values produces a less skewed distribution (figure 6.26 top right), reflected in the reduced skewness of -0.5. Although not ideal, the transformed variable was therefore included in the regression.

Visibility of long barrows

Visibility of long barrows, as discussed in chapter 5, may also have been an important structuring principle of the prehistoric landscape of this area. Chapter 5 suggested that the number of long barrows visible from a location may have influenced the siting of subsequent monuments. It seems possible, therefore, that a similar influence may have been present on the activities responsible for the lithic densities.

The cumulative visibility surface used in chapter 5 was therefore incorporated into the regression model as a possible predictor of lithic density. This provided an index varying between 0 and 22 for the sampled locations, distributed as shown in figure 6.26, centre left. Although there is again some suspicion of skewness to the distribution, supported by the skewness statistic of 1.0, application of a log transform fails to reduce the skewness by any substantial extent (skewness of logged variable is 0.9). It was therefore felt that the transformation would not provide sufficient benefit to be worth performing and the visibility values were included in their untransformed form.

Density of round barrows

The other major ritual monuments of the area are the round barrows. These proved too numerous to characterise in terms of distance and visibility as for long barrows. Nevertheless, their inclusion was felt to be important because they provide the most numerous evidence of ceremonial activity within the area. Visual inspection of the round barrow distributions, as compared with the lithic density maps, also suggested that a relationship may exist.

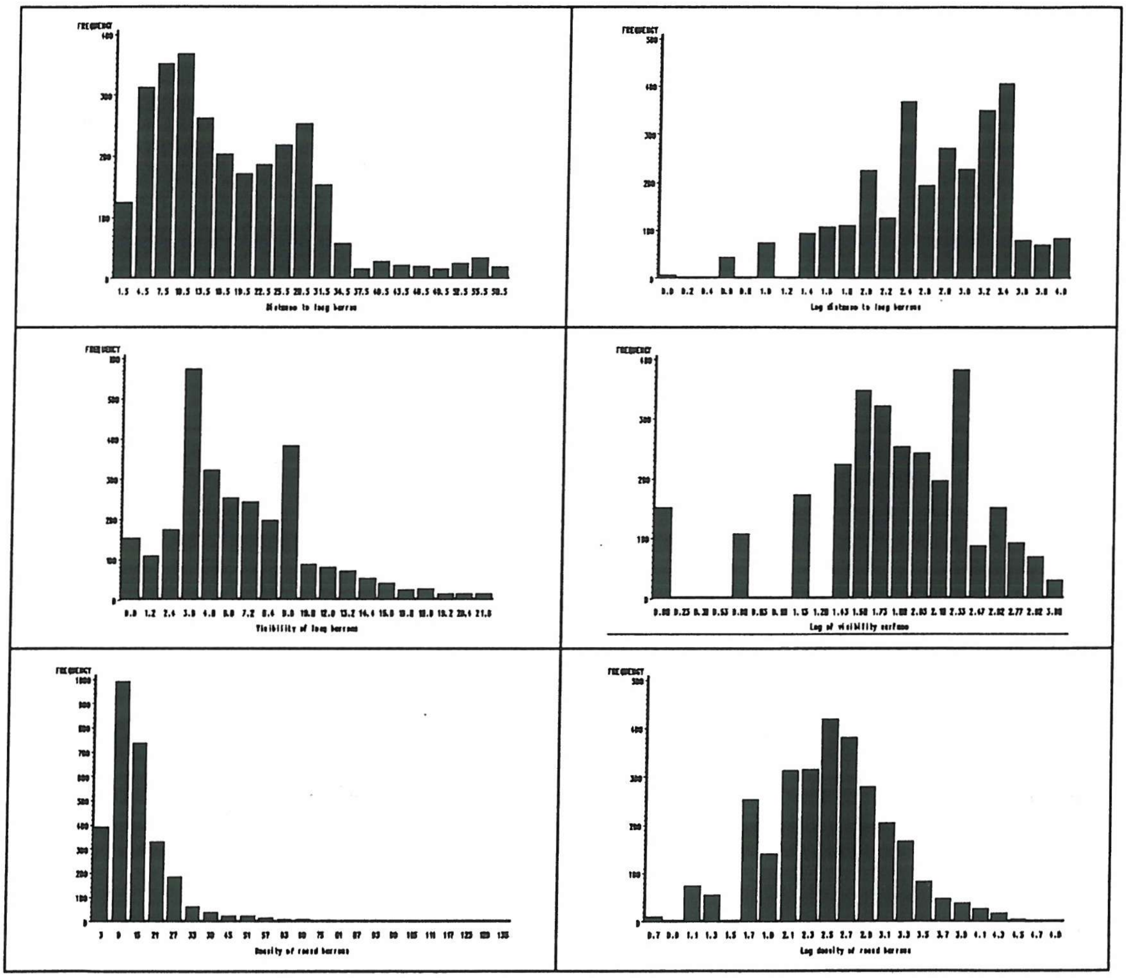


Figure 6.26. Frequency distributions of distance to long barrows (top), visibility of long barrows (centre) and density of round barrows (bottom). Logged variables are on the right.

The index adopted was therefore intended to quantify both the density and proximity of round barrows. This was obtained by regarding the round barrows as point entities with value 1, and then applying an inverse-distance weighed interpolation to the points as discussed above but without restriction on the number of neighbouring points . This results in an index which reflects both the density of barrows at each location within the landscape and, through the distance weighting, the proximity of barrows. The index scores highest for areas which are close to many barrows, moderately high for areas close to a few barrows, and lowest for areas which are far from all barrows. The index varies from 1 to 135 for the sampled areas and a frequency histogram is given in figure 6.26, bottom left. The round barrow index exhibits clear skewing to the right, typical of many spatial variables, and reflected in the skewness statistic of 3.01. Transformation using a natural logarithm provides a variable with a substantially less skewed distribution (6.26 bottom right), skewness statistic -0.04. The transformed index was therefore adopted within the regression.

6.7.4. A linear multiple regression model

Application of the model

Once the dependent and independent indices had been selected for the model, the appropriate overlays were constructed within the GIS. For each index a point-select operation was then performed using the flint density point data from which the density maps were obtained. At the end of this procedure, a point file was obtained for which the attributes were (a) the flint density data discussed above and (b) the independent indices selected for the model. Few GIS currently provide facilities for multiple regression analysis, and it was therefore necessary to export this file for analysis in the SAS statistical package. This was achieved by exporting the point file as an ASCII text file, then reading this into the SAS program.

A stepwise linear multiple regression was then undertaken. This procedure examines each variable as it is included in the model to identify to what extent it contributes to the model. Independent variables were only included in the regression if they proved significant at the 0.15 level. Eleven independent variables met this criterion for inclusion in the model, a summary of the result is given in table 6.2. This reveals that the flint densities show significant correlations with 11 of the variables, which in reality represent 7 of the selected variables once the dummy variables for slope, geology and aspect have been counted out. The partial correlation coefficients (r^2) and F scores for each the variables show that the density of round barrows index has the greatest explanatory power within the model, alone accounting for some 15% of the variation in lithic density while the aspect variables, distance to stream courses, soil indices and visibility of long barrows also carry some explanatory value. Surprisingly perhaps, slope accounts for very little of the variation in lithic density (around 1%) and elevation even less, not passing the 0.15 significance level test for inclusion in the model. Figure 6.27 represents this in graphical format.

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	4.48422092	0.13117657	755.88447063	1168.59	0.0001
LOGSLOPE	-0.07334661	0.03547191	2.76557247	4.28	0.0388
DSTREAMS	-0.03190818	0.00284291	81.48413338	125.97	0.0001
VIEWSUM	0.01457195	0.00420610	7.76370920	12.00	0.0005
LOGDENRB	-0.48342903	0.03451878	126.86697985	196.13	0.0001
GRYREND	0.41090168	0.08041348	16.88931830	26.11	0.0001
BRNREND	0.16700804	0.03942703	11.60594984	17.94	0.0001
TBCEARTH	0.44753602	0.07039656	26.14250292	40.42	0.0001
UCHALK	0.14062157	0.08714878	1.68412841	2.60	0.1067
NE	0.33501178	0.04170020	41.74820486	64.54	0.0001
SW	-0.36472949	0.03930261	55.70483453	86.12	0.0001

Bounds on condition number: 1.852364, 134.823

All variables left in the model are significant at the 0.1500 level.
No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Procedure for Dependent Variable LOGMTOT

Step	Variable Entered	Variable Removed	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F
1	LOGDENRB		1	0.1524	0.1524	348.0545	464.8721	0.0001
2	SW		2	0.0373	0.1897	221.2244	118.7974	0.0001
3	DSTREAMS		3	0.0277	0.2174	127.3268	91.5275	0.0001
4	NE		4	0.0182	0.2356	66.4410	61.4242	0.0001
5	HUMREND		5	0.0067	0.2423	45.3233	22.7708	0.0001
6	TBCEARTH		6	0.0033	0.2456	35.7845	11.4115	0.0007
7	VIEWSUM		7	0.0032	0.2488	26.6531	11.0514	0.0009
8	GRYREND		8	0.0032	0.2521	17.4324	11.1841	0.0008
9	BRNREND		9	0.0019	0.2540	12.9419	6.4831	0.0109
10	LOGSLOPE		10	0.0011	0.2551	11.0785	3.8633	0.0495
11		HUMREND	9	0.0005	0.2546	10.7186	1.6400	0.2004
12	UCHALK		10	0.0008	0.2554	10.1158	2.6036	0.1067

Table 6.2. Summary of stepwise multiple regression model for ln(mean filtered total flint).

Probably the most significant statistic within the result however, and generally one of the most informative statistics in regression analyses, is the model correlation coefficient r^2 of 0.2554, or around 25% for all 7 explanatory variables. This indicates that in all the independent variables can account for only around one quarter of the variation within the lithic density, leaving three-quarters of the variation unexplained by the model.

This is a very low value for r^2 from a regression analysis of this type, but a number of experimental modifications to the analysis through alterations and exclusions independent variables (removal of variables which may not be normally distributed, or present in sufficient numbers) and cases (removal of cases for which no aspect was calculated for example) failed to achieve a value higher than around 27% and as none of these minor experimental alterations improved the methodological rigour of the analysis they are not discussed here in detail.

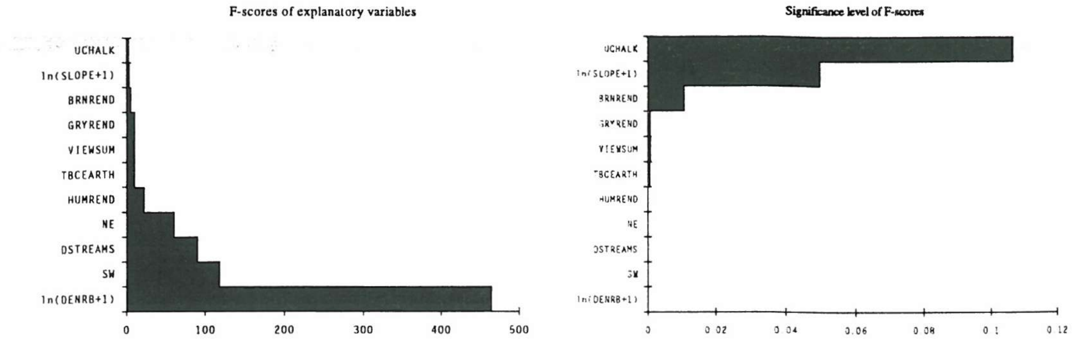


Figure 6.27. Explanatory power of the variables remaining in the regression as F-scores, increasing with explanatory power, and significance of F-scores, reducing with explanatory power.

Accepting the low correlation coefficient for the time being (further attention will be given to this below), the model may now be used to express the predicted lithic density at all locations in the landscape, and then be translated in to a prediction map. Allowing for the transforms used to approximate normally distributed variables, the regression produces the following equation relating the independent variables with lithic density:

```
Flint density =  exp(4.48422092
                 - (0.07334661 x ln(slope+1))      - (0.03190818 x dstreams)
                 + (0.01457195 x bviewsum)         - (0.48342903 x ln(rbardens+1))
                 + (0.41090168 x grend)            + (0.16700804 x brend)
                 + (0.44753602 x tbcearth)         + (0.14062157 x chalk)
                 + (0.33501178 x NE)               - (0.36472949 x SW)
                 ) - 1
```

This equation was then used as the input to a map algebra operation, and this was undertaken within the SPANS GIS to produce the predicted densities for the entire region. The result is shown in figure 6.28, using the same colour coding and scale as figure 6.4 so that the dependent variable and the prediction may be directly compared.

Discussion of the model

The overall character of the prediction is apparent from figure 6.28 which shows that the model generally predicts rather low values for flint throughout the sampled area, with the prediction never reaching more than around 80-90 within the convex hull of the data points. The model seems to provide a fairly close approximation of the observed densities for the region around Normanton Down (56, 88, 84, 55, 79 and 61), and shows slightly higher predictions (although never approaching the true densities) for the surrounding areas. The areas of high density north of the Cursus (52), around Coneybury Henge (51) and north of Stonehenge Down (54) also show up as marginally higher values although, again, these do not approach the real densities for these areas.

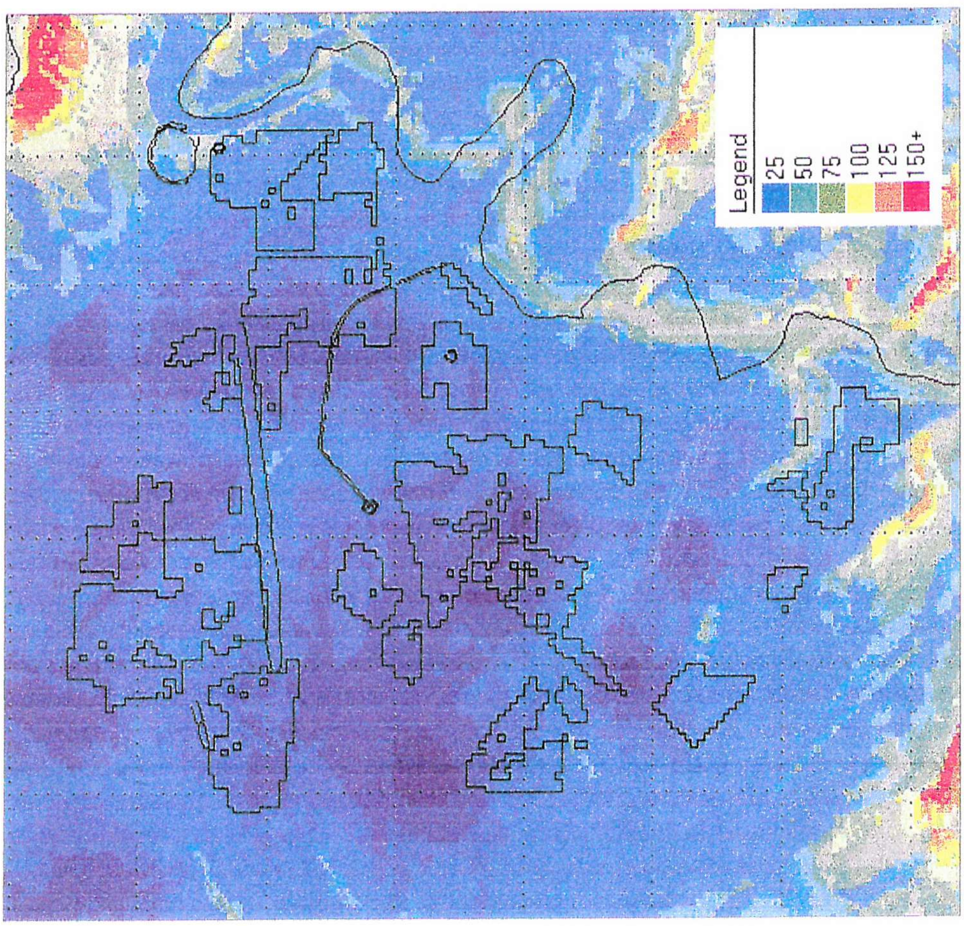


Figure 6.28Linear multiple regression predictive model of total lithic density, the colour coding is identical to figure 6.4

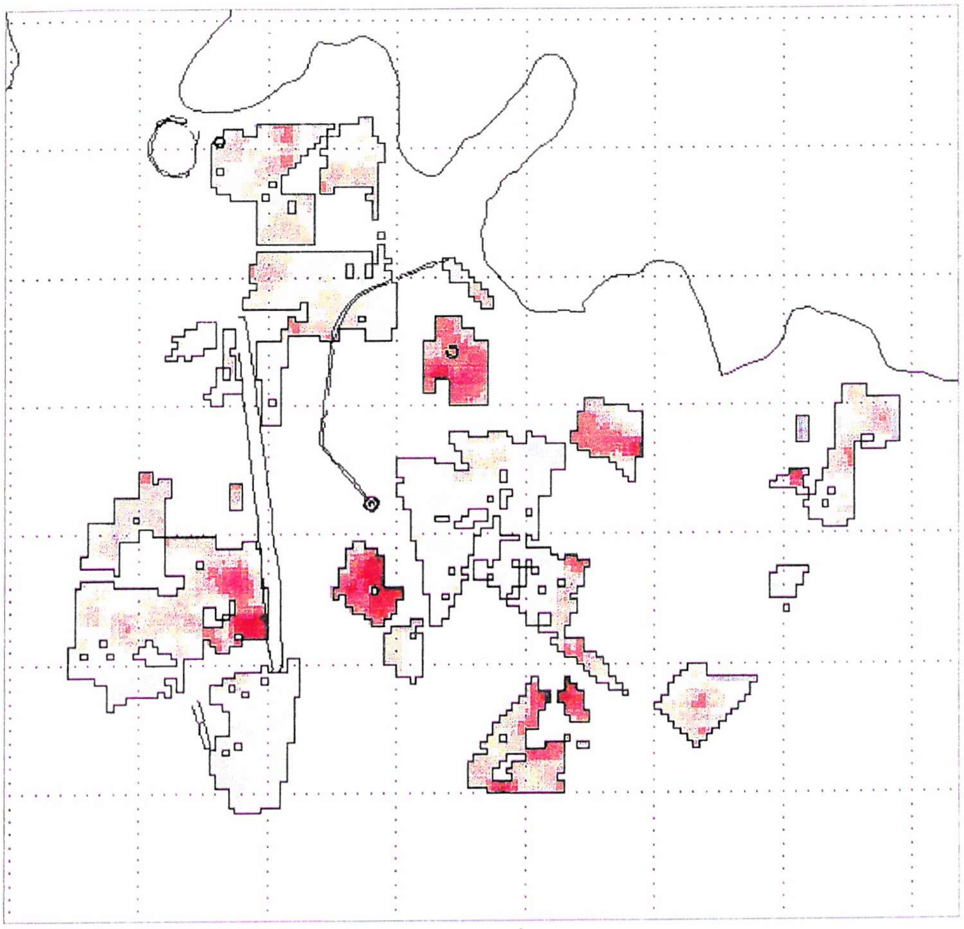


Figure 6.29. Residuals obtained by subtracting figure 6.28 from figure 6.4. Red indicates under-prediction, blue over-prediction

One of the most obvious features of the model are the high values which are predicted for many areas away from the sampled region, for example in the northeast and southwest corners of the image. These seem highly unlikely, and are revealing about the portability of the model which may in fact be no greater than the spatial interpolations discussed above. It may be that the variable most influential in the regression (density of round barrows) is the main factor. It is, for example, a measure which may be particularly prone to edge effects: the model predicts high flint densities at low round barrow densities, and the edge areas of the map may underestimate that index. It is also possible, however, that the model is actually rather 'over-trained' in regard to this variable. The variation of lithics with round barrow density may, at the scale of the sampled area, be as described in the regression equation, but at a larger scale be entirely the reverse. In other words, within the general cluster of monuments on Salisbury Plain the density of lithics are indeed higher in the gaps between the monuments, but nevertheless the flint densities at the edge of the study area might also be expected to decline rather than increase with density of monuments. The high predictions away from the core of the study area seem best explained in this way, therefore, and, put simply, it is that the model has no experience of lithic densities at the edge and extrapolates unreasonably high densities as a result.

The main conclusion which may therefore be deduced from the experiment is that while some of the 'shape' of the distribution is modelled within the core of the study area, this is clearly a poor model in the sense that it fails to adequately account for a large proportion of the variation within the dependent variable and in that it cannot be used outside the scope of the area for which it was designed. There are two possible explanations for this: either the methodology is imperfect, or the model is an accurate reflection that the variables selected as predictors are simply not adequate.

In considering the first of these, it should be remembered that some of the independent variables are questionably normal in distribution, as discussed above, and that the distance from rivers index exhibits more than a suggestion of bimodality. However, of the three variables which might be questioned in this regard (distance to long barrows, distance to river and visibility of long barrows), only the visibility index actually meets the criteria for inclusion within the model. Exclusion of all three variables from the analysis was also attempted but made little significant difference to the outcome. It seems safe to assume that the characteristics of the independent variables are not to blame for their low explanatory value and that the poor predictive value of the result is a genuine reflection of the relationships.

It may also be suggested that the choice of dependent variable is a possible cause of the poor fit of the model, after all the selected dependent variable is a filtered variable and to that extent is somewhat artificial. However, use of the same independent variables was also attempted and produced a model with a rather lower coefficient, and the same relationship was observed between flake density and filtered flake density. This suggested that the case set out above for use of the filtered variable was probably well founded, and that prediction of a generalised density index was indeed the best option.

Having accepted that this is therefore a poor predictive model in that the regression is a poor fit, a useful approach must be to explore the prediction itself, and to examine where in particular the model fails to fit the data. To this end the residuals of the regression were obtained within the GIS by subtracting the result of the application of the regression equation (figure 6.28) from the original observations in the form of the mean filtered density map (figure 6.4). The result is shown in figure 6.29, coded to show areas for which the model under-predicts in red, and areas for which the model over-predicts in blue.

The residual map supports the interpretation presented above, that the model fits adequately the areas which exhibit a low-level and low variation of lithic density. Thus the Normanton Down areas are a good fit, as is the peripheral area southwest of Fargo Wood (62). Practically all the very high areas of flint appear as high residuals. It is possible, therefore, that the model explains the regular and therefore predictable variation within the lithic values which might be termed 'background' variation. The areas which show high residuals must then be interpreted as 'unpredictable' areas which deviate dramatically from the trend. The high residual areas might therefore be interpreted as areas where human intervention was at it greatest, causing either more-than-expected areas of flint to be deposited in araes of extensive use or less-than-expected numbers of flints in areas which were avoided.

Finally it is worth making the point that the lack of success of this approach should not be taken to devalue the experiment itself. The failure of the model to adequately predict absolute values for flint densities with the available independent variables is, of itself, an interesting finding worthy of some explanation. The reasons why this particular approach to prediction might fail in this context will be discussed further below.

6.7.5. Logistic multiple regression models

Given that the accurate prediction of absolute values for flint density did not prove successful, it may still be possible to use logistic procedures to generate predictions of other variables which describe the flint. However, some of the flint variables do not show sufficient range of variation to be used as dependent variables in linear multiple regression analysis. Logistic regression, unlike linear multiple regression, can be used to predict presence/absence of particular classes rather than interval level values. Using logistic approaches, it is therefore possible to turn to the flint classes with low variability, and overall low values, such as core and retouched pieces density, and to define some characteristic of these variables which may be worth predicting. It has already been seen that the relationship between these two classes is of interest, but that they do not provide sufficient variation or range for a profitable linear multiple regression model to be constructed.

Selection of dependent variables

An appropriate generalisation is suggested by the core/retouch ratio images above: these seemed to indicate that there was some differential use of the landscape indicated by different flint classes. High values of cores might be taken to suggest areas of manufacture, while high values of retouched pieces might indicate areas of discard. The core/retouch images indicate that these were not always the same areas, therefore it may be possible to derive methods for predicting which activities might be expected in which parts of the landscape.

To this end two normalised core/retouch ratio thresholds were set which allowed the production of binary maps indicating those areas which seem to have been used for manufacturing activity and those which seem to have been used for discard. The aim of setting the thresholds was to obtain two variables which provided a clear distinction between areas. After a little experimentation, values of 0.9 or less were classed as discard areas, while areas of 1.1 or above were classed as manufacture areas. The resulting maps can be seen in figure 6.30 and 6.31, and seem to provide the best binary summary of the variation suggested by the ratio images.

Application of the models

Once these variables had been generated, they were transferred to the SAS statistical package as for the multiple regression analysis. The same independent variables were adopted for the logistic regressions as for the linear multiple regression experiment, the only modification to the procedure for the linear multivariate model was that all variables were used un-transformed because the logistic procedure does not require normally distributed variables (Rose & Altschul 1988).

The SAS LOGISTIC procedure was then used to generate the model, a stepwise procedure was again used. The procedure uses the iterative maximum-likelihood method for fitting linear logistic regression models to binary or ordinal level response data, obtaining optimum intercept and parameter estimates. Maximum-likelihood estimation is the preferred method for obtaining parameter estimates in computer programs because of statistical problems inherent in the alternative least-squares method (Warren 1990b). Responses for some variables proved to be singular - all of the positive responses for the cases occurred within one category or no responses at all within a particular category - and therefore these were removed from the analysis: failure to do this leads to infinite estimates for some parameters.

Step	Variable Entered	Variable Removed	Number In	Score Chi-Square	Wald Chi-Square	Pr > Chi-Square
1	DLBARS		1	78.6853	.	0.0001
2	DRIVER		2	88.4108	.	0.0001
3	ELEV		3	86.4427	.	0.0001
4	CHALK		4	28.1790	.	0.0001
5	SLOPEF		5	42.2459	.	0.0001
6	SE		6	33.7102	.	0.0001
7	NE		7	98.5428	.	0.0001
8	BREND		8	14.5768	.	0.0001
9	DSTREAMS		9	20.9877	.	0.0001
10	VIEWSUM		10	9.1127	.	0.0025
11	DENSRBAR		11	4.6568	.	0.0309

Analysis of Maximum Likelihood Estimates

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	3.6050	2.1878	2.7152	0.0994	.
ELEV	0.1080	0.0320	11.3757	0.0007	0.574173
SLOPEF	0.5415	0.0710	58.1830	0.0001	0.749372
NE	-2.8268	0.3715	57.8992	0.0001	-0.638711
SE	-2.5466	0.3264	60.8808	0.0001	-0.692974
CHALK	-6.9748	1.2567	30.8042	0.0001	-0.835570
BREND	-1.3809	0.2650	27.1575	0.0001	-0.380738
DSTREAMS	0.1549	0.0289	28.6652	0.0001	0.526867
DRIVER	-0.0547	0.0110	24.7250	0.0001	-0.748168
DLBARS	-0.1525	0.0189	65.4072	0.0001	-1.015312
VIEWSUM	0.2034	0.0577	12.4014	0.0004	0.476920
DENSRBAR	0.0666	0.0296	5.0416	0.0247	0.425465

Association of Predicted Probabilities and Observed Responses

Concordant	= 93.2%	Somers' D	= 0.874
Discordant	= 5.8%	Gamma	= 0.883
Tied	= 1.0%	Tau-a	= 0.085
(390630 pairs)		c	= 0.937

Table 6.3. Summary of logistic multiple regression model for 'manufacture areas'.

Summary of Stepwise Procedure

Step	Variable Entered	Variable Removed	Number In	Score Chi-Square	Wald Chi-Square	Pr > Chi-Square
1	BREND		1	94.3617	.	0.0001
2	DLBARS		2	38.0198	.	0.0001
3	VIEWSUM		3	29.1272	.	0.0001
4	NE		4	23.2454	.	0.0001
5	DRIVER		5	19.0265	.	0.0001
6	DENSRBAR		6	20.4722	.	0.0001
7	ELEV		7	10.0973	.	0.0015
8		VIEWSUM	6	.	0.0711	0.7898
9	TBCEARTH		7	9.6831	.	0.0019

Analysis of Maximum Likelihood Estimates

Variable	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate
INTERCPT	6.1250	1.3314	21.1635	0.0001	.
ELEV	-0.0521	0.0123	17.9639	0.0001	-0.276782
NE	-1.0030	0.2136	22.0439	0.0001	-0.226632
BREND	-2.7628	0.3861	51.1922	0.0001	-0.761737
TBCEARTH	2.7358	1.0944	6.2491	0.0124	0.420462
DRIVER	0.0383	0.00625	37.5140	0.0001	0.524288
DLBARS	0.1045	0.0108	93.1905	0.0001	0.695530
DENSRBAR	0.1170	0.0187	39.2001	0.0001	0.747777

Association of Predicted Probabilities and Observed Responses

Concordant	= 87.5%	Somers' D	= 0.757
Discordant	= 11.8%	Gamma	= 0.762
Tied	= 0.7%	Tau-a	= 0.063
(336660 pairs)		c	= 0.878

Table 6.4. Summary of logistic multiple regression model for 'discard areas'.

The output from the LOGISTIC procedure provides a summary of the results of the regression and gives the parameter estimates for those independent variables remaining in the model. Output for the 'manufacture areas' is given in table 6.3, and for 'discard areas' in table 6.4.

These intercept and parameter estimates were then used to generate an estimate of the *probability* of the event represented by the dependent variable for all locations within the study area. This is achieved by the use of the cumulative logistic distribution function (Kvamme 1988 p371) given in equation 6.1.

$$p = \frac{e^{L}}{1+e^{L}} = \frac{1}{1+e^{(1-L)}} \qquad 6.1$$

Both the models were therefore returned to the GIS, and the parameter estimates from tables 6.3 and 6.4 were used to solve the logistic equation from the landform and cultural overlays for all locations. This was done with the spans modelling language, using modelling equations based on table 6.5 The resulting probability maps are shown in figures 6.32 and 6.33, scaled from low probabilities in blue to high probabilities in red.

```
Prob. of manufacturing area = 1/(1+ exp(1- (3.6050
                                + ELEV ´ 0.1080
                                + SLOPE ´ 0.5415
                                - NE ´ 2.8268
                                - SE ´ 2.5466
                                - CHALK ´ 6.9748
                                - BREND ´ 1.3809
                                + DSTREAMS ´ 0.1549
                                - DRIVER ´ 0.0547
                                - DLBARS ´ 0.1525
                                + VIEWSUM ´ 0.2034
                                + DENSRBAR ´ 0.0666 )))

Prob. of discard area = 1/(1+ exp(1- (6.1250
                                - ELEV ´ 0.0521
                                - NE ´ 1.0030
                                - BREND ´ 2.7628
                                + TBCEARTH ´ 2.7358
                                + DRIVER ´ 0.0383
                                + DLBARS ´ 0.1045
                                + DENSRBAR ´ 0.1170)))
```

Table 6.5. Modeling equations to translate parameter estimates in to probability estimates for manufacturing areas (top) and discard areas (bottom).

Discussion of the models

The manufacturing model is complex result, incorporating all the independent variables to generate the response. Increasing values of *slope*, *elevation*, *distance to streams*, *visibility of long barrows*, and *density of round barrows* increase the probability while increasing *distance to river* and *distance to long barrows* both decrease the probability. The presence of *northeast* and *southeast aspects*, *brown rendzina*, and *chalk* rather than *clay with flints* all seem to reduce the probability of a site.

The manufacturing areas model generally shows high values where they would be expected close to the region of the sampled areas: north of the Cursus (52) and around the southern rim of Normanton Down (south of area 67). Low values on Normanton Down (61, 79, 55), around Coneybury (51), south of Durrington Walls (60, 71, 69) and at the eastern end of the Cursus (76, 66 and 85) are also consistent with the data.

However the model predicts manufacturing areas in an unlikely proportion of the area south-east of the river. Examination of the equation, and of the independent variables (see figure 6.25) suggests that the high values in the southeast are primarily influenced by the *distance to long barrows* and the *density of round barrows* indices. Both show extreme values in this area, and both have the type of dual relationship discussed above in relation to the linear multiple regression model. The area of high value in the northwest of the study area may be related to *distance from the river Avon*, although this seems to be also be partly an effect of the *distance to long barrows* index. The equation indicates that both *slope* and *aspect* are significant factors. Higher values of slope are modelled as increasing the probability although the rather localised nature of the slope variable means that this is only apparent in small regions on the edge of the river floodplain. The effect of aspect is quite the opposite, providing continuous variability throughout the study region. Westerly aspects increase the probability of a response. *Elevation, soil, geology* and *visibility of long barrows* are indicated as affecting the probability but their influence seems to be minor.

For the same reasons as for the linear regression therefore, it seems likely that predictions away from the surveyed areas are unreliable and the model must be regarded as internal to the cluster of monuments rather than portable (although see section 6.6).

In contrast to the manufacturing model, which seems to have been influenced by all ten independent variables, the discard model shows responses to only six. The presence of *brown rendzina* soils clearly increases the probability of a discard site when compared with *brownearth*, although this could obviously be a reflection of taphonomic rather than archaeological processes, or of collection bias due to differences in colour or tractability.

Within this soil variation, the other main influences seem to have been *aspect*, with probability reduced for northeastern aspects. The equation indicates less obvious influences from two other landform indices: *elevation* (lower elevations increase site probability) and *distance from rivers*, (probability decreases with distance from the Avon). The *slope* and *distance from streams* indices are not sufficiently correlated to be present in the model.

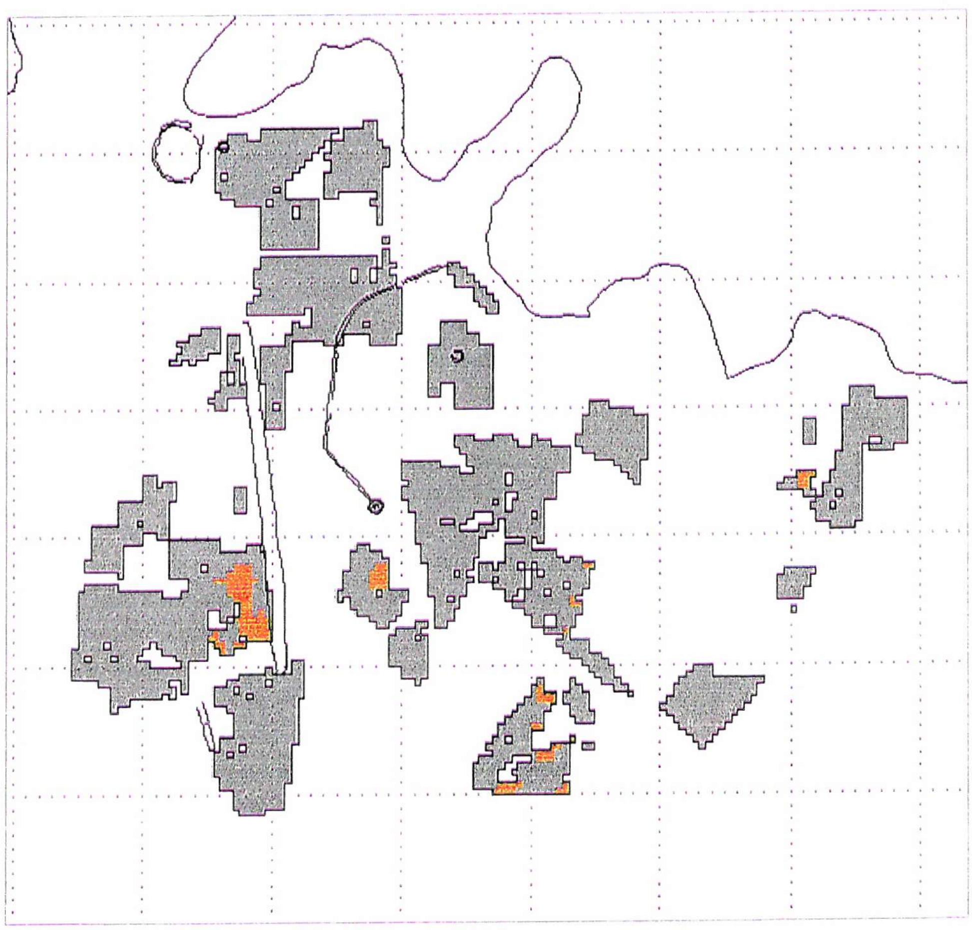


Figure 6.30. 'Manufacturing areas' obtained by reclassification from mean filtered, normalised core : retouch ratio image, threshold 1.1

Figure 6.31. 'Discard areas' obtained by reclassification from mean filtered normalised core : retouch ratios, threshold 0.9

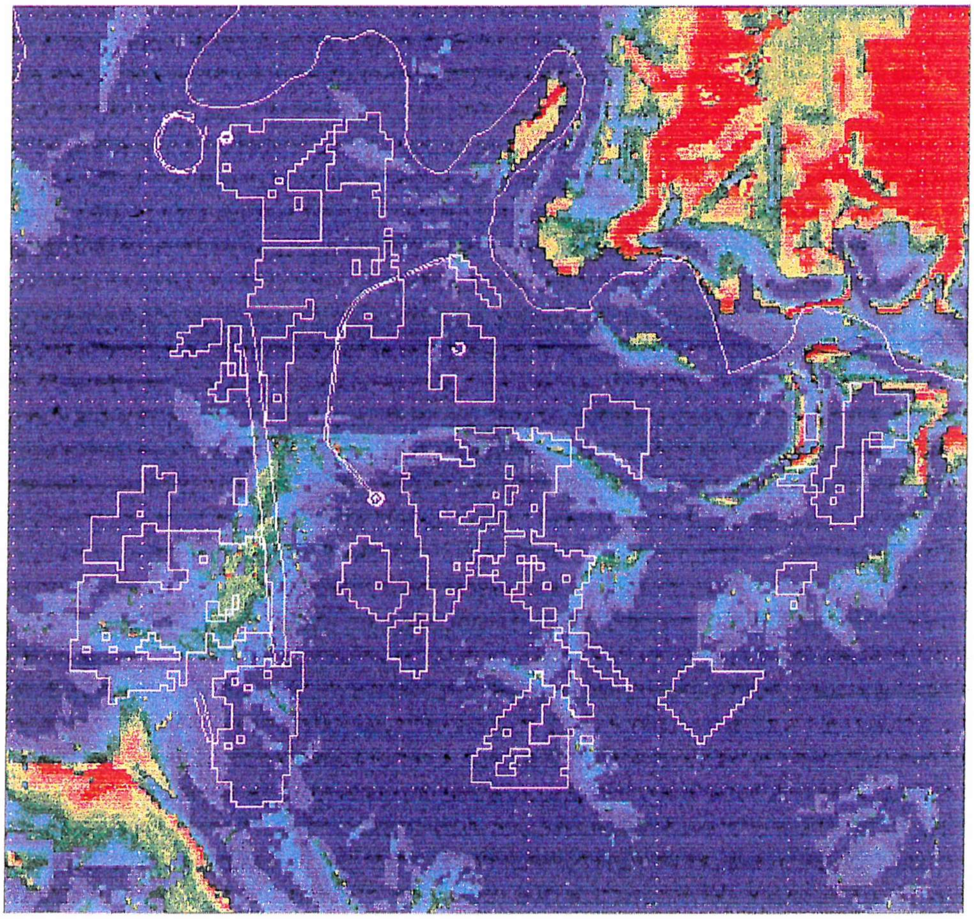


Figure 6.32. Probability of a location being a 'Manufacturing area'

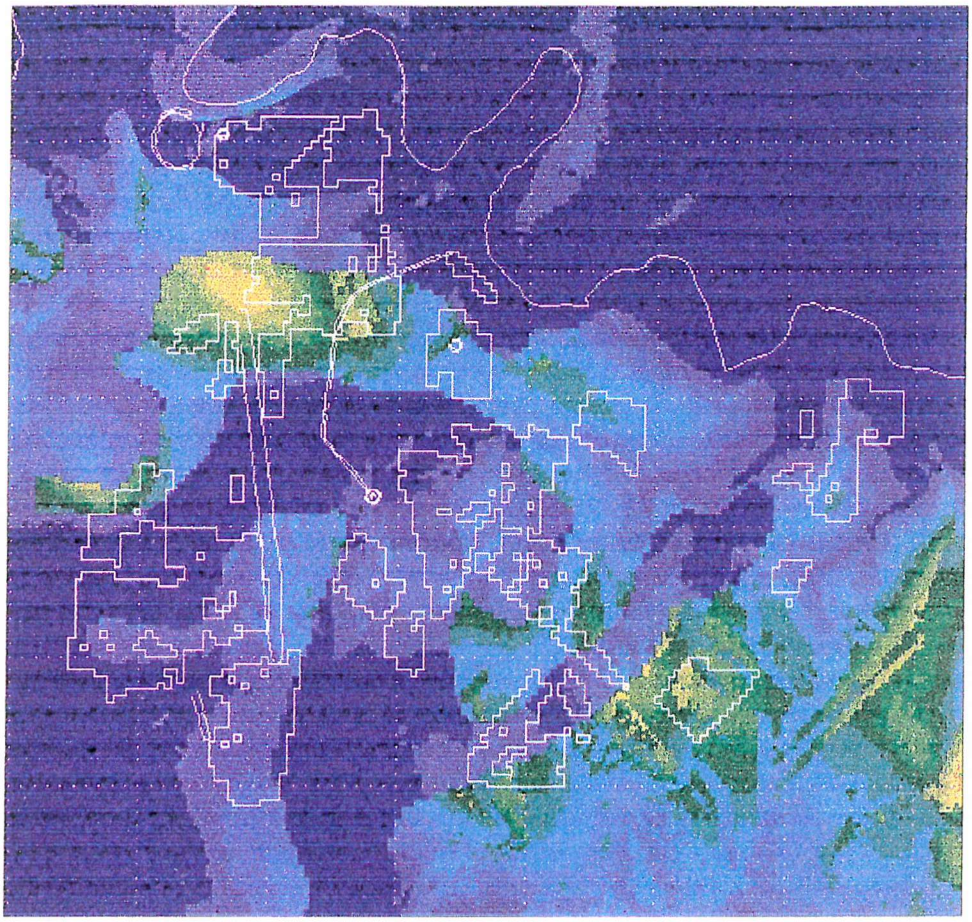


Figure 6.33. Probability of a location being a 'Discard area'

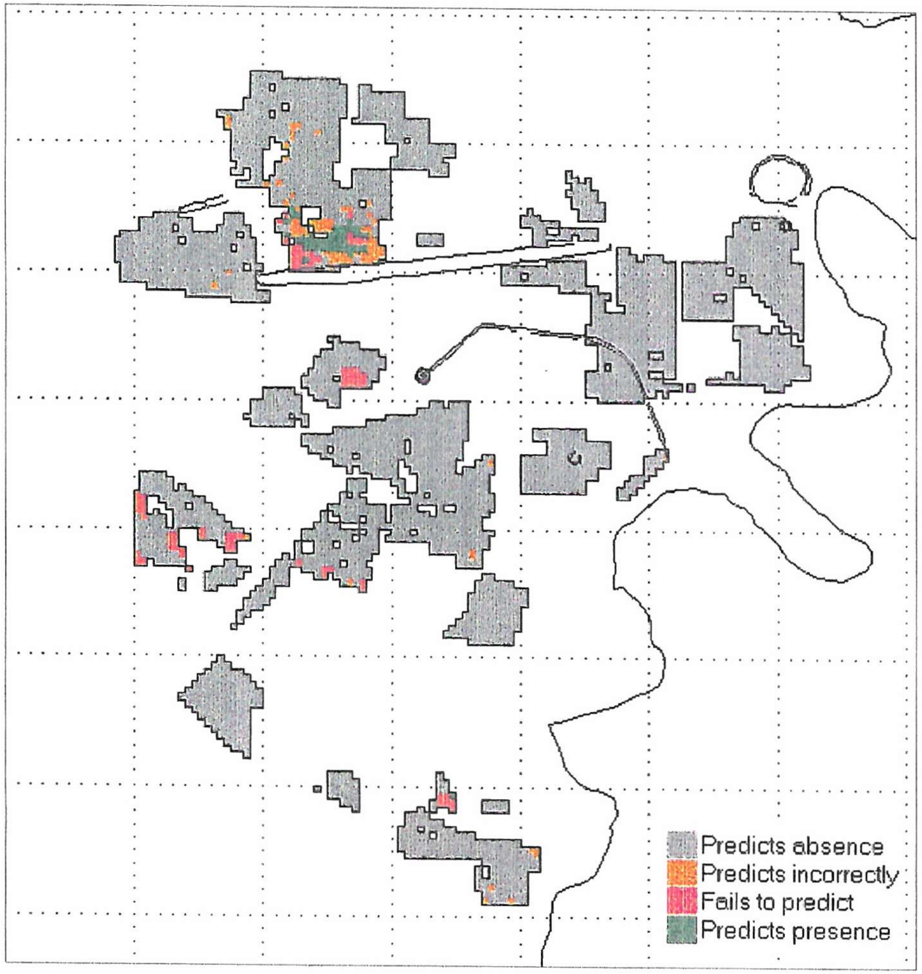


Figure 6.34. Comparison of prediction of 'manufacturing areas' with observed model responses

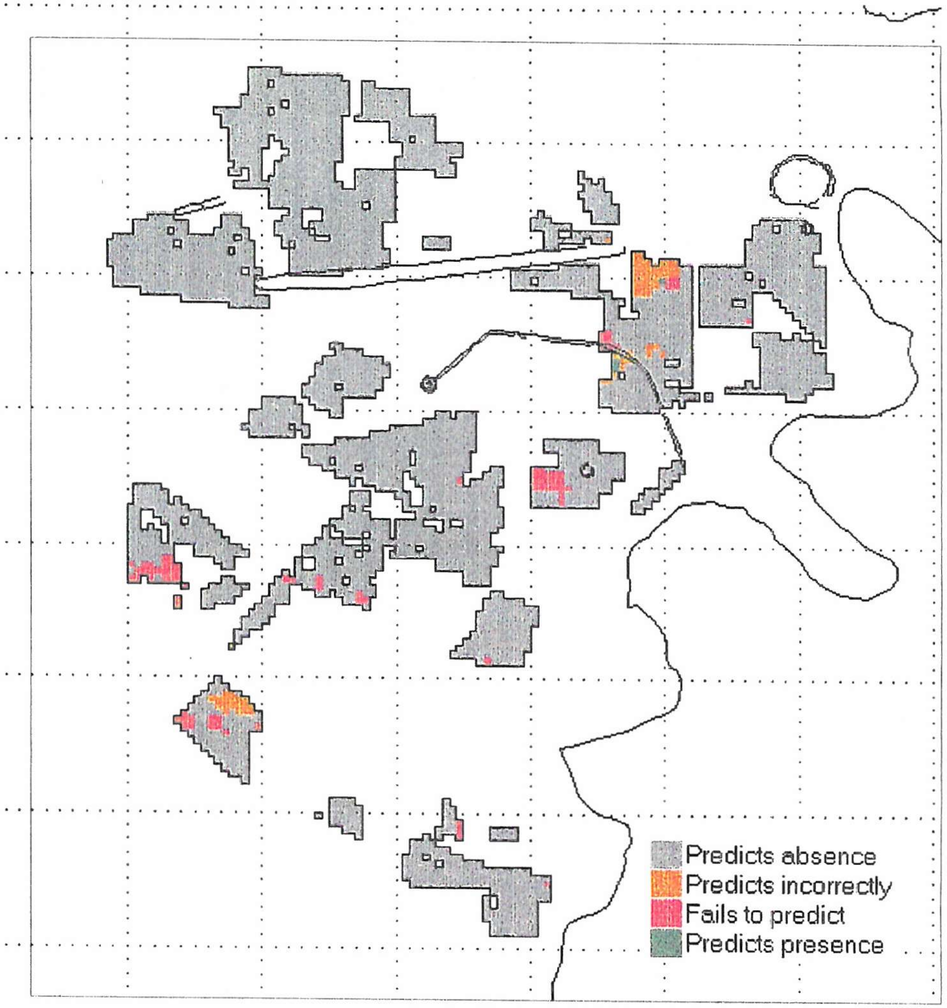


Figure 6.35. Comparison of prediction of `discard areas' with observed model responses

Both *density of round barrows* and *distance to long barrows* decrease the probability of a site, which is curious as the two are inversely related. This is in contrast with the manufacturing area model, although the effects of these two variables may be marginal compared to *soil*. The marginal influence within the model of the cultural variables is enforced by the lack of *long barrow visibility* which shows insufficient correlation to appear in the model.

In spatial distribution, the 'discard' model also shows some encouraging features: high probabilities of discard sites around the elbow of the avenue (87), south of Normanton Down (67) and south of Winterbourne Stoke Crossroads (50) each seem to fit the data. Low probabilities north of the Cursus (52) contrast with the high values of the manufacturing model as should be expected as this forms the major axis of variation between the two areas selected as dependent variables.

Optimisation and assessment of performance

Assessment of the degree of confidence which should be placed in the predictions is difficult. In an ideal situation, further samples would be taken throughout the study area and the results compared with the predicted outcome for those locations. This is rarely possible, however, and it is perhaps slightly ironic that in situations where this were possible, there would then be rather less point in constructing a model.

One source of data concerning model performance is the ratio of observed responses to predictions within the data itself, and this is provided with the output from the procedure. In the case of these models, this indicates that the manufacturing area model makes 93% correct predictions against nearly 6% incorrect, while the discard area model is a poorer fit with 87.5% correct against nearly 12% incorrect predictions. However this is widely recognised as an extremely optimistic assessment of the performance of the probability model and Warren (1990b) recommends withholding a random control sample of observations from the prediction and then comparing the predictions with these controls. Carmichael (1990), used the control procedure advocated by Warren and found that a 72% correct prediction rate amongst the sites used for prediction produced only 55% correct prediction amongst the controls.

This control procedure was not followed here for two main reasons. Firstly the procedure removes some of the cases from the model and therefore inevitably reduces the performance of the procedure. The more cases which are removed, the better the assessment of the model performance but the more effect it will have on the predictive potential of the model. If the available number of cases is high enough then it may be worthwhile sacrificing some of the predictive power of the model for improvement in ability to assess the model performance. In this case, however, the samples are grouped tightly together within walked areas so that any randomly selected subset of points would still have fallen within the same surveyed areas as the cases included in the study.

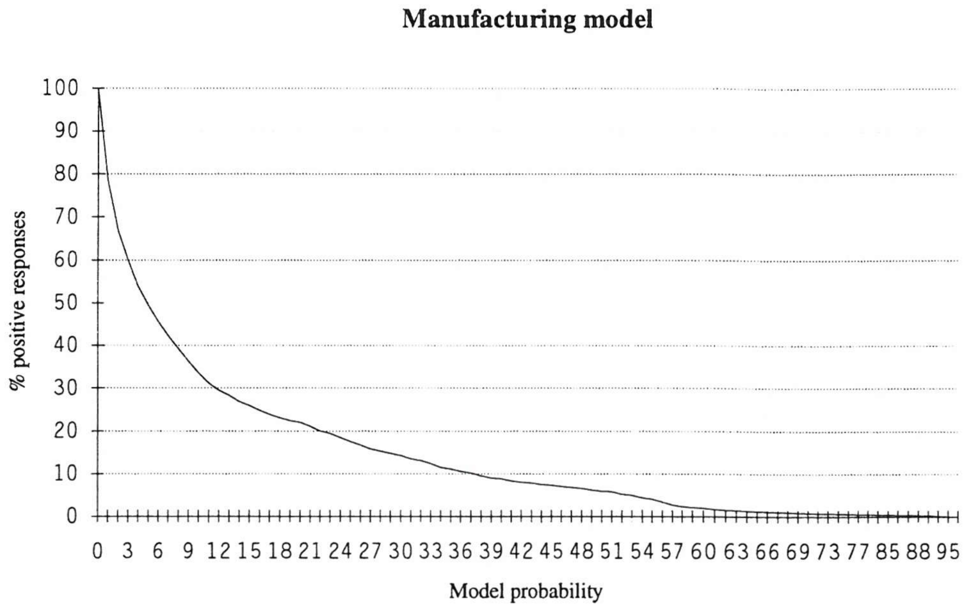


Figure 6.36. The proportion of the surveyed area which is predicted as a 'manufacturing area' for each probability result from the manufacturing model.

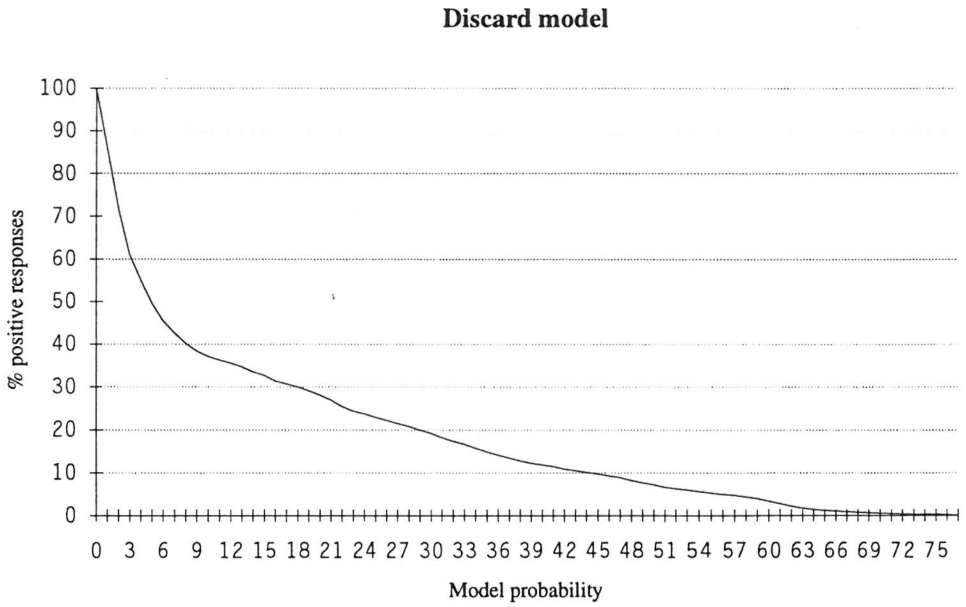


Figure 6.37. The proportion of the surveyed area which is predicted as a 'discard area' for each probability result from the discard model.

Intuition, and experience with the linear regression model suggests that these are likely to be the best performing areas of the model. Consequently little confidence could be held in any assessment of the model based on such a sample of sites, and it was felt that insufficient benefit would probably be obtained to offset the removal of the cases.

For example, Carmichael's (1990) result of 55% may be a better estimate of the performance of the model than 72%, but cannot be taken to indicate either that a model which included the controls would produce a 72% result, or that 55% of the sites in unsurveyed areas would be correctly predicted: the only way to establish this would be to survey those areas.

One method to assess the performance of the model is to force it to predict the same percentage of the surveyed area as a positive response as occurs within the sample data, and then comparing this prediction with the actual result - a procedure analogous in many ways to the examination of residuals of linear regressions. To 'force' the prediction, a threshold was selected for each probability map which generated the same percentage of positive responses within the worked areas as the original dependent variables. An appropriate threshold was obtained from the graphs of probability against proportion of data predicted as a positive response (figures 6.36 and 6.37).

The result of this procedure can then be compared with the original dependent variable, with four possible outcomes. The model may (1) correctly predict no site, (2) correctly predict a site, (3) incorrectly predict a site or (4) fail to predict a site. These four possible outcomes are mapped in figures 6.34 and 6.35.

For the manufacturing model, the prediction is excellent for negative predictions. Of those land units which do not meet the criteria for being a manufacturing zone, 97.4% are correctly predicted by the model (grey) and only 2.6% (orange) are predicted as manufacture zones. On the other hand the model is very poor at predicting the presence of manufacturing activity: of the land units which qualify as manufacturing areas, the model predicts only 33%, while predicting the remaining 66% as negatives.

The prediction result for the discard model is worse. Although it too predicts a high proportion of the location which do not qualify as discard sites (97.6% correct, 2.4% incorrect), the co-incidence of predicted discard areas with actual discard areas is minute. Of the discard areas, the model predicts only 3.7% correctly, classifying the remaining 96.3% as nonsites.

Part of the reason for this poor performance is the influence of the 'nonsite' locations over the 'site' locations. Because there are a far larger number of negative responses than positive ones, the model is heavily biased towards prediction of these. Two possible mitigation strategies are possible (Warren 1990b); the first is simply to select a random sample of negative responses so that there are the same number of positive and negative responses.

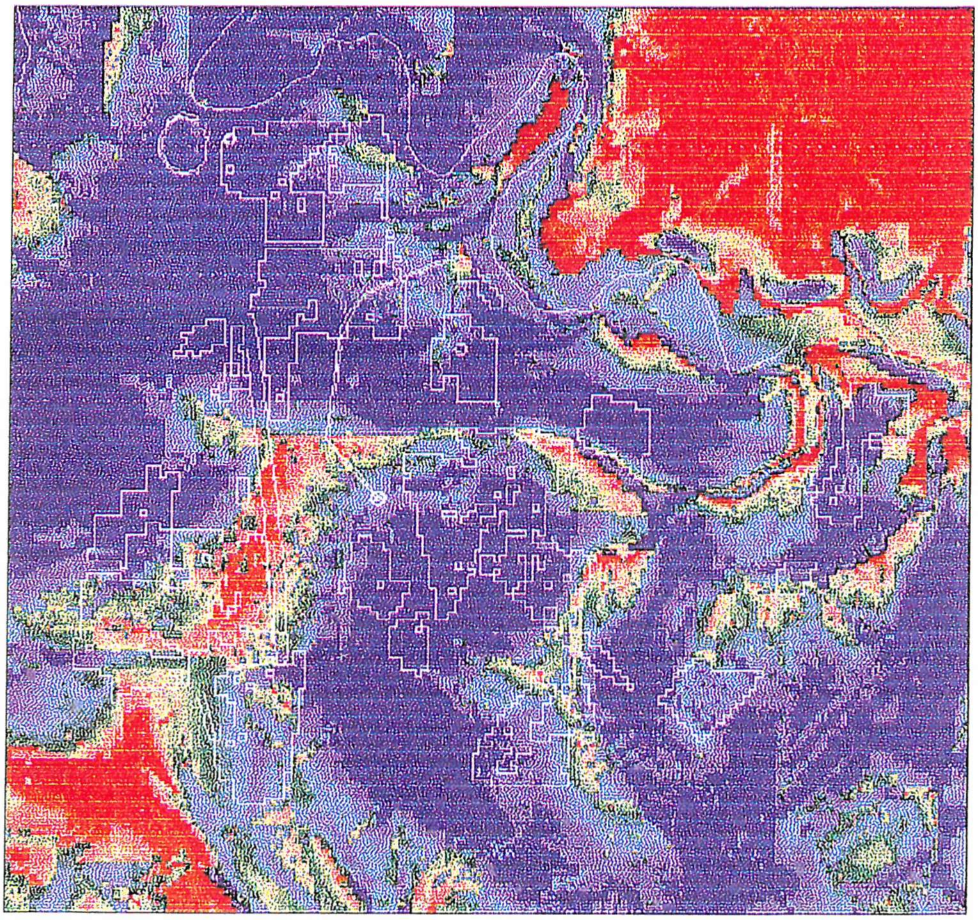


Figure 6.38. Probability of a location being a 'manufacturing area' adjusted for sample size bias

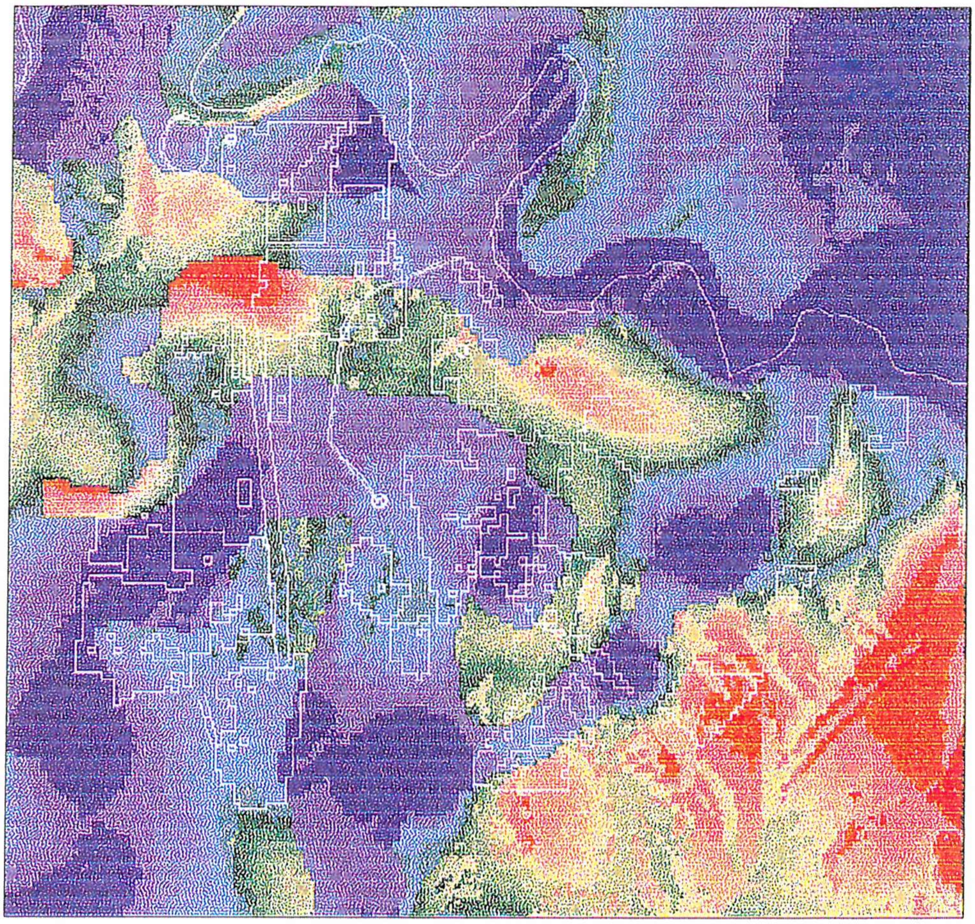


Figure 6.39. Probability of a location being a 'discard area' adjusted for sample size bias

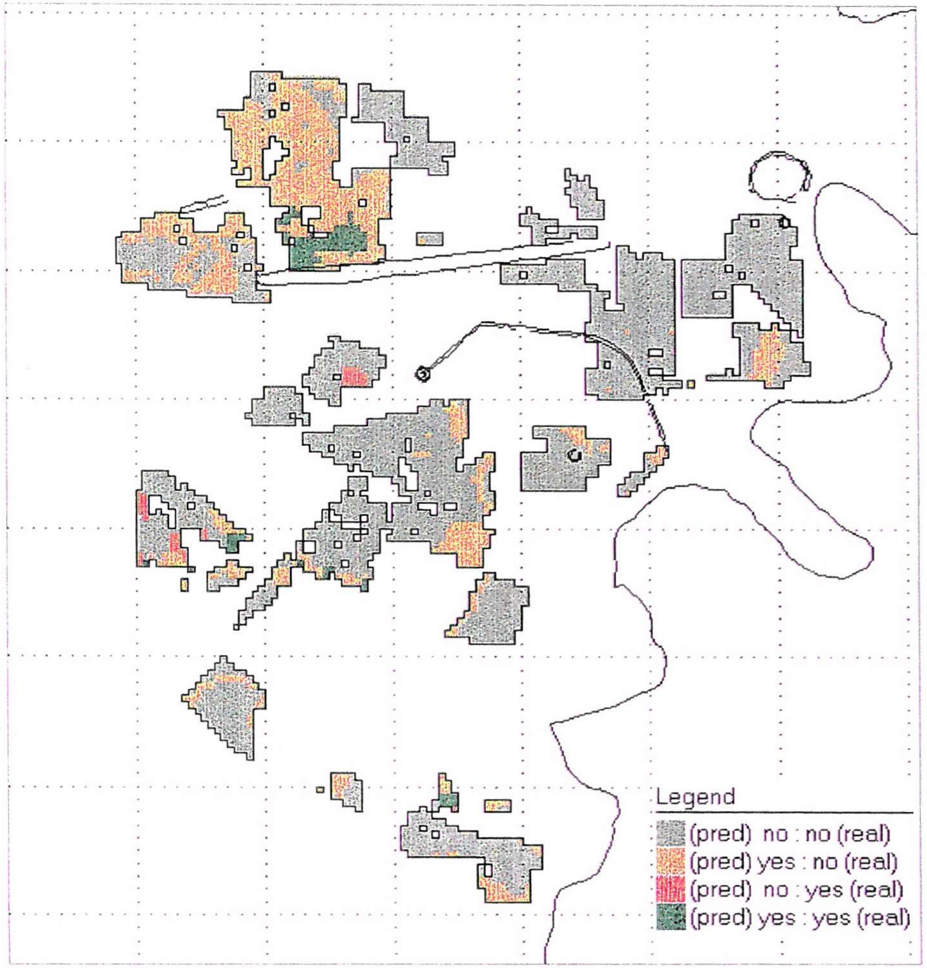


Figure 6.40. Comparison of optimum prediction of 'manufacturing areas' with observed responses

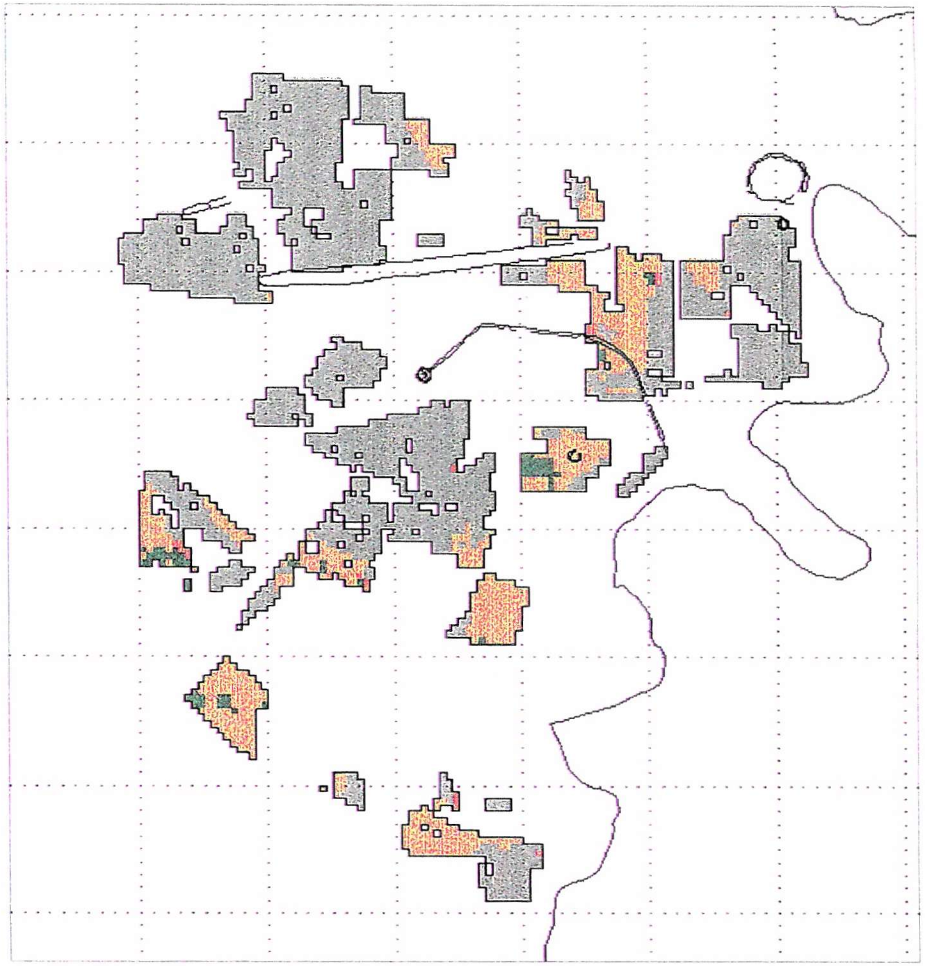


Figure 6.41. Comparison of optimum prediction of 'discard areas' with observed responses

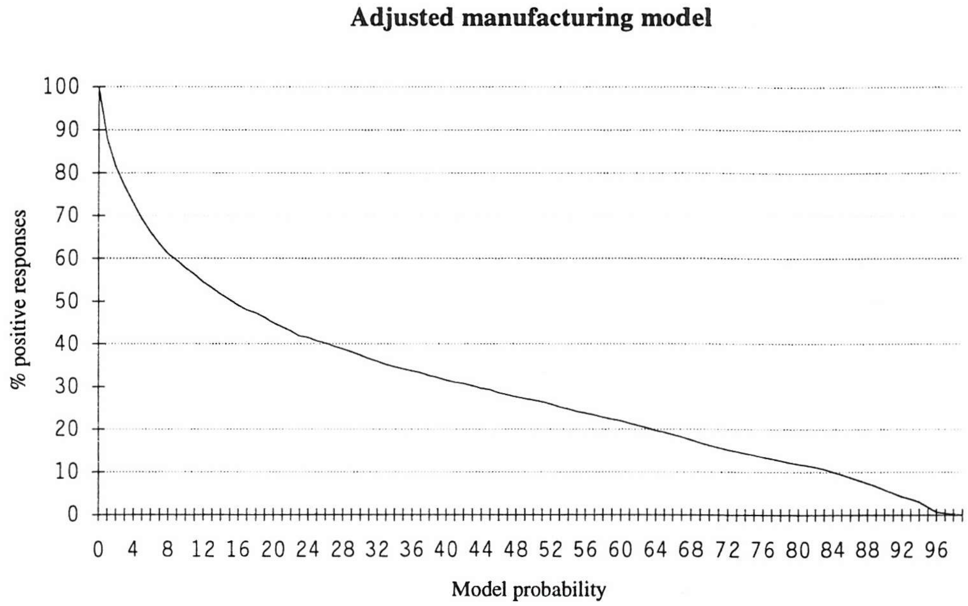


Figure 6.42. The proportion of the surveyed area which is predicted as a 'manufacturing area' for each probability result from the adjusted manufacturing model.

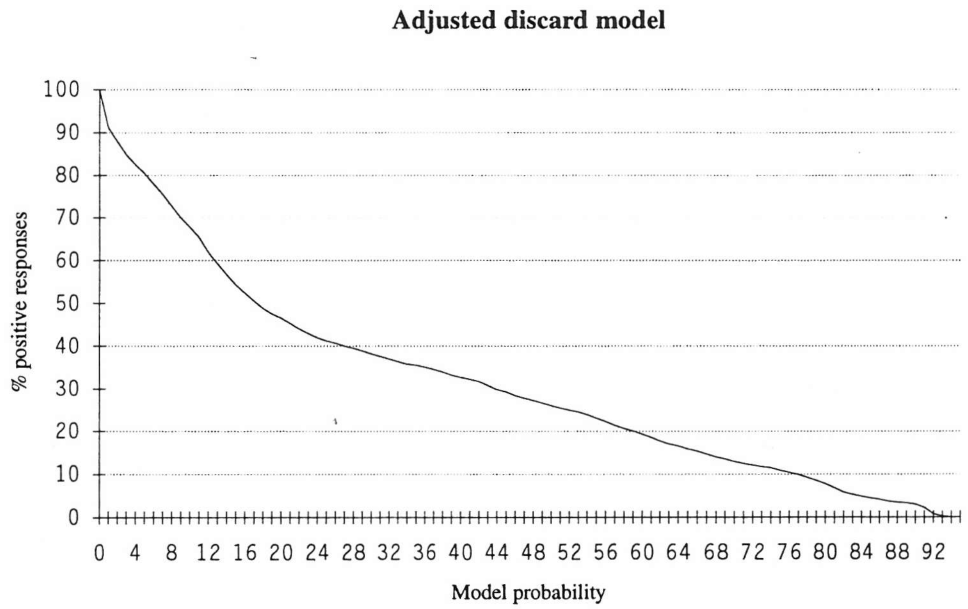


Figure 6.43. The proportion of the surveyed area which is predicted as a 'discard area' for each probability result from the adjusted discard model.

In this case, this would involve discarding approximately 90% of the data which is clearly unacceptable. The second approach is to attempt to correct for the sample size bias after running the model. This can be done by adjusting the intercept parameter by the natural log of the ratio of the sample sizes (Kvamme 1983b). Adjusting the two models in this way produces the results shown in figures 6.38 and 6.39. Response curves for these adjusted models are shown in figures 6.42 and 6.43, showing the gentler response slope which results.

It remains to decide on an appropriate threshold for prediction of sites from the adjusted models. It has already been seen that prediction of the 'correct' number of positive responses will not provide an good prediction of presence. Increasing the threshold which is used as a prediction increases the number of incorrect predictions of negative responses but increases the number of correct predictions of positive responses.

The optimum solution must therefore be sought from the models, and occurs when the proportion of correct predictions for positive responses is equal to the proportion of correct predictions for negative responses. This probability value is referred to as the cut-off for the model, and can be obtained by graphing the observed proportions of correct positive and negative responses against the predicted probability as in figures 6.44 and 6.45. Observations for these graphs were made by repeated application of the models at appropriate intervals. The point where the positive and negative response curves cross provides both the optimum point of the model (which will be 50% in a model with equal sample sizes) and the percentage of correct predictions. Using the closest round numbers obtained from figures 6.44 and 6.45, the cut-off point for the manufacturing model occurs at 15%, where the model correctly predicts 71% of positive responses (manufacturing areas) and 72% of negative responses. For the discard model, the cut-off occurs at 44% where the model predicts 73% of positive responses (discard areas) and 74% of negative responses. This response can be mapped as above, and is shown in figures 6.40 and 6.41.

Using the logistic models

The final models shown in figures 6.38 and 6.39, provide a method for assessing how likely it is that any location within the landscape would contain a lithic assemblage with either of two particularly interesting charactaristics. Clearly the most obvious application of such a model is to the management of the archaeological resource in which they might be used as a method for assessing the relative impacts of alternative courses of action on the (unknown) archaeological resource. However, although using them for this purpose is valid, the models must be understood before they can be best utilised. For one thing, the predictions for manufacturing areas in the far northwest and southeast of the study area are spurious, and should be disregarded. An alternative model may be developed for areas which are marginal or outside the cluster of monuments, but these two are not applicable for the reasons outlined above.

The models can be used to make predictions in a sophisticated way. The optimum predictions which were obtained from the cut-off points of the models may not be the most useful predictions from an archaeological point of view and alternative thresholds might be chosen in order to predict more sites at the expense of nonsites. Given enough resources, of course, an archaeological management strategy would not require a model at all but include provision to survey all areas. In reality, however, management strategies are constrained by resources which in turn restrict the area which may be surveyed. Depending on this, thresholds may be defined which progressively reduce the area which needs to be surveyed, while maximising the likelihood that the interesting areas will be within them. For example, although the manufacturing model predicts roughly 70% of the sites at the cutoff point of 45%, a threshold of a 20% may be selected to obtain a prediction which accounts for 90% of the manufacturing areas at the expense of predicting 40% of the non-site areas as manufacturing areas also.

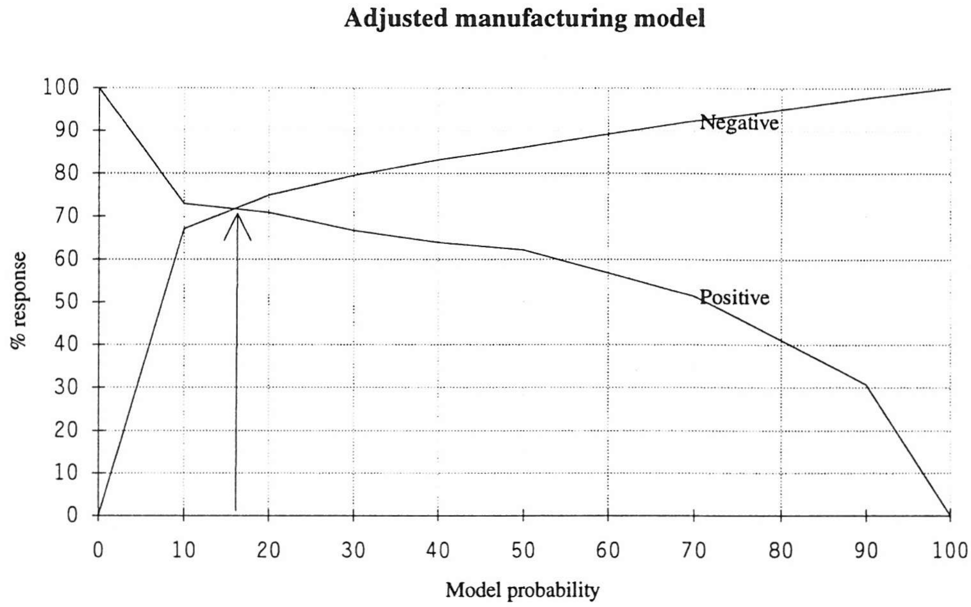


Figure 6.44. Cut-off point for the adjusted manufacturing model

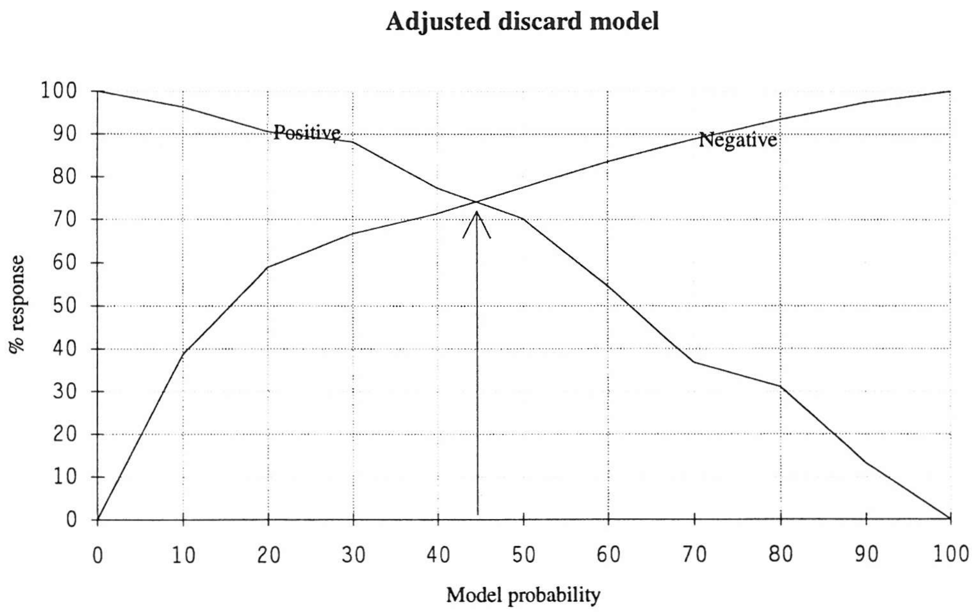


Figure 6.45. Cut-off point for the adjusted discard model

6.8. Conclusions

The aims of this chapter were set out in section 6.1, and were to identify whether GIS could be of value in the interpretation of extensive survey data. The three areas of interest were identified as visualisation of the lithic data, in order to identify interesting trends within the lithic data itself, interpolation of the lithic data from the surveyed areas into unsurveyed areas, and correlation of the lithic data with other data types in order to construct predictive models.

Section 6.4 showed how the visualisation of the lithic density data within GIS can provide valuable summaries of the spatial patterns within the data. The ability to use image-processing derived functions to smooth noisy data aided the observation of lithic clusters, and allowed some preliminary suggestions as to what the visible pattern may represent. It should be emphasised that although the printed output presented here is useful and flexible, it does not begin to reflect the utility of interactive visualisation within the system, with the ability to zoom and pan, to alter colour combinations and to experiment with different combinations of variables.

In general the density maps confirms the interpretation of Richards (1990), although the core/retouch ratio perhaps suggests that the study area exhibits a tripartite division of the landscape. Firstly there is clear evidence for the 'manufacturing zone' at the eastern end of the Cursus, which may extend southwards as suggested by Richards. Secondly, there seem to be a series of discrete 'discard areas' to the southeast of Stonehenge itself (51, 87, 57), the precise interpretation of which is open to question. Thirdly, the dry valley between Winterbourne Stoke Crossroads and Normanton bottom may have been an area of mixed activity, and shows signs of both manufacture and discard activities.

In order to extend interpretations beyond the areas which were surveyed, both interpolation and modelling were attempted. The interpolation results from inverse distance weighting seem to provide the best estimates of lithic densities close to the surveyed areas, and are useful as estimations of the extent of observed lithic scatters. However the results of both the interpolation approaches tried rapidly decline in value with distance from the observations.

The result of the attempt to use linear multiple regression to predict the lithic densities from correlated variables is, from this point of view, disappointing. That only 25% of the variability within the flint data can be accounted for in this way suggests strongly that the types of deterministic models applied with some success to Mesolithic data will not provide explanations for the behaviour of Neolithic and Bronze Age people in very structured landscapes. It is particularly revealing, however, that although it is overall low in predictive power, the majority of the predictive power of the model is accounted for by cultural indices, particularly distance to barrows, rather than the indices of landsform which have been used by, for example, Kvamme (1989). The explanatory power of the logistic regression models is far more difficult to assess, but is likely to be of the same order as the multiple regression because the same predictors were used for both.

In both cases, the portability of the models is compromised by the use of predictors which seem to have complex relationships with the data. The extreme predictions in both the linear multiple regression model, and both logistic models seem best explained as 'over training', in the sense that the model is too specifically related to the variation *within* the cluster of monuments to provide a useful estimate of the density of lithics *between* monument clusters. In a sense this is a problem of scale: at the local scale, the relationship between density of monuments and activity may be that the activity takes place away from monuments while at a regional scale this relationship cannot be sustained.

One last, but currently untested observation might stem from this conclusion. Although it seems that it is of no use extrapolating the predictions beyond the cluster of monuments on Salisbury Plain, because this is altering the scale of the prediction required, it might be quite appropriate to use the same model to predict lithic densities within another monument cluster such as Avebury or the Dorset Ridgeway - although to do so would require some evidence that the same relationships hold for these areas also. If this proved to be the case, then it may well be that portability of scale must be regarded as a different problem to portability of region, and that the former is at least as great a problem as the latter.

Chapter 7
Conclusions

7.1. Introduction

It is the aim of this final chapter to draw upon the interpretations and conclusions which accompany the case studies in order to develop some points which are of wider interest. Several themes will have become apparent in the previous chapters, many of which do not fit comfortably within the artificial division of the thesis in to 'archaeological' and 'technical' chapters, and into the study of 'monuments' and 'field data'.

The objective of this thesis has not been to dictate the direction that archaeological applications of geographic information technology should take. Rather it has been to explore the nature of current practice, and to suggest that there are other areas of archaeology which might be developed and enriched by the use of technologies such as GIS. If the case studies prove nothing else, therefore, it is to be hoped that they provide a provocation and an incitement to other archaeologists who share an interest in the application of information technology. Likewise, the archaeological aim of the thesis has not been to present a holistic explanation of the Neolithic of Wessex, or to settle once and for all the debates about the period. Instead, the approach has been to highlight some areas of the archaeological record which deserve greater attention than they have been given to date and to use the insights gained from GIS methods to provide a new reading of the archaeological remains.

7.2. Scale and resolution

Scale, in a variety of senses, has been a continuing theme of this work. Most obviously, the thesis has been concerned with geographic scale and the associated issue of spatial resolution. Intuitively, the larger the spatial scale at which we wish to study archaeology, the lower the spatial resolution of our inferences must be. GIS methods for representing archaeological material reinforce this opposition through limitations of processing speed and storage capacity: in the raster approach used extensively here, for example, the larger the spatial area of interest, the larger the units of landscape must be in order to manage it effectively.

Because of this limitation on spatial resolution, GIS studies of cultural phenomena cannot ask the same questions of archaeological evidence at one scale as at another. In a small-scale study, it may be possible to treat a long barrow as an entity with length and extent, but at a larger scale the barrow must become a point location: the barrow is represented as being all in one place. The representation of space is therefore different depending at which scale we wish to consider the material.

Contrary to expectation, however, this is by no means a disadvantage of the GIS way of doing things. Both theoretically and methodologically it is not an unfortunate by-product of the restrictions of computer storage but a positive advantage. To consider the example above, of the long barrow, the GIS representation of a barrow at a small scale actually mirrors the experience of being close to the monument: it is large, it takes time to walk around it and it has spatial extent. Constructing it may take some land out of production so the area of the monument at this scale is an important consideration. Thus at a small, local scale it is right that the barrow must be represented in a way which recognises that it has length, breadth and height. At a larger scale, however, perhaps the scale of a traveller moving about the entire region the barrow will be considered differently. It may be a marker on a journey, for example, so that a traveller will set out on a walk knowing that when a particular barrow is passed the journey is half completed. At this scale of reference, therefore, the human agent is also like the GIS in that the extent of the barrow is forgotten and it is regarded as a single place of reference. Cognitive space, therefore, is not the same as physical space: cognitive space varies the representation of objects according to the scale and context. The GIS should be used to represent models of cognitive space if is to have any real interpretative power in archaeology and in a very real sense this is what it does.

Methodologically also the scale of study must be a central consideration. In the analysis of lithic densities around Stonehenge this problem was brought into sharp focus by a relationship between lithic density and proximity to monuments. At the scale of the Stonehenge Environs Project database the relationship seemed to be simple: higher lithic densities tended to occur in the spaces between the monument. This relationship cannot, however, be extended to the analysis of lithic densities at a different scale because this does not imply that flint densities will be at their highest between the major monument clusters. At this scale, we could generalise a little and claim that the lithic densities were at their highest close to the monuments. This conclusion regarding the relationship between monument locations and flint working activity means that comments regarding such relationships must be tempered by the scale at which the relationship is claimed. Thus when survey data is interpreted by Holgate (1988) to indicate that

> *'At the start of the Neolithic, domestic activity shifted on to the uplands and coincided with a phase of tomb construction in these areas'*

then we are entitled to ask at what scale the tomb construction is 'in these areas' and whether that relationship can be sustained at other scales of reference. Similarly Gardiner's (1984) conclusion that, in Sussex the distribution of early Neolithic leaf-shaped arrowheads clusters significantly around long barrows as compared with enclosure sites, and that virtually all surface scatters and stray arrowheads are located on or near to clay-with-flints deposits can only be accepted as a conclusion which holds for the particular scale of study which was conducted.

The flint densities raise the issue of portability as well as scale. Clearly the relationship between distance from monuments and lithic density was helpful in the construction of a model within the Stonehenge Environs area itself, but equally clearly the results could not be extended far beyond the edges of this region. Not only is the relationship between flint and monuments not scalable, therefore, in that it is of little direct value in the interpretation of the larger regional complex, but it is not *portable* either. If an area of the same size as the Stonehenge Environs Project database were to be delineated in, for example, the Vale of Pewsey then we could have no confidence that the same relationship would hold. It follows from this that GIS methods which archaeologists borrow or derive to analyse the spatial distribution of cultural materials must be related to both the scale of study and to the context of the study.

7.3. Scale and social organisation

Scale in a rather different sense is also an issue in the consideration of the distribution of the long barrows. The social-evolutionary approach which led to the definition of the 'tribes' and 'chiefdoms' of the Wessex chalklands (Renfrew 1973) was concerned only with the scale of socio-economic groups. Nowhere in the explanation of the spatial pattern is there a concern with human agency: those individual actions and practices which formed the archaeological record. In this sense, therefore, the Neolithic tribes and chiefdoms are a product of a social macro-scale interpretation.

I would claim that the view presented here is built from a smaller social scale than this. Where the chiefdom model is imposed from the top down by importing a completed model and testing for correlates, this reading of the Neolithic monuments is built from the bottom up. The social scale of the study considers each activity at each barrow as an event within a sequence of barrows. The model of barrow intervisibility presented in chapter 5 relates the practice of selecting locations for monumentalisation with the rest of the cultural landscape. Each single choice of location is considered at this scale because it assumes individual conscious choice of place is followed by an individual conscious and meaningful action. The consequences of these actions put together over a long period of time may be observed, but the interpretation of the pattern is that of practices which

> *'maintained or developed particular lines of authority and inheritance between the living, lines of authority which could be drawing upon in other fields of social action.* (Barrett 1988 p40)

The difficult part is to obtain insight into the motivations and the *intended* consequences of the actions from the study of the effects of all the consequences, both *intended* and *unintended.*

Specifically, though, the thesis should suggest that the locations of the long barrows on Salisbury Plain are not best understood as 'territorial markers'. If they were territorial we might expect them to show the pattern which Lock and Harris (forthcoming) have claimed for the barrows around Danebury would be apparent - each of the barrows would be sited to exert control over a single territory and intervisibility would be suppressed. Instead the barrows should be seen partly as visual references to the ancestors and to the power and legitimacy which they stood for.

More generally, the evidence of the monuments does not provide a 'record' whose characteristics may be correlated with those of past social institutions such as tribes or chiefdoms. Rather the monuments, like all archaeological remains, provide evidence of social practices and actions. Through the interpretation of these practices by observing their effects over long periods of time, it may be possible to consider the nature of social organisation. In this context, it is hoped, the approach to regional analysis accords with the desire of Barrett to

> '*... develop detailed regional analyses which attempt to build syntheses out of all the available data which relate to these different fields. In this way we may learn something about the processes of social reproduction without recourse to general, and spurious, models of social totalities.*(Barrett 1988 p40)

7.4. Temporality and chronological resolution

Superficially, a genuinely temporal approach to GIS analysis might initially the generation of 'time slice' overlays to represent the different phases to which material remains could be ascribed. In the same way that finer and finer resolution of raster data will eventually allow the modelling of a discrete model of continuous space, finer and finer resolution in temporal time slices would eventually allow a discrete model of time in the GIS. Once stored in the system, this could be used to perform analyses of chronological process much as current GIS methodology allows performance of spatial interactions between layers.

Of course spatial processes are not the same as chronological processes: unlike most spatial phenomena the processes associated with time are unilinear so that while events at one time will influence future events, this relationship does not hold in reverse because the future cannot influence the past. This, of course, assumes that the western scientific understanding of measured linear time is the most appropriate model for analysis of chronological process in GIS. Castleford (1992) has suggested that it need not be: time as considered by the Hopi Indians of the south-western US, for example, has only two components: those things which are happening and those which have happened. In this framework the future is considered as part of a separate, spiritual world. Australian aboriginals also hold a wholly different view of time to the European measurable 'commodified' view of time which has prevailed since the industrial revolution.

Even if the western idea of time is most appropriate, however, the methodology for such explicitly chronological modelling does not currently exist in GIS, and some theorists have come to the conclusion that *'the present structure of gis is atemporal and consequently only* *able to deal with spatial phenomena'* (Castleford 1992) and that any archaeological analysis of any temporal sophistication with GIS would therefore have to await the development of a Temporal Geographic Information System or TGIS.

This is a naive simplification, however. It is not necessary to create a TGIS to adopt a temporal approach to archaeological material. The analysis of the long barrows in chapter 5 is explicitly temporal in approach even though no such linear chronological model was presented. The idea that barrows constructed in the past may have influenced the choice of location for new barrows is inherently temporal - it is a model which cannot be sustained unless the monuments form a chronological sequence. The analysis also requires that there is a change in the function of barrows not just from one place to another place but from one time to another time. When they are constructed they function (among other ways) as the place from which other barrows may be viewed; time then passes and the monuments themselves act as the source of the visual impact. The model does not provide a description of the chronological sequence of barrows, it is true, because such a sequence is not available with the current level of knowledge about these monuments. Instead the analysis concentrates on an explanation of the individual practices which are evidenced by the selection of place and the construction of the tomb in terms of the cumulative effects of such practices over a long period of time.

This raises another issue, that of chronological scale. As in the discussion of spatial scale and resolution above, the 'poor chronological resolution' of the archaeological record might be a problem in the reconstruction of the development of Neolithic societies or it may be a reflection of the different methods and approaches which are required in order to understand change over long and short periods of time. It is the difference in the time scales which archaeology deals which causes the greatest theoretical difficulties, archaeology constantly grapples with the problem of:

> *'different scales of analysis revealing different cycles of historical time, ranging from the short-term histories of moments, events and individual lives, to the medium-term cycles of economic and political process and, finally, to the long-term, underlying currents of natural evolution and ecological change.'* (Barrett 1994 p2).

This difference in time scales is highlighted in the difference in approach which is required for the monument locations considered in chapter 5 and the lithic densities of chapter 6. In this case the spatial scale of the survey data analysis of chapter 6 is smaller than that of chapter 5, while the temporal scale of the explanation must be greater because of the nature of extensive survey data. Although neither of these sources of archaeological data are of a very high chronological resolution (at least in comparison with excavation data, for example) the lithic survey data is not amenable even to the level of chronological explicitness of the monuments. While we know, in general, the sequence of construction of long barrow, enclosures, cursus, henges and round barrows we cannot hope to identify the same sort of sequences in the extensive survey data used in chapter 6 without recourse to the very small number of type artefacts. Although the resolution of the explanation is different, however, the goal remains the same in both studies: to understand what individual practices (discard or manufacture) were undertaken in different parts of the landscape over a long period of prehistory.

7.5. Economic and ceremonial landscapes

7.5.1 Enclosures

The interpretation of the long barrows and lithic scatters around Stonehenge has suggested that existing GIS technology is more than adequate to undertake meaningful archaeological analysis if an appropriate theoretical basis is adopted. The 'cognitive models' of the enclosure sites in chapter 5, particularly, have suggested that GIS might successfully be extended beyond the interpretation of economic remains and of non-sedentary societies. Rather the organisation of the sort of 'ritual landscapes' which survive on the Wessex chalklands may be attempted.

These reveal how the three enclosures of Avebury region, far from being the 'central places' which Renfrew's (1973) chiefdom models of the social development of the period assume, are (geographically at least) entirely marginal in nature. This finding also weighs against the economic explanation of the enclosures advocated by Barker and Webley (1978), who argue that the locations were chosen because of the availability of suitable pasture or agricultural territory which surround them. Barker and Webley's use of site catchment analysis is clear in its assumption that the enclosures are the centres of economic exploitation territories: a view which can no longer be sustained.

This should not be a surprising finding when the artefact assemblages of these places are considered: exotic 'oolitic' and 'gabroic' pottery was found at Windmill Hill, for example, and to treat the enclosures as economic or social centres is to ignore the considerable evidence that the enclosure sites were not primarily domestic in function but were places where a variety of activities took place including feasting, rites of passage and the deliberate burial of artefacts. Even the concentric low ditches and banks of the sites seem to emphasise the desire to enclose and incorporate rather than to impose. The situation of the three enclosures at edge of the anthropogenic landscape may be more psychological than functional, they may have been places where 'dangerous' things from outside the local community could pass into the community safely after appropriate rites of passage. The evidence of the ritual feasting, breaking of pottery and the existence of non-native pottery at these places is therefore entirely in agreement with their geographic position at the margins of the landscape, the places mark the physical delineation of 'inside' the community and 'outside' which artefacts symbolically pass through.

If the enclosures are marginal rather than central, another possibility which was briefly considered in chapter 3 cannot be ruled out. That is that some of the sites might have had a defensive function. Certainly this is less convincing in the case of Windmill Hill than for Rybury or Knap Hill. These latter two sites are perched on the edge of a steep chalk escarpment and must be approached up very steep slopes from outside the Avebury area but far less so from the North and from the direction of the Avebury 'centre'. The time contour maps, however, actually suggest that the sites are easiest to approach from the south rather than the north, and while this might reflect that they are strategically positioned to control the approaches to the Avebury area, it is strange that they are not situated to make an easier approach from the north.

7.5.2 Flint working and ceremonial monuments

This division of landscape (and material things) into 'inside' and 'outside' can be compared with the evidence of the for different activities in different parts of the Salisbury plain landscape. Here the relationship between lithic densities and monument locations may be a division between areas appropriate for the construction of barrows and areas appropriate for flint working activity. There is a parallel with Cranborne Chase, where there also seems to be a separation between the areas of greatest monument-building activity and of highest concentration of flint scatters (Bradley *et al* 1984) but superficially this stands in contrast to the situation in Sussex where Neolithic flint arrowheads seem to cluster around barrows, and avoid enclosures (Gardiner 1984).

This division of the landscape into flint working areas and monumental areas may be a distinction between the 'sacred' and the 'profane': in other words the pattern may be evidence that the Stonehenge landscape was divided into ritual and domestic locations. This difference, however, is observed at a large chronological scale suggesting that those patterns of behaviour which were established in the earlier Neolithic with the construction of the first monuments persisted for many hundreds of years, and over many generations.

The 'visual impact' map of Salisbury plain might also be considered in light of the lithic density analysis. There is an area towards centre of the Cursus which shows an extremely low visual impact of long barrows and this is interestingly close to the manufacturing area revealed by the lithic density analysis. It is tempting to speculate from this that the manufacture of flint was deliberately conducted in a place away from the gaze of the long barrows. It is, of course, possible that this association is entirely spurious and that the low visual impact is caused by the low-lying nature of the terrain which in turn caused the selection of the area for manufacture of flint. However, if it could be supported, then such a relationship would reveal much not only about the division of the geographic landscape, but about the spatial separation of practices in the Neolithic. At the very least, and in accordance with the relationships suggested by Bradley *et al* (1984) for Cranborne Chase and by Gardiner (1984) for Sussex, this study suggests that relationships between deposition of worked flint and monuments are a recurrent feature of this period.

As well as differentiation between 'flint areas' and 'ceremonial areas', there is also clear evidence for further differentiation of the landscape within the flint areas. Here the division seems to be into areas of 'discard' (which could be domestic sites or the locations of deliberate deposition of knapped flint) and areas of 'manufacture'. The division between the zone of manufacturing and the area of discard is reinforced by the position of the cursus which, as well as acting as a processual way, seems to form a visual barrier between the area of manufacture and of discard.

Although the lithic density data does not extend sufficiently far north to draw any real conclusions the relationship between lithic distributions and the enclosure, the explanation put forward above may not entirely accord with the view of Thomas (1991) that the Cursus monument was constructed to separate the barrows north of cursus (which are associated with the activities at the enclosure) from those to the south which are associated with domestic activity. The evidence of the majority of the lithic assemblage makes it more likely that the barrows around Stonehenge are actually *disassociated* with flint working altogether and that the cursus separates the landscape not into 'domestic' flint working areas and 'enclosure related' areas but, perhaps, into areas which relate to manufacturing of flint artefacts and deposition of flint artefacts. However, only an extension of the systematic fieldwalking to the north can genuinely add more light to the obviously complex relationship between flint, barrows and enclosures.

7.6. GIS technology and methods in archaeology

Whatever the value or otherwise of the case studies for the particular study of the Neolithic archaeology of the study areas, some conclusions regarding the application of GIS to British archaeology can now be presented. It was one of the stated aims of the thesis to develop methods of approaching archaeological material which stemmed from archaeological questions, rather than to import methods which were developed to answer other kinds of questions into archaeology. This was achieved in chapter 5 in two ways. The archaeological interpretation of the social function of the enclosure sites was based on assumptions about their centrality, therefore an approach to the definition of centrality was derived and used to argue for the marginality (for want of a better word) of the causewayed enclosures in the study areas. Similarly Cumulative Viewshed Analysis, was not and could not have been devised within another field of enquiry, because it was created to investigate the fundamentally archaeological question 'why did people choose to build this monument here rather than elsewhere?'. It was a side-effect of trying to answer this question that produced the method.

Admittedly, the methodological components from which these methods were constructed already existed, and these were not devised to answer archaeological questions: cost-surface analysis was developed, and continues to be developed, primarily to model market catchments for retail outlets, and the binary line-of-sight maps which were needed for cumulative viewshed analysis of barrows have been used for a variety of purposes by, for example, architects and soldiers. But by regarding these components not as methods in their own rights, but as parts of the technology of GIS, it is possible to use them to generate entirely new methods for archaeological analysis such as those in chapters 5 and 6. In a sense, therefore, the central conclusion of the thesis must be that it is the *technology* of GIS which offers most to archaeologists, while the associated *methodology* should be regarded with suspicion, although some GIS methods can now be regarded as part of the technology.

In many ways the issue of data-driven ('bottom-up') versus theory-driven ('top-down') archaeological method becomes redundant within this framework. The aims of adopting either a 'top-down' strategy in which theories about archaeological material lead to data collection, or a 'bottom-up' strategy in which patterns in the data itself suggest theories which may then be tested are both unattainable. GIS applications in archaeology cannot be driven from the bottom up by the data itself because all data is selected within a theoretical framework (explicit or not) and represents selected, and hence subjective, observations: data is theory-laden. Similarly theories about past cultural phenomena cannot arise out of nothing in the absence of archaeological observations. The approach to GIS applications in archaeology adopted here is to accept that the relationship between theory and data is complex and bi-directional, but that this relationship produces questions which can then be used to devise methods, and to obtain data to apply them. The second fundamental conclusion of this thesis is therefore that archaeologists who wish to use GIS technology in archaeological investigations should adopt neither a theory-driven nor data-driven approach but, instead, should adopt a question-driven approach. Within this general approach, several aspects of GIS technology have shown themselves to be of particular value to archaeological analysis, while some have problems.

Chapter 6 revealed how the *data visualisation* capabilities of GIS can be used to great advantage in the interpretation of large archaeological datasets. It is a characteristic of archaeological field survey that it produces a very large number of observations, and the sheer weight of numbers can make the observation of patterns within these datasets extremely difficult. One of the criticisms occasionally made of GIS applications is that it produces very attractive graphical results which often have little concrete analysis underlying this. If 'analysis' in this sense can be taken to mean 'processing', then this is not a criticism at all but a useful characteristic of the method: that it allows large quantities of data to be rapidly assimilated and presented as a graphical summary. The generation of the core/retouch ratios in chapter 6 is a good example of how little 'number crunching' actually needs to be done in order to visualise a pattern within a large dataset which might be obscured by the sheer number of observations.

The role of predictive modelling methods within the context of British archaeology is far from clear at this time. Although these techniques, and particularly approaches based on probability models, have apparently been applied with some success to archaeological management contexts in the US, the experience of chapter 6 suggest that there may be little to be gained in a management sense from their importation into British archaeological management practice. In part this may reflect the difference in character between the archaeological material to which predictive modelling has been applied in the US and Britain: in the US there are large geographic areas in which large proportions of the areas are entirely unsurveyed. In these situations, such as the Arizona dessert, there can be no question that the entire area could be surveyed prior to development and the aim of the model is to gain a 'coarse resolution' picture of the distribution of archaeological material without, necessarily, a high degree of certainty. In Britain, however, the areas which archaeological managers are concerned with are generally smaller and have often been surveyed and re-surveyed for generations. The problems of 'knowing where the archaeology is' are therefore entirely different in nature in these areas: it is not those places in which archaeologists *expect* to find archaeological material which cause greatest management problems but those where archaeological remains are located *unexpectedly*.

On the other hand, there is considerable research merit in extending our understanding of why particular distribution patterns of archaeological material are observed. In this the ability of GIS to manipulate large amounts of observations is again central, because it allows predictions and models to be constructed from the majority of the archaeological assemblage rather than a small selected part of it. The main conclusion of chapter 6 may therefore be that, while predictive models might fail to translate directly from the US context to British archaeological management, they do offer considerable insights in the area of 'offsite archaeology' (Foley 1981).

But if GIS allows the archaeologist the option of interpreting archaeological remains without reference to 'sites', as recommended by Foley (1981) and Schofield (1988), the links made above between chapters 5 and 6 also provides an example of how to then relate the findings of an offsite approach to the evidence of more traditional site-based archaeological practices. While behavioural, off-site archaeological analyses such as those in chapter 6 are clearly of interest, the results take on a new perspective when they are placed in the wider context of 'Landscape Archaeology' (e.g. Roberts 1987). It is therefore the final and overriding conclusion of this thesis that GIS offers archaeologists the opportunity to bring together within one analytical framework the insights of functional, behavioural approaches to the distribution of cultural material and the interpretation of ceremonial and ritual practices within the landscape.

Appendix I
IDR2SPA

I.i. Description

The IDR2SPA program was written to translate idrisi format vector files to spans archive vector files (each is an ascii format for vector data). The program is written in the Clipper dialect of the dBase language, and was compiled with the Summer '87 release of Clipper. An executable file which runs on all MS-DOS machines may be obtained via the internet by anonymous ftp, or gopher protocols from AVEBURY.ARCH.SOTON.AC.UK (give 'anonymous' as username and your e-mail address as the password). It is situated in the /pub/gis/utils directory as both source IDR2SPA.PRG and binary executable IDR2SPA.EXE.

I.ii. Code listing

```
* ------------------------------------------------
* idr2spa.PRG
* ------------------------------------------------
* Reads an Idrisi  (Version 2.0 Ascii) vector file, writes SPANS
* .vec (vector) file and .veh (vector header) file.

* Constants
* ---------
set exact on
setcancel(.t.)
linefeed=chr(13)+chr(10)


* Parameters and parameter checking
* --------------------------------
? "IDR2SPA - Dave Wheatley 1992"
?
if pcount()=2
        parameters infile,outfile
        if infile=outfile
                error(1)
        endif
        if ("."$infile) .or. ("."$outfile)
                error(2)
        endif
else
        if pcount()=1
                parameters switch
                if switch="/h" .or. switch="/?" .or. switch="-h" .or. switch="-?"
                        do helptext
                endif
        endif
        do syntaxtext
        quit
endif

* Open input files
* ----------------
? "Reading IDRISI files ..."
ivecname=alltrim(infile)+".vec"
idvcname=alltrim(infile)+".dvc"
ivechandle=fopen(ivecname,0)
        if ferror()<>0
             error(3)
        endif
idvchandle=fopen(idvcname,0)
        if ferror()<>0
             fclose(ivechandle)
             error(4)
        endif

* Read .dvc file into memory
* ---------------------------
declare dvcline[13]
```

```
for counter=1 to 13
        dvcline[counter]=freadline(idvchandle)
next
fclose(idvchandle)

* Open output files
* -----------------
svecname=alltrim(outfile)+".vec"
svehname=alltrim(outfile)+".veh"
svechandle=fcreate(svecname,0)
        if ferror()<>0
             fclose(ivechandle)
             error(5)
        endif
svehhandle=fcreate(svehname,0)
        if ferror()<>0
             fclose(ivechandle)
             fclose(svechandle)
             error(6)
        endif


* Get info from IDRISI .dvc file
* ------------------------------
if .not. ("ascii"$fieldvalue(dvcline[3]))
        fclose(ivechandle)
        fclose(svechandle)
        fclose(svehhandle)
        ? error(7)
endif

* Obtain object type and min/max dimensions.
* ------------------------------------------
idrtype=fieldvalue(dvcline[4])
if idrtype="polygon"
        objecttype="P"
        ? "Objects are polygons."
elseif idrtype="line"
        objecttype="A"
        ? "Objects are arcs."
else
        fclose(ivechandle)
        fclose(svechandle)
        fclose(svehhandle)
        ? error(8)
endif

for counter=5 to 13
    if fieldtitle(dvcline[counter])="min. X" .or. fieldtitle(dvcline[counter])="min X"
        minx=round(mathtran(fieldvalue(dvcline[counter])),4)
        maxx=round(mathtran(fieldvalue(dvcline[counter+1])),4)
        miny=round(mathtran(fieldvalue(dvcline[counter+2])),4)
        maxy=round(mathtran(fieldvalue(dvcline[counter+3])),4)
        exit
    endif
next

? "minx ="+alltrim(str(minx))+ ", maxx ="+alltrim(str(maxx))
? "miny = "+alltrim(str(miny))+ ", maxy = "+alltrim(str(maxy))

* Translate .vec file into spans format
* -------------------------------------
? "Writing SPANS files ..."
set cursor off
fwrite(svechandle,"ARCS"+linefeed)
numarcs=0

firstline=freadline(ivechandle)
if (" "$firstline)
        * Version 3 onwards format
        * ------------------------
        ? "Version 3 format (two values per line) detected."
        ? "Arcs completed:    0"
        continue=.t.
        do while continue=.t.
                if numarcs<>0
                        firstline=freadline(ivechandle)
                endif
                arcvalue=getfirst(firstline)
                numpoints=getsecond(firstline)
                if arcvalue=0 .and. numpoints=0
                        continue=.f.
                else
                        numarcs=numarcs+1
                fwrite(svechandle,alltrim(str(numarcs))
                               +" "+alltrim(str(numpoints))
```

```
                              +" "+alltrim(str(arcvalue))+linefeed)
                        for counter=1 to numpoints
                              firstline=freadline(ivechandle)
                              fwrite(svechandle,alltrim(str(getfirst(firstline)))+" ")

        fwrite(svechandle,alltrim(str(getsecond(firstline)))+linefeed)
                        next
                endif
                ?? chr(8)+chr(8)+chr(8)+chr(8)
                ?? str(numarcs,4,0)
        enddo
else
        * Version 2 format
        * ----------------
        ? "Version 2 format (one value per line) detected."
        ? "Arcs completed:    0"
        continue=.t.
        do while continue=.t.
                if numarcs<>0
                        firstline=freadline(ivechandle)
                endif
                arcvalue=val(firstline)
                numpoints=val(freadline(ivechandle))
                if arcvalue=0 .and. numpoints=0
                        continue=.f.
                else
                        numarcs=numarcs+1
                        fwrite(svechandle,alltrim(str(numarcs))+" "+str(numpoints)+"
"+str(arcvalue)+linefeed)
                        for counter=1 to numpoints
                              fwrite(svechandle,freadline(ivechandle)+" ")
                              fwrite(svechandle,freadline(ivechandle)+linefeed)
                        next
                endif
                ?? chr(8)+chr(8)+chr(8)+chr(8)
                ?? str(numarcs,4,0)
        enddo
endif
fclose(ivechandle)
fclose(svechandle)

* Write SPANS .VEH file
* ---------------------
fwrite(svehhandle,"FILE_TYPE "+objecttype+" X C 0 1"+linefeed)
fwrite(svehhandle,"VERSION 4.02"+linefeed)
arcsheader= alltrim(str(int(numarcs)))+" 3 "+str(minx)+" "+str(miny)+" "+str(maxx)+"
"+str(maxy)+" 3 3 0 0 0 0 0"
fwrite(svehhandle,"ARCS_HEADER "+arcsheader+linefeed)
fwrite(svehhandle,"EXTENTS "+" "+str(minx)+" "+str(miny)+" "+str(maxx)+"
"+str(maxy)+linefeed)
fwrite(svehhandle,"REFERENCE 1 1 3 1 1 0"+linefeed)
fclose(svehhandle)
? "Done."


function errorbeep
        tone(300,1)
        tone(499,5)
        tone(700,5)
        return .t.

procedure helptext

? "Translates IDRISI vector files into SPANS vector archive format" ?
? "IDRISI files must be ascii, V2.0 format files with the extensions "
? ".vec and .dvc for the vector and vector file respectively "
? "Line and Polygon files are detected, and translated as SPANS"
? "Arc or Polygon files."
?
? "SPANS filename must not be the same as IDRISI filename"
?
? " - (if polygons) that they are contiguous and non-overlapping"
? " - (if arcs) that they are 'clean' with snapped endpoints"
? " - that all references are in projection coordinates."
?

procedure syntaxtext
      ? "Syntax:"
      ? "-------"
      ? "idr2spa infile, outfile"
      ?
      ? "where infile  = Idrisi Vector format file        (     No     )"
      ? "      outfile = filename for spans vector file   ( extensions )"
      ?

function error
```

```
        parameters errornum
        ? "ERROR: "
        do case
            case errornum=1
                ?? "Infile and Outfile cannot be the same."
            case errornum=2
                ?? "Do not specify file extensions, program assumes"
                ?? ".vec/.dvc (IDRISI) and .vec/.veh (SPANS)."
            case errornum=3
                ?? "Unable to read IDRISI .vec file."
            case errornum=4
                ?? "Unable to read IDRISI .dvc file."
            case errornum=5
                ?? "Unable to create SPANS .vec file."
            case errornum=6
                ?? "Unable to create SPANS .veh file."
            case errornum=7
                ?? "IDRISI Vector file must be version 2 format ASCII file."
            case errornum=8
                ?? "IDRISI file is type POINT - convert to a SPANS table."
            case errornum=9
                ?? "Only IDRISI V2.0 files supported at present."
            case errornum=10
                ?? ""
            case errornum=11
                ?? ""
            case errornum=12
                ?? ""
        endcase
         ? "IDR2SPA /h for help."
        ?
        quit

function freadline
        parameters filehandle
        filestring=""
        filechar=" "
        check=fread(filehandle,@filechar,1)
        if check=1
                do while (filechar=chr(13) .or. filechar=chr(10)) .and. check=1
                        check=fread(filehandle,@filechar,1)
                enddo
                do while (filechar<>chr(13) .and. filechar<>chr(10)) .and. check=1
                        filestring=filestring+filechar
                        check=fread(filehandle,@filechar,1)
                enddo
        endif
        return alltrim(filestring)

function fieldtitle
        parameters textline
        return alltrim(substr(textline,1,12))

function fieldvalue
        parameters textline
        textlength=len(textline)
        return alltrim(substr(textline,14,(textlength-13)))

function getfirst
        parameters instring
        instring=ltrim(instring)
        instring=substr(instring,1,at(" ",instring))
        return mathtran(instring)

function getsecond
        parameters instring
        instring=ltrim(instring)
        instring=substr(instring,at(" ",instring))
        return mathtran(instring)

function mathtran
        parameters mathstring
        stem=val(mathstring)
        plusminus=substr(mathstring,at("E",mathstring)+1,1)
        expvalue=substr(mathstring,at("E",mathstring)+2,2)
        if plusminus="+"
                outvalue= stem*( 10^(val(expvalue)) )
        elseif plusminus="-"
                outvalue= stem*( 10^(val(expvalue)-2*val(expvalue)) )
        else
                outvalue=stem
        endif
        return outvalue
```

Appendix II
GETCELL

II.i. Description

The GETCELL program was written to extract discrete 20km cells from a tape of NTF format data relating to many tiles. The program requires the following syntax:

```
getcell <infile> <outfile> <tile>
```

where <infile> is the name of the source file, <outfile> is the name of the file to create (both including extensions) and <tile> is the name of the tile to extract in letter-number form (e.g. SU04). The program is written in the Clipper dialect of the dBase language, and was compiled with the Summer '87 release of Clipper. An executable file which runs on all MS-DOS machines may be obtained via the internet by anonymous ftp, or gopher protocols from AVEBURY.ARCH.SOTON.AC.UK (give 'anonymous' as username and your e-mail address as the password). It is situated in the /pub/gis/utils directory as both source GETCELL.PRG and binary executable GETCELL.EXE.

II.ii. Code listing

```
* GETCELL.PRG
* -----------
* Removes a specified cell of OS Digitised Contour Data
* from a big file, and writes it to its own file.

* Call with: getcell filein, fileout, xx99
* where xx99 is an existing grid square ie. 'SU26'


* Obtain names of input and output files.
* ---------------------------------------
if pcount()=3
        parameters infile,outfile,cellname
else
        ? "Call with:"
        ? "getcell infile.ext, outfile.txt cellname"
        quit
endif

* Open input and output files.
* ----------------------------

inhandle=fopen(infile,0)
outhandle=fcreate(outfile,0)

dummy=" "
declare size[1]
adir(infile,dummy,size)
numrecs=int(size[1]/80)
? "   0%"

check=0
frecord=space(80)
check=fread(inhandle,@frecord,80)
processed=1
do while check=80 .and. inkey()<>27
        if substr(frecord,1,2)="07"
                if substr(frecord,3,4)=cellname
                        ?? " -Cell Found- "
                        ?? str((processed/numrecs)*100,3,0)+"%"
                        fwrite(outhandle,frecord)
                        check=fread(inhandle,@frecord,80)
                        processed=processed+1
                        do while (substr(frecord,1,2)<>"07") .and. inkey()<>27 .and.
check=80
```

```
                    fwrite(outhandle,frecord)
                    check=fread(inhandle,@frecord,80)
                    processed=processed+1
                    ?? chr(8)+chr(8)+chr(8)+chr(8)
                    ?? str((processed/numrecs)*100,3,0)+"%"
               enddo
               ? " End."
               quit
          endif
     endif
     ?? chr(8)+chr(8)+chr(8)+chr(8)
     ?? str((processed/numrecs)*100,3,0)+"%"
     check=fread(inhandle,@frecord,80)
     processed=processed+1
enddo
fclose(inhandle)
fclose(outfile)
```

Appendix III
NTF2IDR

III.i. Description

The NTF2IDR program was written to extract named line types from an NTF file and write an IDRISI GIS vector file as output. The program requires the following syntax:

```
ntf2idr <infile> <outfile> <code> [<A/P>]
```

where infile is the NTF file containing the source data (such as the output from GETCELL) including file extension, <OUTFILE> is the root filename for idrisi vector file (without the .vec extension), <CODE> is the ntf feature code for the lines to extract (i.e. 200 for spotheights, 201 for contours) and the optional <A/P> parameter indicates whether to output arcs (lines), or polygons. If this last parameter is omitted, the program promts for the information. The program outputs both the data (.vec) and header (.dvc) files, obtaining the names by adding the appropriate extension to the <OUTFILE> parameter.

III.ii. Code listing

```
* -------------------------------------------------
* ntf2spa.PRG
* -------------------------------------------------
* Reads a National Transfer Format (Level 1) vector
* file, searches for line data and writes IDRISI
* .vec (vector) file and .dvc (vector header) file.

*----------------------------------------
* MAIN PROGRAM
*----------------------------------------

set exact off
setcancel(.t.)

if pcount()=3
       parameters infile,outfile,feat_code
       filetype="z"
elseif pcount()=4
       parameters infile,outfile,feat_code,filetype
else
if pcount()=1
   parameters switch
   if switch="/h" .or. switch="/?" .or. switch="-h" .or. switch="-?"
? "Translates Ordnance Survey digital contour data (which uses"
? "a subset of the National Transfer Format level 1) into IDRISI"
? "vector file format (version 3.2 with a documentation file)"
? "All record pairs (23=line info. record, 21=coordinate records)"
? "with the code given as the third parameter (ie 201 = contours)"
? "are treated as lines in the new vector file, the height of the"
? "contour is output as the line attribute."
       endif
endif
? "Syntax:"
? "ntf2idr infile, outfile, code[, A/P]"
?
? "where infile  = ntf format file (with extension)"
? "outfile = filename for idrisi vector file (WITHOUT extension)"
? "code = ntf feature code (ie 200=spotheights, 201=contours etc)"
? "[A/P]    = optional filetype A=arcs (lines), P=polygons."
quit
endif

* Open input and output files & read 1st record.
* ----------------------------------------------
inhandle=fopen(infile,0)
vechandle=fcreate(alltrim(outfile)+".vec",0)
linefeed=chr(13)+chr(10)
```

```
check=1
frecord=space(80)
check=fread(inhandle,@frecord,80)
processed=1
numarcs=0

* Ask for arcs or polygons:
* -------------------------
filetype=upper(filetype)
if filetype<>"A" .and. filetype<>"P"
        set cursor on
        ?? "Do you want Arc or Polygon output? (A/P): "
        do while filetype<>"A" .and. filetype<>"P"
                filetype=chr(inkey(0))
                filetype=upper(filetype)
                if filetype<>"A" .and. filetype<>"P"
                        errorbeep()
                end
        enddo
        set cursor off
endif
if filetype="A"
        filetype="line"
else
        filetype="polygon"
endif

* Find a cell header record
*---------------------------
do while (substr(frecord,1,2)<>"07")
        check=fread(inhandle,@frecord,80)
        processed=processed+1
enddo
if .not. substr(frecord,1,2)="07"
        ? "Cell header record not found"
        quit
endif


* While not eof or file error
* ---------------------------
processed=0
dummy=""
declare size[1]
adir(infile,dummy,size)
numrecs=int(size[1]/80)
fcode=val(feat_code)
?
? "NTF2SPA. Processing "+upper(infile)+" : "+alltrim(str(size[1]/1000,5,0))+"k"
? alltrim(str(numrecs))+" records."
? "Feature code "+alltrim(str(fcode))

* Get base coords. from cell Header Record:
* ----------------------------------------
basex=substr(frecord,44,9)
basey=substr(frecord,56,9)
minx=val(basex)
miny=val(basey)
? "Base is : "+basex+", "+basey
? "Done:   0%"
set cursor off
do while check<>0 .and. inkey()<>27

        * find a line record with appropriate feature code:
        * ------------------------------------------------
        do while substr(frecord,1,2)<>"23".and. check=80
                check=fread(inhandle,@frecord,80)
                processed=processed+1
        enddo
        if check=80 .and. substr(frecord,17,19)=feat_code
                numarcs=numarcs+1
                height=val(substr(frecord,35,6))
                if height<1
                        height=-1
                endif
                heightstr=alltrim(str(height))

                * Read the first geometry record:
                * -------------------------------
                check=fread(inhandle,@frecord,80)
                processed=processed+1
                numstring=substr(frecord,30,4)

                * Write the arc value and number of points:
                * ---------------------------------------------------
                fwrite(vechandle,heightstr+linefeed)
```

```
                fwrite(vechandle,numstring+linefeed)
                fwrite(vechandle,str(val(subst(frecord,37,9))+minx)+linefeed)
        fwrite(vechandle,str(val(subst(frecord,47,9))+miny)+linefeed)
        fwrite(vechandle,str(val(subst(frecord,58,9))+minx)+linefeed)
        fwrite(vechandle,str(val(subst(frecord,68,9))+miny)+linefeed)

                * Read continuation records:
                * --------------------------
                check=fread(inhandle,@frecord,80)
                processed=processed+1
                do while substr(frecord,1,2)="00" .and. check=80
                        tempstring = substr(frecord,7,9)
                        fwrite(vechandle,str(val(tempstring)+minx)+linefeed)
                        tempstring = substr(frecord,17,9)
                        fwrite(vechandle,str(val(tempstring)+miny)+linefeed)
                        tempstring = substr(frecord,28,9)
                        if tempstring<>space(9)
                                fwrite(vechandle,str(val(tempstring)+minx)+linefeed)
                        endif
                        tempstring = substr(frecord,38,9)
                        if tempstring<>space(9)
                                fwrite(vechandle,str(val(tempstring)+miny)+linefeed)
                        endif
                        tempstring = substr(frecord,49,9)
                        if tempstring<>space(9)
                                fwrite(vechandle,str(val(tempstring)+minx)+linefeed)
                        endif
                        tempstring = substr(frecord,58,9)
                        if tempstring<>space(9)
                                fwrite(vechandle,str(val(tempstring)+miny)+linefeed)
                        endif
                        check=fread(inhandle,@frecord,80)
                        processed=processed+1
                        ?? chr(8)+chr(8)+chr(8)+chr(8)
                        ?? str((processed/numrecs)*100,3,0)+"%"
                enddo
        else
                check=fread(inhandle,@frecord,80)
                processed=processed+1
                ?? chr(8)+chr(8)+chr(8)+chr(8)
                ?? str((processed/numrecs)*100,3,0)+"%"
        endif

enddo
set cursor on
? alltrim(str(numarcs))+" arcs written."
?
fclose(infile)
fwrite(vechandle,"0"+linefeed+"0"+linefeed)
fclose(outfile+".vec")
dvchandle=fcreate(alltrim(outfile)+".dvc",0)

fwrite(dvchandle,"title        : "+linefeed)
fwrite(dvchandle,"data type    : integer"+linefeed)
fwrite(dvchandle,"file type    : ascii"+linefeed)
fwrite(dvchandle,"object type  : "+filetype+linefeed)
fwrite(dvchandle,"coord. span  :   1.0000000000E+01"+linefeed)
fwrite(dvchandle,"coord. unit  : m"+linefeed)
fwrite(dvchandle,"min X        :    "+str(minx)+linefeed)
fwrite(dvchandle,"max X        :    "+str(minx+20000)+linefeed)
fwrite(dvchandle,"min Y        :    "+str(miny)+linefeed)
fwrite(dvchandle,"max Y        :    "+str(miny+20000)+linefeed)

fclose(outfile+".dvc")


function errorbeep
        tone(300,1)
        tone(499,5)
        tone(700,5)
        return .t.
```

Appendix IV
NTF2SPA

IV.i. Description

The NTF2SPA program is was written to extract named line types from an NTF file and write an SPANS GIS vector archive format file as output. The program requires the following syntax:

```
ntf2spa <infile> <outfile> <code> [<A/P>]
```

where infile is the NTF file containing the source data (such as the output from GETCELL) including file extension, <OUTFILE> is the root filename for spans vector file (without the .vec and .veh extensions), <CODE> is the ntf feature code for the lines to extract (i.e. 200 for spotheights, 201 for contours) and the optional <A/P> parameter indicates whether to output arcs (lines), or polygons. If this last parameter is omitted, the program promts for the information. The program outputs both the data (.veh) and header (.veh) archive files, obtaining the names by adding the appropriate extension to the <OUTFILE> parameter. These may then be copied to the SPANS directory and imported into SPANS as with all other vector archive files.

IV.ii. Code listing

```
* -----------------------------------------------
* ntf2spa.PRG
* -----------------------------------------------
* Reads a National Transfer Format (Level 1) vector
* file, searches for line data and writes SPANS
* .vec (vector) file and .veh (vector header) file.

*----------------------------------------
* MAIN PROGRAM
*----------------------------------------

set exact off
setcancel(.t.)

if pcount()=3
        parameters infile,outfile,feat_code
        filetype="z"
elseif pcount()=4
        parameters infile,outfile,feat_code,filetype
else
   if pcount()=1
     parameters switch
     if switch="/h" .or. switch="/?" .or. switch="-h" .or. switch="-?"
? "Translates Ordnance Survey digital contour data (which uses"
? "a subset of the National Transfer Format level 1) into SPANS"
? "vector file format (version 4.02 with a documentation file)"
? "All record pairs (23=line info. record, 21=coordinate records)"
? "with the code given as the third parameter (ie 201 = contours)"
? "are treated as lines in the new vector file, the height of the"
? "contour is output as the line attribute."
      endif
   endif
? "Syntax:"
? "ntf2spa infile, outfile, code[, A/P]"
?
? "where infile  = ntf format file (with extension)"
? "      outfile = filename for spans vector file (WITHOUT extension)"
? "      code    = ntf feature code"
quit
endif

* Open input and output files & read 1st record.
* ----------------------------------------------
inhandle=fopen(infile,0)
vechandle=fcreate(alltrim(outfile)+".vec",0)
```

```
linefeed=chr(13)+chr(10)
check=1
frecord=space(80)
check=fread(inhandle,@frecord,80)
processed=1
numarcs=0

* Ask for arcs or polygons:
* -------------------------
filetype=upper(filetype)
if filetype<>"A" .and. filetype<>"P"
        set cursor on
        ?? "Do you want Arc or Polygon output? (A/P): "
        do while filetype<>"A" .and. filetype<>"P"
                filetype=chr(inkey(0))
                filetype=upper(filetype)
                if filetype<>"A" .and. filetype<>"P"
                        errorbeep()
                end
        enddo
        set cursor off
endif

* Find a cell header record
*---------------------------
do while (substr(frecord,1,2)<>"07")
        check=fread(inhandle,@frecord,80)
        processed=processed+1
enddo
if .not. substr(frecord,1,2)="07"
        ? "Cell header record not found"
        quit
endif


* While not eof or file error
* ---------------------------
processed=0
dummy=""
declare size[1]
adir(infile,dummy,size)
numrecs=int(size[1]/80)
fcode=val(feat_code)
?
? "NTF2SPA. Processing "+upper(infile)+" : "+alltrim(str(size[1]/1000,5,0))+"k"
? alltrim(str(numrecs))+" records."
? "Feature code "+alltrim(str(fcode))

* Get base coords. from cell Header Record:
* -----------------------------------------
basex=substr(frecord,44,9)
basey=substr(frecord,56,9)
minx=val(basex)
miny=val(basey)
? "Base is : "+basex+", "+basey
? "Done:   0%"
set cursor off
fwrite(vechandle,"ARCS"+linefeed)
do while check<>0 .and. inkey()<>27

        * find a line record with appropriate feature code:
        * -------------------------------------------------
        do while substr(frecord,1,2)<>"23".and. check=80
                check=fread(inhandle,@frecord,80)
                processed=processed+1
        enddo
        if check=80 .and. substr(frecord,17,19)=feat_code
                numarcs=numarcs+1
                height=val(substr(frecord,35,6))
                if height<1
                        height=-1
                endif
                heightstr=alltrim(str(height))

                * Read the first geometry record:
                * -------------------------------
                check=fread(inhandle,@frecord,80)
                processed=processed+1
                numstring=substr(frecord,30,4)

       * Write arc identifier, number of points and value:
       * -------------------------------------------------
       fwrite(vechandle,alltrim(str(numarcs))+" ")
       fwrite(vechandle,numstring+" ")
       fwrite(vechandle,heightstr+linefeed)
       fwrite(vechandle,str(val(subst(frecord,37,9))+minx)+" ")
```

```
      fwrite(vechandle,str(val(subst(frecord,47,9))+miny)+linefeed)
      fwrite(vechandle,str(val(subst(frecord,58,9))+minx)+" ")
     fwrite(vechandle,str(val(subst(frecord,68,9))+miny)+linefeed)

              * Read continuation records:
              * -------------------------
              check=fread(inhandle,@frecord,80)
              processed=processed+1
              do while substr(frecord,1,2)="00" .and. check=80
                      tempstring = substr(frecord,7,9)
           fwrite(vechandle,str(val(tempstring)+minx)+" ")
                      tempstring = substr(frecord,17,9)
             fwrite(vechandle,str(val(tempstring)+miny)+linefeed)
                      tempstring = substr(frecord,28,9)
                      if tempstring<>space(9)
           fwrite(vechandle,str(val(tempstring)+minx)+" ")
                      endif
                      tempstring = substr(frecord,38,9)
                      if tempstring<>space(9)
           fwrite(vechandle,str(val(tempstring)+miny)+linefeed)
                      endif
                      tempstring = substr(frecord,49,9)
                      if tempstring<>space(9)
           fwrite(vechandle,str(val(tempstring)+minx)+" ")
                      endif
                      tempstring = substr(frecord,58,9)
                      if tempstring<>space(9)
           fwrite(vechandle,str(val(tempstring)+miny)+linefeed)
                      endif
                      check=fread(inhandle,@frecord,80)
                      processed=processed+1
                      ?? chr(8)+chr(8)+chr(8)+chr(8)
                      ?? str((processed/numrecs)*100,3,0)+"%"
              enddo
      else
              check=fread(inhandle,@frecord,80)
              processed=processed+1
              ?? chr(8)+chr(8)+chr(8)+chr(8)
              ?? str((processed/numrecs)*100,3,0)+"%"
      endif

enddo
set cursor on
? alltrim(str(numarcs))+" arcs written."
?
fclose(infile)
fclose(outfile+".vec")
vehhandle=fcreate(alltrim(outfile)+".veh",0)
fwrite(vehhandle,"FILE_TYPE "+filetype+" X C 0 1"+linefeed)
fwrite(vehhandle,"VERSION 4.02"+linefeed)
arcsheader= alltrim(str(int(numarcs)))+" 3 "+str(minx)+" "+str(miny)+
              " "+str(minx+20000)+" "+str(miny+20000)+" 3 3 0 0 0 0 0"
fwrite(vehhandle,"ARCS_HEADER "+arcsheader+linefeed)
fwrite(vehhandle,"EXTENTS "++str(minx)+" "+str(miny)
              +" "+str(minx+20000)+" "+str(miny+20000)+linefeed)
fwrite(vehhandle,"REFERENCE 1 1 3 1 1 0"+linefeed)
fclose(outfile+".veh")


function errorbeep
      tone(300,1)
      tone(499,5)
      tone(700,5)
      return .t.
```

Appendix V
IDROUT.LSP

V.i. Description

IDROUT is an AutoLISP program to extend the capabilities of AutoCAD to export IDRISI format vector files from the AutoCAD command line. The program must be located in a directory which is on the AutoCAD search path, and then loaded into the AutoCAD program with the following syntax:

```
Command:> (load "idrout")
IDROUT v1.2 (c) Dave Wheatley, Jan 1994
```

The program allows point, line or area (whole polygon) data to be exported to an IDRISI vector file. The dialogue allows the selection of the output type, filename, accuracty of the coordinates and accuracy of the attribute values. The program also allows three methods for designating the attribute values of the entities to be output (a) as fixed values (b) using the z-coordinate of the AutoCAD entities or (c) using the layer of the AutoCAD entities. The program does not produce the documentation file and must be documented within IDRISI prior to display or analyisis. The syntax for using the program is extremely similar to the internal AutoCAD DXFOUT function on which it is modelled. A typical use of the program for export of line data to the file 'outfile.vec' is as follows:

```
Command: idrout
Points lines or areas (p/l/a)?: l
Filename for output?: outfile.vec
Enter decimal places for coordinates <6>: 2
Enter decimals for attributes (0 for integer) <0>: 0
Attribute method [Fixed, Layer or Z]?: f
Attribute value <0>: 10

Select objects: w

First corner: Other corner: 1 found

Select objects:
Working . done.nil
```

The only known problem with the program is that it will crash if it is asked to write to a write-protected disc drive. It is therefore important that if the current directory is write-protected (such as a network drive) then the *full pathname* of the output file be specified on the command line.

The lisp file IDROUT.LSP which runs on all versions of AutoCAD for UNIX and MS-DOS machines (so far as it has been possible to test it) may be obtained via the internet by anonymous ftp, or gopher protocols from AVEBURY.ARCH.SOTON.AC.UK (give 'anonymous' as username and your e-mail address as the password). It is situated in the /pub/autocad/autolisp directory as IDROUT.LSP.

V.ii. Code listing

```
; IDROUT
; Generic export function from AutoCAD drawings to IDRISI format
; vector files. Modelled on the AutoCAD DXFOUT function

; Allows output of Points, Lines and Areas.
; Checks that all Area output is composed of closed polygons.
; Allows the user to specify the decimal accuracy of coordinate values.
; Allows attribute values to be either
;       (a) Specified as a fixed value for the whole idrisi file
;       (b) Obtained from the z-coordinates of the autocad entities
;       (c) Obtained from the layer name of the entities
; Allows the user to specify the decimal accuracy of the attribute values

; ----------------------
; FUNCTIONS
; ----------------------

; DXF - Returns the data part of the dotted pair
; ----------------------------------------------
(defun dxf (dxfcode entitydata)
    (cdr (assoc dxfcode entitydata))
);dxf


; GETMETHOD - Returns "Z", "L" or an integer as a string
; -----------------------------------------------------
(defun getmethod (/ zvaltype)
   (while (not (or  (= zvaltype "Z") (= zvaltype "F") (= zvaltype "L") ))
        (setq zvaltype (strcase (getstring "Attribute method [Fixed, Layer or Z]?: ")) )
    ) ; while
    (if (= zvaltype "F")
        (progn (setq zvaltype 0)
               (setq zvaltype (rtos (getreal "Attribute value <0>: ") 2 numattdecimals))
        )
        (setq zvaltype zvaltype)
    );if
);getmethod


; writeidrpoly - Write a polyline to the file fn
; ------------------------------------------------
; call with (writeidrpoly <name_of_head_entity> <L, Z or integer> <decimals> <A(rc) or
L(ine)> <Filehandle>)
(defun writeidrpoly (headname zval outputtype fn  / vname vdata xcoord ycoord zcoord)

         (setq headdata (entget headname))
         (setq vname (entnext headname))


         ; Get attribute value for pline
         ; -----------------------------
         (if (= zval "Z")
             (setq zval (rtos (caddr (dxf 10 headdata)) 2 numattdecimals))
         );if
         (if (= zval "L")
             ;(if (= (type (dxf 8 headdata)) "INT")
             ;    (setq zval (rtos (dxf 8 headdata) 2 numattdecimals))
             ;    (setq zval "0")
             ;)
             (setq zval (dxf 8 headdata))
         );if

         ; Count the vertices:
         ; -------------------
         (setq tempname vname)
         (setq tempdata (entget tempname))
         (setq nverts 0)
         (while (= (dxf 0 tempdata) "VERTEX")
             (setq nverts (+ nverts 1))
             (setq tempname (entnext tempname))
             (setq tempdata (entget tempname))
          ); end while
          (if (= outputtype "A")
             (setq nverts (+ nverts 1))
          );
          (setq nverts (rtos nverts 2 0))
```

```
         ; Write the polyline header:
         ; --------------------------
         (princ (strcat zval " " nverts "\n") fn)

         ; Step throuh the vertices:
         ; -------------------------
         (setq vdata (entget vname))
         (setq xcoord1 (car (dxf 10 vdata)))
         (setq ycoord1 (cadr (dxf 10 vdata)))
         (while (= (dxf 0 vdata) "VERTEX")
                (setq xcoord (rtos (car (dxf 10 vdata)) 2 numdecimals))
                (setq ycoord (rtos (cadr (dxf 10 vdata)) 2 numdecimals))
                (princ (strcat xcoord " " ycoord "\n") fn)
                (setq vname (entnext vname))
                (setq vdata (entget vname))
          ); end while

          ; Repeat the first coordinates to ensure closed polygons.
          ; -------------------------------------------------------
          (if (= outputtype "A")
(princ (strcat (rtos xcoord1 2 numdecimals) " " (rtos ycoord1 2 numdecimals) "\n") fn)
          );if
);writepoly


; ----------------------
; MAIN PROGRAM
; ----------------------

(defun C:idrout()

   ; Get output type (point, line or area):
   ;-----------------------------------
   (setq outputtype "")
   (while (not (or  (= outputtype "L") (= outputtype "P") (= outputtype "A")))
       (setq outputtype (strcase (getstring "Points lines or areas (p/l/a)?: ")) )
   ) ; while

   ; Get output filename and open file:
   ;---------------------------------
   (setq fname (getstring "Filename for output?: "))
   (setq fn (open fname "w"))

   ; Get number of decimal places
   ;----------------------------
   (setq numdecimals (getint "Enter decimal places for coordinates <6>: "))
   (if (= numdecimals nil) (setq numdecimals 6))
   (setq numattdecimals (getint "Enter decimals for attributes (0 for integer) <0>: "))
   (if (= numattdecimals nil) (setq numattdecimals 0))

   ; For LINE or AREA data:
   ; ========================================
   (if (or (= outputtype "L") (= outputtype "A"))
        (progn

            ; Get method for obtaining line attribute:
            ;----------------------------------------
            (setq zvaltype (getmethod))


            ; Get set of entities:
            ; ---------------------
            (setq sel_set (ssget))

            ; Step through the selection set:
            ;---------------------------------
            (princ "Working ")
            (repeat (sslength sel_set)
               (princ ".")
               (setq ename (ssname sel_set 0))
               (setq ent1 (entget ename))
               (if (= (dxf 0 ent1) "POLYLINE")
                   (writeidrpoly ename zvaltype outputtype fn)
                );endif
                (ssdel ename sel_set)
             ) ;end repeat loop
            (princ "0 0\n" fn)
            (princ " done.")
        );progn

    ; Else for point data:
    ; ========================================
    (progn

        ; Get method for obtaining point attribute:
        ;-----------------------------------------
```

```
        (setq zvaltype (getmethod))
        (setq pvalue zvaltype)

        ; Get set of entities:
        ; ----------------------
        (setq sel_set (ssget))

       ; Step throuh the selection set:
       ;--------------------------------
       (princ "Working ")
       (repeat (sslength sel_set)
            (progn
               ; Get a entity from set:
               ; -----------------------
               (setq ename (ssname sel_set 0))
               (princ ".")
               (setq ent1 (entget ename))
               (if (= (dxf 0 ent1) "POINT")
                   (progn

                       ; Get z value of pline
                       ; -------------------
                       (if (= zvaltype "Z")
                         (progn
                             (setq pvalue (rtos (caddr (dxf 10 ent1)) 2 numattdecimals))
                         );progn
                       );if
                       (if (= zvaltype "L")
                            (setq pvalue (dxf 8 ent1))
                       );if

                       (princ (strcat pvalue " 1\n") fn)

                       ; Get coordinates of point:
                       ; -------------------------
                       (setq xcoord (rtos (car (dxf 10 ent1)) 2 numdecimals) )
                       (setq ycoord (rtos (cadr (dxf 10 ent1)) 2 numdecimals) )

                       (setq outp (strcat xcoord " " ycoord "\n"))
                       (princ outp fn)
                   );progn
               );endif
               (ssdel ename sel_set)
            ) ;progn
        ) ;end repeat loop
        (princ "0 0\n" fn)
        (princ " done.")
    );progn
    );if

(close fn)
); end idrout()
(princ "\nIDROUT v1.2   (c) Dave Wheatley, Jan 1994")
(princ)
```

Appendix VI
DIGITOOL.LSP

VI.i. Description

DIGITOOL.LSP contains five utility programs which can be a great assistance for digitising within the AutoCAD program for export to a gis. It must be loaded into the AutoCAD program in exactly the same way as IDROUT.LSP (see appendix V) prior to use. The five utilities are as follows:

- PLINK allows the selection of any number of polyline entities, then the selection of a point. The utility then moves the closest endpoint of all the polylines to that point. This therefore removes the possibility that the lines do not meet correctly at a node.

- WEED reduces vertex density of a drawing by selectively discarding vertices of polylines (except the endpoints). The program asks for a threshold, in drawing units, and then steps through all selected polylines discarding all vertices which are less than the specified minimum distance from the last recorded vertex.

- CLOSEALL converts all open polylines in a drawing to closed polylines. This is essential prior to export of polyline information as closed polygon areas to a GIS.

- LN2PLINE converts all line entities within a drawing to two-vertex polylines. This can be used to convert drawings which have been supplied as a series of 'line' entities to polylines (which are required by the IDROUT program. First, this program should be used to convert all lines to polylines, then the (j)oin function of the pedit command can be used to join the lines into multi-vertex polylines.

- VTXSNAP relocates all vertices of a polyline to a coordinate 'round number'. It is equivalent to a retrospective use of the 'snap' function.

The lisp file DIGITOOL.LSP which runs on all versions of AutoCAD for UNIX and MS-DOS machines (so far as it has been possible to test it) may be obtained via the internet by anonymous ftp, or gopher protocols from AVEBURY.ARCH.SOTON.AC.UK (give 'anonymous' as username and your e-mail address as the password). It is situated in the /pub/autocad/autolisp directory as DIGITOOL.LSP.

VI.ii. Code listing

```
;================================================
; DIGITOOLS v1.0 Digitizing tools for AutoCAD
;
; (c) 1993, 1994 David Wheatley, Department of
;     Archaeology, University of Southampton,
;     Southampton,England, UK SO9 5NH
```

```
;
; PLINK, WEED, CLOSEALL, and LN2PLINE
;================================================

;================================
; Subroutines used by the tools
;================================

; Rounds off <vreal> to the nearest <vround>
; ----------------------------------------
(defun roundoff(vreal vround)
       (* vround (atoi (rtos (/ vreal vround) 2 0)))
); roundoff


; Moves start point of polyline to a new point
; -------------------------------------------
(defun move_start(entname nodepoint / x1 y1 z1 new10list)
       (setq   edata (entget ename)
               vname (entnext ename)
               vdata (entget vname)
               x1 (car nodepoint)
               y1 (cadr nodepoint)
               z1 (caddr nodepoint)
       )
       (setq new10list (list 10 x1 y1 z1))
       (entmod (subst new10list (assoc 10 vdata) vdata))
); move_start

; Moves last point of polyline to a new point
; -------------------------------------------
(defun move_end(entname lastvtxname nodepoint / x1 y2 z1 new10list)
       (setq   edata (entget ename)
               vdata (entget lastvtxname)
               x1 (car nodepoint)
               y1 (cadr nodepoint)
               z1 (caddr nodepoint)
       )
       (setq new10list (list 10 x1 y1 z1))
       (entmod (subst new10list (assoc 10 vdata) vdata))
); move_end

;-------------------------------------------
; PLINK - aligns the endpoints of polylines
; -----------------------------------------

; Allows selection of any number of plines, then a user-entered
; node point, then moves the nearest endpoint of each pline to
; the given node point.


(defun C:plink()

       ; Get set of polylines
       ; --------------------
       (setq s_set (ssget))

       ; Get point to snap endpoints to
       ; ------------------------------
       (setq endnode (getpoint "Place node"))

       ; Step through the selection set:
       ;--------------------------------
       (repeat (sslength s_set)

               ; Get head and first vertex:
               ; ---------------------------
               (setq ename (ssname s_set 0)
                     edata (entget ename)
                     vname (entnext ename)
                     vdata (entget vname)   )

               ; Get point of first vertex
               ; --------------------------
               (setq vtx1 (cdr (assoc 10 vdata)) )

               ; Step throuh the vertices:
               ; ------------------------
               (while (= (cdr (assoc 0 vdata)) "VERTEX")

                       ; Get point
                       ; ----------
                       (setq vtx2 (cdr (assoc 10 vdata)) )

                       ; Move to next vertex
                       ; ------------------
```

```
                        (setq   lastvname vname
                                vname (entnext vname)
                                vdata (entget vname)    )

                ); end while

                ; Move endpoint or startpoint to node
                ; ------------------------------------
                (if (< (distance vtx1 endnode) (distance vtx2 endnode))
                        (move_start ename endnode)
                        (move_end ename lastvname endnode)
                );endif
                (entupd ename)
                (ssdel ename s_set)
        );repeat
        (princ)
);plink

; --------------------------------------
; WEED - weeds out vertices from plines
; --------------------------------------
; Discards all points which are below the threshold distance
; away from the last recorded point (distance in *either* x
; or y axes).

(defun C:weed()

    ; Get threshold value for weeding:
    ;----------------------------------
    (setq thresh (getreal "Threshold?: "))

    ; Get set of polylines:
    ; ----------------------
     (setq ss_layer (ssget))

    ; Step through the selection set:
    ;-----------------------------------
    (princ "Working ")
    (repeat (sslength ss_layer)

        ; Get head entity and first vertex:
        ; --------------------------------
        (princ ".")
        (setq ename (ssname ss_layer 0)
               edata (entget ename)
               vname (entnext ename)
               vdata (entget vname)   )

        ; Make new head and first vertex
        ; ------------------------------
        (entmake edata)
        (entmake vdata)

        ; Get x and y of first vertex
        ; ---------------------------
        (setq x1 (car (cdr (assoc 10 vdata)) ))
        (setq y1 (cadr (cdr (assoc 10 vdata)) ))

        ; Move to second vertex
        ; ---------------------
        (setq vname (entnext vname))
        (setq vdata (entget vname))

        ; Step throuh the vertices:
        ; -------------------------
        (while (not (= (cdr (assoc 0 vdata)) "SEQEND"))

                ; Get x and y
                ; -----------
                (setq x2 (car (cdr (assoc 10 vdata)) ))
                (setq y2 (cadr (cdr (assoc 10 vdata)) ))

                ; Get distance from last written vertex
                ; -------------------------------------
                (setq dist (max (abs (- x2 x1)) (abs (- y2 y1)) ))

                ; If far enough away, make new vertex
                ; ----------------------------------
                (if (> dist thresh)
                   (progn
                        (entmake vdata)
                        (setq x1 x2)
                        (setq y1 y2)
                   );progn
                );if
```

```
                ; Move to next vertex
                ; -------------------
                (setq lastvdata vdata)
                (setq vname (entnext vname))
                (setq vdata (entget vname))
        ); end while

        ; Make tail entity
        ; ----------------
        (entmake lastvdata)
        (entmake vdata)

        ; Delete original entity
        ; ----------------------
        (entdel ename)

        ; Redraw new entity
        ; -----------------
        (redraw (entlast))

        (ssdel ename ss_layer)
    ) ;end repeat loop
    (princ)
); end weed()


; ----------------------------------------
; CLOSEALL - Closes all selected polylines
; ----------------------------------------

(defun C:closeall()

    ; Initialize counter
    ; -------------------
    (setq numnotplines 0)

    ; Get set of entities:
    ; ----------------------
     (setq ss_layer (ssget))

    ; Step through the selection set:
    ;----------------------------------
    (princ "Working ·")
    (repeat (sslength ss_layer)
        (progn

                ; Get entity
                ; -----------
                (setq ename (ssname ss_layer 0)
                      edata (entget ename)       )

                ; If entity is a polyline, get vertex and continue
                ; ------------------------------------------------
                (if (= "POLYLINE" (cdr (assoc 0 edata)))
                   (progn
                        (setq vname (entnext ename)
                              vdata (entget vname)   )

                        (entmake edata)   ;} Make new head and
                        (entmake vdata)   ;} first vertex

                        (setq x1 (car (cdr (assoc 10 vdata)) ))    ;} Remember first x
                        (setq y1 (cadr (cdr (assoc 10 vdata)) ))   ;} and y of polyline
                        (setq firstvdata vdata)

                       ; Step throuh the vertices:
                        (while (not (= (cdr (assoc 0 vdata)) "SEQEND"))
                                  (entmake vdata)             ; Make new vertex then ...
                                 (setq lastvdata vdata)       ; ... move to next vertex
                                 (setq vname (entnext vname))
                                 (setq vdata (entget vname))
                        ); end while

                        (setq x2 (car (cdr (assoc 10 lastvdata)) ))   ;} Get the x and y
                        (setq y2 (cadr (cdr (assoc 10 lastvdata)) ))  ;} of the end vert

                       ; If first and last vertices are not the same ...
                        (if (not (and (= x1 x2) (= y1 y2)))
                              (entmake firstvdata)             ; ... make a new vertex
                        );endif

                        (entmake vdata)       ; Make new seqend (tail) entity
                        (entdel ename)        ; Delete original entity
                        (redraw (entlast))    ; Redraw new entity
                        (princ "·")
                );progn
```

```
                ; else it's not a polyline ...
                (setq numnotplines (+ numnotplines 1))    ; ... so count it
        );endif

      ; Remove ename from selection set
      ; ------------------------------
      (ssdel ename ss_layer)

   )); end progn and repeat

   (if (not (= numnotplines 0)) (progn
       (princ "· ")
       (princ numnotplines)
       (princ " entities were ignored because they were not polylines.")
   ));progn and if

   (princ)
); end closeall()


; VTXSNAP - Snaps all pline vertices to given accuracy
; ---------------------------------------------------

; Moves all vertices of every pline to the nearest <thresh>.

(defun C:vtxsnap()

    ; Get layer name and selection set:
    ;----------------------------------
    (setq vlayer (getstring "What layer?: "))

    ; Get threshold value for weeding:
    ;----------------------------------
    (setq thresh (getint "Threshold?: "))
    (graphscr)

    ; Get set of polylines:
    ; ----------------------
    (setq ss_layer (ssget "X" (list (cons 8 vlayer) (cons 0 "POLYLINE"))))

    ; Step through the selection set:
    ;----------------------------------
    (princ "Working ")
    (repeat (sslength ss_layer)

        ; Get first vertex:
        ; -----------------
        (setq   ename (ssname ss_layer 0)
                vname (entnext ename)
                vdata (entget vname)     )

        ; Step throuh the vertices:
        ; ------------------------
        (while (not (= (cdr (assoc 0 vdata)) "SEQEND"))

                ; Change vertex
                ; -------------
                (setq x1 (car (cdr (assoc 10 vdata)) ))
                (setq y1 (cadr (cdr (assoc 10 vdata)) ))
                (setq z1 (caddr (cdr (assoc 10 vdata)) ))
                (setq new10list (list 10 (roundoff x1 thresh) (roundoff y1 thresh)
(roundoff z1 thresh)))
                (entmod (subst new10list (assoc 10 vdata) vdata))

                ; Move to next vertex
                ; ------------------
                (setq vname (entnext vname))
                (setq vdata (entget vname))
        ); end while
        (entupd ename)
        (ssdel ename ss_layer)

   ) ;end repeat loop
   (princ)
); end vtxsnap()


; ---------------------------------------------
; Selects all lines within specified layer, then
; converts them into plines.
; ---------------------------------------------
(defun C:ln2pline()

   ; Get layer name and selection set:
   ;----------------------------------
   (setq vlayer (getstring "What layer?: "))
```

```
    (setq ss_layer (ssget "X" (list (cons 8 vlayer) (cons 0 "LINE"))))

    ; Step through the selection set:
    ;---------------------------------
    (princ "Working ")
    (repeat (sslength ss_layer)
         (princ ".")
         (setq ename (ssname ss_layer 0))
         (setq edata (entget ename))
         (setq vtx1 (cdr (assoc 10 edata)) )
         (command "pedit" vtx1 "y")
         (ssdel ename ss_layer)
    ) ;end repeat loop
); end ln2pline()
```

Appendix VII
Batch processing of visibility calculations

VII.i. Description

The repetitive nature of procedure to generate cumulative viewshed maps may be reduced by using a batch processing strategy to call the individual parts of the process for each barrow automatically. This batch process can also be instructed to keep a running total, and therefore to remove the unwanted intermediate steps as the analysis proceeds. The analysis as a whole then occupies substantially less disk space, and may be left to run overnight; the resulting cumulative viewshed surface examined the following day.

The following pair of batch files were written for the MS-DOS operating system, to run the IDRISI GIS. They generate a cumulative viewshed map from a series of vector files, each of which contains a single point entity representing a long barrow. The files assume the vector files are labelled *bar_1.vec ... bar_n.vec*, and that there is an empty raster image called *blank250.img* (which is copied for each barrow). The elevation model is contained in the file *dem250.img*, and all documentation files must also be present and correct.

VII.ii. Code listing

CVA.BAT is the main control file which contains the list of the point files to process, and calls the viewshed routine for each of these:

```
@echo off
rem == CALCULATES CUMULATIVE, VIEWSHED MAP for
rem == the group of sites given in the (brackets)

rem == Create the running total file:
copy blank250.* runtot.*

rem == Do the viewshed calculation for each barrow:
for %a in
(1,2,3,4,5,6,7,8,9,10,11,12,14,15,16,17,18,19,20,22,23,24,25,26,27,28,29,30,31,32,33) do
call view.bat %a

rem == Rename the running total:
ren runtot.* viewsum.*
```

and VIEW.BAT contains the line-of-sight procedure call, adds the result to the running total and then deletes the intermediate files:

```
@echo off
rem == CALLED FROM CVA.BAT
rem == Obtains a visual impact map for one monument, then adds
rem == it to the running total of visual impact.
rem == Needs (1) point vector files for each monument
rem ==       (2) viewshed maps for each monument (two classes)

rem == Rasterize barrow location:
copy blank250.* bar_%1.*
pointras x bar_%1 bar_%1 3

rem == Do viewshed calculation:
viewshed x dem250 bar_%1 18000 2 bview_%1
```

```
rem == Add result to running total:
overlay 1 bview_%1 runtot temp
del runtot.*
ren temp.* runtot.*

rem == Delete image files for this point:
del *_%1.*
```

Allen KMS, 1990 Modelling early historic trade in the eastern Great Lakes using geographic information systems IN Allen KMS, Green S & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology*. London: Taylor and Francis. pp.319-329.

Allen KMS, Green S & Zubrow EBW (eds) 1990 Interpreting Space: GIS and Archaeology. London: Taylor and Francis.

Allen MJ, 1988 Archaeological and Environmental Aspects of Colluviation in South-East England. IN Groenman van Waateringe W & Robinson M (eds) *Man Made Soils*. Oxford: British Archaeological Reports International Series 410 pp.67-92.

Altschul JH, 1988 Models and the modeling process. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past* Denver: US Department of the Interior Bureau of Land Management. pp.61-96.

Altschul JH, 1990 Red Flag Models: the Use of Modelling in Management Contexts. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology*. London: Taylor & Francis. pp.226-238.

Anderson JM, 1981 Ecology for Environmental Sciences. London: Edward Arnold.

Arnold III JB, 1979 Archaeological Applications of Computer Drawn Contour and Three-Dimensional Perspective Plots. IN Upham S (ed) *Computer graphics in archaeology : statistical cartographic applications to spatial analysis in archaeological contexts*. Arizona State University Anthropological Research Papers no 15. pp.1-15.

Ashbee P, 1963 The Wilsford Shaft. *Antiquity* XXXVII pp.116-120.

Ashbee P, 1966 The Fussell's Lodge Long Barrow Excavations. *Archaeologia* 100, pp.1-80.

Ashbee P, 1984 The Earthern Long Barrow in Britain. (Second edition: 1st ed. 1970) Cambridge: Cambridge University Press.

Ashbee P, Smith IF, & Evans JG, 1979 Excavation of Three Long Barrows near Avebury, Wiltshire. *Proceedings of the Prehistoric Society* 45 pp.207-300.

Barker CT, 1985 The Long Mounds of the Avebury Region. *Wiltshire Archaeological and Natural History Magazine* 79. pp.7-38.

Barker G, 1985 Prehistoric Farming in Europe. Cambridge: Cambridge University Press.

Barker G & Webley D, 1978 Causewayed Camps and Early Neolithic Economies in Central Southern England. *Proceedings of the Prehistoric Society* 44 pp.161-186.

Barrett JC, 1987 Fields of discourse. Reconstructing a Social Archaeology. *Critique of Anthropology* 7(3). pp.5-17.

Barrett JC, 1988 The living, the dead and the ancestors: Neolithic and early Bronze Age mortuary practices. IN Barrett JC & Kinnes I (eds) *The archaeology of context in the Neolithic and early Bronze Age: recent trends*. Sheffield: Sheffield University Press. pp.30-41.

Barrett JC, 1989 Towards an archaeology of Ritual. IN Garwood P, Jennings D, Skeates R & Toms J (eds) *Sacred and Profane*. Oxford University Committee for Archaeology Monograph No. 32. pp.1-9.

Barrett JC, 1994 Fragments from antiquity. An archaeology of social life in Britan, 2900-1200BC. Oxford: Blackwell.

Barrett JC, Bradley R & Green M, 1991 Landscape, Monuments and Society, The Prehistory of Cranborne Chase. Cambridge: Cambridge University Press.

Bell M, 1981 Valley Sediments and Environmental Change. IN Jones M & Dimbleby G (eds) *The Environment of Man: The Iron Age to the Anglo Saxon Period*. Oxford: British Archaeological Reports British Series 87. pp.75-91.

Bell M, 1982 The Effects of Land-Use and Climate on Valley Sedimentation. IN Harding AF (ed) *Climate Change in Later Prehistory*. Edinburgh: Edinburgh University Press. pp.127-142.

Bell M, 1983 Valley Sediments as Evidence of Prehistoric Land-use on the South Downs. *Proceedings of the Prehistoric Society* 49. pp.119-150.

Bell M, 1985 Archaeological Evidence for the Date, Cause and Extent of Soil Erosion on the Chalk. *South East Soils Discussion Group Journal* 3. pp.72-83.

Berry JK, 1987 Fundamental operations in computer aided map analysis. *International Journal of Geographic Information Systems* 1. pp.119-136.

Bradley R, 1978 The Prehistoric Settlement of Britain. London: Routledge and Kegan Paul.

Bradley R, 1984a Regional Systems in Prehistoric Britain. IN Bradley R & Gardiner J (eds) *Neolithic Studies: A Review of Current Work.* Oxford: British Archaeological Reports British Series 133. pp.5-14.

Bradley R, 1984b The Social Foundations of Prehistoric Britain. London: Longman.

Bradley R, 1991 The Pattern of Change in British Prehistory. IN Earle T (ed) *Chiefdoms: Power Economy and Ideology.* Cambridge: Cambridge University Press. pp.44-70.

Bradley R, Cleal R, Gardiner J, Green M, & Bowden M 1984 The Neolithic Sequence in Cranborne Chase. IN Bradley R & Gardiner J (eds) *Neolithic Studies: A Review of Current Work.* Oxford: British Archaeological Reports British Series 133. pp.87-105.

Brewer R, 1979 Principles of Ecology. Philadelphia: WB Saunders.

Burrough PA, 1986 Principles of Geographic Information Systems for Land Resources Management. Oxford: Clarendon Press.

Carter S, 1987 The Incorporation of Land Snail Death Assemblages into Calcareous Soils. Institute of Archaeology, London PhD Thesis.

Carmichael DL, 1990 GIS Predictive Modelling of Prehistoric Site Distributions in Central Montana. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology.* London: Taylor & Francis. pp.216-225.

Carr C, 1985 Introductory remarks on regional analysis. IN Carr C (ed) *For concordance in archaeological analysis.* Kansas City: Westport Publishers. pp114-127.

Case H, 1969 Neolithic Explanations. *Antiquity* 43. pp.176-186.

Castleford J, 1992 Archaeology, GIS, and the Time Dimension: an Overview. IN Lock G & Moffett J (eds) *Computer and Quantitative Methods in Archaeology 1991.* British Archaeological Reports International Series 577. Oxford: Tempus Reparatum. pp.95-106.

Catt JA, 1978 The Contribution of Loess to Soils in Lowland Britain. IN Limbrey S & Evans JG (eds) *The Effect of Man on the Landscape: the Lowland Zone.* London: Council for Brittish Archaeology. pp.12-20.

Chapman G, 1992 Do-it-yourself reconstruction modelling. IN Lock G and Moffett J (eds) *Computer Applications in Archaeology 1991.* British Archaeological Reports International Series 577. Oxford: Tempus Reparatum. pp.213-218.

Chapman RW, 1981 The Emergence of Formal Disposal Areas and the 'Problem' of Megalithic Tombs in Prehistoric Europe. IN Chapman RW, Kinnes I and Randsborg K (eds) *The Archaeology of Death.* Cambridge: Cambridge University Press pp71-81.

Chandler TJ & Gregory S (eds), 1976 The Climate of the British Isles. London: Longman.

Clarke K, 1990 Analytical and computer cartography. London: Prentice Hall.

Cleere H, 1984 Great Britain. IN Cleere H (ed) Approaches to the Archaeological Heritage: a Comparative Study of World Cultural Resource Management Systems. Cambridge:Cambridge University Press. pp55-62.

Cliff AD and Ord JK, 1973 Spatial autocorrelation. *Monographcs in spatial and environmental analysis.* London: Pion.

Connah G, 1961 Excavations at Knap Hill, Alton Priors, 1961. *Wiltshire Archaeological Magazine* 60 pp.1-23.

Crumley CL and Marquardt WH, 1990 Landscape: a unifying concept in regional analysis. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology.* London: Taylor & Francis. pp.73-79.

Curtis LF, Courtney FM & Trudgill ST, 1976 Soils of the British Isles. London: Longman.

Darvill T, 1992 Valuing Britain's Archaeological Resource. British Property Federation Chair in Archaeology and Development Inaugural Lecture. Unpublished.

Darvill T, Saunders A & Startin W, 1987 A Question of National Importance: Approaches to the Evaluation of Ancient Monuments for the Monuments Protection Programme in England. *Antiquity* 61. pp.393-408.

Dean SD & Robinson WJ, 1979 Computer Cartography and the Reconstruction of Dendroclimatic Variability in the American Southwest. IN Upham S (ed) *Computer graphics in archaeology : statistical cartographic applications to spatial analysis in archaeological contexts.* Arizona State University Anthropological Research Papers no 15. pp.79-94.

Devereux P, 1991 Three dimensional aspects of apparent relationships between selected natural and artificial features within the topography of the Avebury complex. *Antiquity* 65. pp.894-898.

Dimbleby GW, 1976 Climate, Soil and Man. *Philosophical Transactions of the Royal Society of London* 275. pp.197-208.

Dimbleby GW, 1984 Anthropogenic Changes from Neolithic through Medieval Times. *New Phytologist* 98. pp.55-72.

Dimbleby GW & Evans JGE, 1974 Pollen and Land-Snail Analysis of Calcareous Soils. *Journal of Archaeological Science* 1. pp.117-133.

DoE, 1987 Handling geographic information. Report to the Secretary of State for the Environment from the Committee of Enquiry into the Handling of Geographic Information, chaired by Lord Chorley. London: HMSO.

DoE, 1992 Planning Policy Guidance Note 16. London: HMSO.

Downie NM & Heath RW, 1965 Basic Statistical Methods. Harper & Row:London.

Earle T (ed), 1991 Chiefdoms: Power Economy and Ideology. Cambridge University Press.

Earle T, 1991 Property rights and the evolution of Chiefdoms IN Earle T (ed) *Chiefdoms: Power Economy and Ideology.* Cambridge: Cambridge University Press. pp.71-99.

Earle T, 1993 Paths and roads in evolutionary perspective. IN Trombold CD (ed) *Ancient road networks and settlement hierarchies in the New World.* Cambridge: Cambridge University Press. pp.10-16.

Eastman RJ, 1992 IDRISI V4.0 User Guide. Worcester, Massachusets: Clarke University.

Eastman RJ, 1993 IDRISI V4.1 Update Manual. Worcester, Massachusets: Clarke University.

Ebert JI, 1988 Remote sensing in archaeological projection and prediction. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past.* Denver: US Department of the Interior Bureau of Land Management. pp.429-492.

Ebert JI & Kohler TA, 1988 The theoretical basis of archaeological predictive modeling and a consideration of appropriate data-collection methods. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past.* Denver: US Department of the Interior Bureau of Land Management. pp.97-171.

English Heritage, 1991 Exploring our Past: Strategies for the Archaeology of England. London: Historic Buildings and Monuments Commission (England).

Entwhistle R, 1989 Relativism and interpretation in prehistoric archaeology : some thoughts on the formulation of archaeological evidence, with special reference to the use of palaeoevironmental data. University of Reading PhD Thesis.

Evans C, 1988 Acts of Enclosure: A Consideration of Concentrically Organised Causewayed Enclosures. IN Barret JC, and Kinnes IA (eds) *The Archaeology of Context in the Neolithic and Bronze Age: Recent Trends.* Sheffield: University of Sheffield. pp.85-96.

Evans JG, 1970 Interpretation of Land Snail Faunas. *Bulletin of the Institute of Archaeology* 8 pp.109-116.

Evans JG, 1972 Land Snails in Archaeology. London: Seminar Press.

Evans JG, 1978 Introduction to Environmental Archaeology. London: Granada Publishing.

Evans JG, 1983 Excavations at Cherhill, North Wiltshire, 1967. *Proceedings of the Prehistoric Society* 49 pp.43-117.

Evans JG, 1984 Stonehenge - The Environment in the Late Neolithic and Early Bronze Age and a Beaker-Age Burial. *Wiltshire Archaeological and Natural History Magazine* 78 pp.7-30.

Evans JG, 1990 Notes on some Late Neolithic and Bronze Age Events in Long Barrow Ditches in Southern and Eastern England. *Proceedings of the Prehistoric Society* 56. pp.111-116.

Evans JG & Jones H, 1979 Mount Pleasant and Woodhenge: The Land Molusca. IN Wainwright G, *Mount Pleasant, Dorset, Excavations 1970-1971.* Report of the Research Committee of the Society of Antiquaries. London: Thames and Hudson. pp.190-213.

Evans JG, Limbrey S, Mate I & Mount RJ, 1988a Environemntal Change and Land-use History in a Wiltshire River Valley in the Last 14000 years. IN Barrett JC & Kinnes IA (eds) *The Archaeology of Context in the Neolithic and Bronze Age: Recent Trends.* Sheffield: University of Sheffield. pp.97-103.

Evans JG, Pitts MW & Williams D, 1985 An Excavation at Avebury, Wiltshire, 1982. *Proceedings of the Prehistoric Society* 51. pp.305-320.

Evans JG, Rouse AJ & Sharples NM, 1988b The Landscape Setting of Causewayed Camps: Recent Work on the Maiden Castle Enclosure. IN Barrett JC & Kinnes IA (eds) *The*

Archaeology of Context in the Neolithic and Bronze Age: Recent Trends. Sheffield: University of Sheffield. pp.73-84.

Farley JA, Limp WF & Lockhart J, 1991 The Archaeologists Workbench: Integrating GIS, Remote Sensing, EDA and Database Management IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology*. London: Taylor & Francis. pp.141-164.

Fasham PJ, 1980 Excavations at Bridgets and Burntwood Farms, Itchen Valley Parish, Hampshire. *Proceedings of the Hampshire Field Club and Archaeological Society* 36. pp. 37-86.

Feder KL, 1979 Geographic patterning of tool types as elicited by trend surface analysis. IN Upham S (ed) *Computer graphics in archaeology: statistical cartographic applications to spatial analysis in archaeological contexts*. Arizona State University Anthropological Research Papers No. 15. pp.95-102.

Firth A, 1995 (forthcoming) Ghosts in the machine. IN Cooper M, Carmen J, Firth A & Wheatley DW (eds) *Managing Archaeology*. London: Routledge.

Fisher P, 1991 First experiments in viewshed uncertainty: the accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing* 57(10). pp.1321-1327.

Fisher P, 1992 First experiments in viewshed uncertainty: simulating fuzzy viewsheds. *Photogrammetric Engineering and Remote Sensing* 58(3). pp.345-352.

Fisher PF, 1991 The Physical Environment of Cranborne Chase. IN Barret J, Bradley R and Hall M (eds) *Papers on the Prehistoric Archaeology of Cranborne Chase*. Oxford: Oxbow Books Monograph 11. pp.11-19.

Fleming A, 1971 Territorial patterns in Bronze Age Wessex. *Proceedings of the Prehistoric Society* 37(I) pp.138-166.

Foley R, 1981 A Model of Regional Archaeological Structure. *Proceedings of the Prehistoric Society* 47 pp.1-17.

Fowler PJ & Evans JG, 1967 Plough-marks, Lynchets and Early Fields. *Antiquity* XLI. pp.289-301.

Gaffney V & Stancic Z, 1991 GIS Approaches to Regional Analysis: A Case Study of the Island of Hvar. Ljubljana: Filozofska fakulteta.

Gaffney V & Stancic Z, 1992 Diadorus Siculus and the Island of Hvar, Dalmatia: Testing the Text with GIS. IN Lock G & Moffett J (eds) *Computer and Quantitative Methods in Archaeology 1991*. British Archaeological Reports Int. Series S577 Oxford: Tempus Reparatum. pp.113-126.

Gaffney V & Stancic Z, (forthcoming) title unavailable IN Aldenderfer M & Maschner H (eds), *The anthropology of human behaviour through geographic information and analysis.* New York: Oxford University.

Gardiner JP, 1984 Lithic Distributions and Neolithic Settlement Patterns in Central Southern England. IN Bradley R & Gardiner J (eds) *Neolithic Studies: A Review of Current Work.* Oxford: British Archaeological Reports British Series 133 pp.177-187.

Gorenflo LJ and Bell TL, 1993 Network analysis and the study of past regional organization. IN Trombold CD (ed) Ancient road networks and settlement hierarchies in the New World. Cambridge: Cambridge University Press. pp.80-98.

Goudie A 1981, The Human Impact on the Natural Environment. Oxford: Blackwell.

Groenman-van Waateringe W, 1983 The Earliest Agricultural Utilisation of the Irish Landscape: The Last Word on the Elm Decline? IN Reeves-Smythe, T and Hamond, F (eds) *Landscape Archaeology in Ireland.* Oxford: British Archaeological Reports British Series 116. pp.217-232.

Grinsell LV, 1958 The Archaeology of Wessex. London.

Groube LM & Bowden MCB, 1982 The Archaeology of Rural Dorset. *Dorset Natural History and Archaeology Society Monograph 4.* Dorchester: Friary Press.

Harris TM, 1986 Geographic Information System Design for Archaeological Site Information Retrieval. IN Laflin S (ed) *Computer Applications in Archaeology 1986.* Birmingham: University of Birmingham. pp.148-161.

Harris TM, 1987 Digital Terrain Modelling and Three-dimensional Surface Graphics for Landscape and Site Analysis in Archaeology and Regional Planning. IN Ruggles C and Rahtz SPQR (eds) *Computer and Quantitative Methods in Archaeology 1987.* British Archaeological Reports International Series 393. Oxford: British Archaeological Reports. pp.161-170.

Harris TM, & Lock GR, 1990 The Diffusion of a New Technology: a Perspective on the Adoption of Geographic Information Systems Within UK Archaeology. IN Allen

KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology.* London: Taylor & Francis. pp.33-53.

Hasenstab RJ & Resnick B, 1990 GIS in Historical Predictive Modelling: the Fort Drum Project. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology*. London: Taylor & Francis. pp.284-306.

Hodder I and Orton C, 1976 Spatial analysis in archaeology. Cambridge: Cambridge University Press.

Holgate R, 1988 A review of Neolithic domestic activity in Southern Britan. IN Barrett JC & Kinnes I (eds) *The archaeology of context in the Neolithic and early Bronze Age: recent trends*. Sheffield: Sheffield University Press. pp.104-112.

Hunt ED, 1992 Upgrading site catchment analysis with GIS. *World Archaeology* 24(2) pp.284-309.

Ingold T, 1993 The temporality of the landscape. *World Archaeology* 25(2) pp.152-167.

Iverson J, 1941 Landnam i Danmarks Stenalder *Danmarks Geologiske Undersøgelse* 4(66), pp.20-68.

Jacobi RM, 1976 Britain Inside and Outside Mesolithic Europe. *Proceedings of the Prehistoric Society* 42 pp.67-84.

Jacobi RM, 1979 Early Flandrian Hunters in the South West. *Proceedings of the Devon Archaeological Society* 37, pp.48-93.

Jarman MR, Bailey GN & Jarman HN, 1982 Early European Agriculture. Cambridge: Cambridge University Press.

Jermann JV & Dunnell R, 1979 Some Limitations of Isopleth Mapping in Archaeology. IN Upham S (ed) *Computer graphics in archaeology: statistical cartographic applications to spatial analysis in archaeological contexts.*. Arizona State University Anthropological Research Papers no 15. pp.31-60.

Kincaid C, 1988 Predictive modeling and its relationship to cultural resource management applications. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past.* Denver: US Department of the Interior Bureau of Land Management. pp.549-569.

Kohler TA, 1988 Predictive locational modeling: history and current practice. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past.* Denver: US Department of the Interior Bureau of Land Management. pp.19-59.

Kohler TA & Parker SC, 1986 Predictive models for archaeological resource location. IN Schiffer M (ed) *Advances in Archaeological Method and Theory* Vol 9. New York: Academic Press. pp.397-452.

Kristiansen K, 1984 Denmark. IN Cleere H (ed), *Approaches to the Archaeological Heritage: a Comparative Study of World Cultural Resource Management Systems.* Cambridge: Cambridge University Press. pp.55-62.

Kvamme KL, 1983a Computer processing techniques for regional modeling of archaeological site locations. *Advances in Computer Archaeology* VI. pp.26-52.

Kvamme KL, 1983b A manual for predictive site location models: examples from the Grand Junction District, Colorado. Colorado: Bureau of Land Management, Grand Junction District.

Kvamme KL, 1985a Fundamentals and Potential of Geographic Information System Techniques for Archaeological Spatial Research. Paper presented at 50th meeting of Society of Americal Archaeology, Denver.

Kvamme KL, 1985b Determining empirical relationships between the natural environment and prehistoric site locations: a hunter-gatherer example IN Carr C (ed) *For concordance in archaeological analysis.* Kansas City MO: Westport Publishers. pp.208-238.

Kvamme KL, 1988 Using existing archaeological survey data for model building. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past* Denver: US Department of the Interior Bureau of Land Management. pp.301-428.

Kvamme KL, 1989 Geographic Information Systems in Regional Archaeological Research and Data Management. IN Schiffer MB (ed) *Archaeological Method and Theory.* Volume 1 Tuscon: University of Arizona Press. pp.139-203.

Kvamme KL, 1990a The Fundamental Principles and Practice of Predictive Archaeological Modelling. IN Voorips A, (ed) *Mathematics and Information Science in Archaeology: A Flexible Framework. Studies in Modern Archaeology.* Volume 3. Bonn: Holos-Verlag. pp.257-295.

Kvamme KL, 1990b One-Sample Tests in Regional Archaeological Analysis: New Possibilities Through Computer Technology. *American Antiquity* 55(2) pp367-381.

Kvamme KL, 1990c GIS Algorithms and their effects on regional archaeological analyses. IN Allen KMS, Green SW & Zubrow EBW *Interpreting Space: GIS and Archaeology.* London: Taylor & Francis. pp.112-125.

Kvamme KL, 1990d Spatial autocorrelation and the Classic Maya collapse revisited: refined techniques and new conclusions. *Journal of Archaeological Science* 1990 pp.197-207.

Kvamme KL, 1992a Geographic Information Systems in Archaeology. IN Lock G & Moffett J (eds) *Computer and Quantitative Methods in Archaeology 1991.* British Archaeological Reports International Series 577. Oxford: Tempus Reparatum. pp.77-84.

Kvamme KL, 1992b Terrain Form Analysis of Archaeological Location Through Geographic Information Systems. IN Lock G & Moffett J (eds) *Computer and Quantitative Methods in Archaeology 1991.* British Archaeological Reports International Series 577. Oxford: Tempus Reparatum. pp.127-136.

Kvamme KL, 1993 Spatial statistics and GIS: an integrated approach. IN Andresen *et al.* (eds) *Computing the Past - CAA92.* Aarhus: Aarhus University Press. pp.91-103.

Kvamme KL and Jochim MA, 1989 The Environmental Basis of Mesolithic Settlement. IN Bonsall C (ed) *The Mesolithic in Europe: Papers Presented at the Third International Symposium, Edinburgh 1985.* Edinburgh: John Donald Publishers pp.1-12.

Lang NAR, 1993 From model to machine: procurement and implementation of Geographic Information Systems for county sites and monuments records. IN Andresen *et al.* (eds) *Computing the Past - CAA92.* Aarhus: Aarhus University Press. pp.167-176.

Limbrey S, 1975 Soil science and archaeology. London: Academic Press.

Lock GR & Harris TM, 1991 Visualising Spatial Data: The Importance of Geographic Information Systems. IN Reilly P & Rahtz P (eds) *Communication in Archaeology: a Global View of the Impact of IT.* One World Archaeology Series. London: Unwin Hyman. pp.81-96.

Lock GR & Harris TM, (forthcoming) Danebury revisited: an English Iron Age hillfort in a digital landscape. IN Aldenderfer M & Maschner H (eds), *The anthropology of human behaviour through geographic information and analysis.* New York: Oxford University Press.

Magurran AE, 1988 Ecological Diversity and Its Measurement. London: Croom Helm.

Marble DF, 1987 The Computer and Cartography. *The American Cartographer* 14 pp.101-103.

Marble DF, 1988 Approaches to the efficient design of spatial databases at a global scale. IN Mountsey H & Tomlinson R *Building databases for global science*. London: Taylor & Francis. pp.49-65.

Marble DF, 1990 The Potential Methodological Impact of Geographic Information Systems on the Social Sciences. IN Allen KMS, Green SW & Zubrow EBW *Interpreting Space: GIS and Archaeology*. London: Taylor & Francis. pp.9-21.

Martin D, 1991 Geographic Information Systems and their Socioeconomic Applications. London: Routledge.

Maskell JE, 1993 From ploughzone to pixel. University of Southampton: MSc Thesis.

Mercer R, 1980 Hambledon Hill: a neolithic landscape. Edinburgh: Edinburgh University Press.

Middleton R & Winstanley D, 1993 GIS in a landscape archaeology context. IN Andresen *et al.* (eds) *Computing the Past - CAA92*. Aarhus: Aarhus University Press. pp.151-158.

Mizoguchi K, 1993 Time in the reproduction of mortuary practices. *World Archaeology* 25, 2 pp.223-235.

Molyneaux B, 1992 Virtuality to actuality: archaeological site simulation. IN Reilly P and Rahtz SPQ *Archaeology and the information age: a global perspective*. London: Routlege. pp.312-322.

Mount RJ, 1991 The environmental history of the Upper Kennett River valley: some implications for human communities. University College Cardiff PhD Thesis.

Ozawa K, 1993 Reconstruction of Japanese ancient tombs and villages. IN Andresen *et al.* (eds) *Computing the Past - CAA92*. Aarhus: Aarhus University Press. pp.415-423.

Palmer R, 1976 Interrupted Ditch Enclosures in Britain: The Use of Aerial Photography for Comparative Studies. *Proceedings of the Prehistoric Society* 42 pp.161-186.

Peacock DPS, 1969 Neolithic pottery production in Cornwall. *Antiquity* 43 pp.145-149.

Peuquet DJ, 1988 Issues involved in selecting appropriate data models for global databases. IN Mountsey H & Tomlinson R *Building databases for global science*. London: Taylor & Francis. pp.66-78.

Piggot S, 1962 The West Kennet Long Barrow. London: HMSO.

Rackham O, 1978, Trees and Woodland in the British Landscape. London: Dent.

Reilly P, 1991 Towards a virtual archaeology. IN Lockyear K and Rahtz *Computer Applications in Archaeology 1990.* British Archaeological Reports International Series 565 Oxford: Tempus Reparatum. pp.133-140.

Reilly P, 1991 Visualising the problem: advancing graphic systems in archaeological analysis. IN Ross S, Moffett J and Henderson J *Computing for Archaeologists.* Oxford: Oxford University Committee for Archaeology Monograph No. 18. pp.131-151.

Reilly P, 1992 Three-dimensional modelling and primary archaeological data. IN Reilly P and Rahtz SPQ *Archaeology and the information age: a global perspective.* London: Routlege. pp.147-173.

Renfrew C, 1973 Monuments Mobilization and Social Organization in Neolithic Wessex. IN Renfrew, C (ed) *The Explanation of Culture Change: Models in Prehistory.* London: Duckworth. pp.539-558.

Rhind D, 1977 Computer-aided Cartography. *Transactions of the Institute of British Geographers* 2. pp.71-96.

Richards J, 1984 The Development of the Neolithic Landscape in the environs of Stonehenge. IN Bradley R & Gardiner J (eds) *Neolithic Studies: A Review of Current Work.* Oxford: British Archaeological Reports British Series 133 pp.177-187.

Richards J, 1990 The Stonehenge Environs Project. Southampton: Hobbs.

Richards JD, 1986 Computers in archaeological theory and practice. *Science and Archaeology* 28. pp.51-55.

Roberts BK, 1987 Landscape Archaeology. IN Wagstaff JM (ed) *Landscape and Culture.* Oxford: Blackwell. pp.77-95.

Robertson-Mackay ME, 1980 A Head and Hooves Burial Beneath a Round Barrow, with other Neolithic and Bronze Age Sites, on Hemp Knoll, Near Avebury, Wiltshire. *Proceedings of the Prehistoric Society* 46 pp.123-176.

Rose MR and Altschul JH, 1988 An overview of statistical method and theory for quantitative model building. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting* *the past.* Denver: US Department of the Interior Bureau of Land Management. pp.173-255.

Ruggles C, 1992 Abstract Data Structures for GIS Applications in Archaeology. IN Lock G & Moffett J (eds) *Computer and Quantitative Methods in Archaeology 1991*. British Archaeological Reports International Series 577 Oxford: Tempus Reparatum. pp.107-112.

Ruggles C, Medyckyj-Scott DJ & Gruffydd A, 1993 Multiple Viewshed analysis using GIS and its archaeological application: a case study in northern Mull. IN Andresen *et al.* (eds) *Computing the Past - CAA92*. Aarhus: Aarhus University Press. pp.125-132.

Sahlins MD, 1958, Social Stratification in Polynesia. Seatle: University of Washington.

Salisbury EJ & Jane FW, 1940 Charcoals from Maiden Castle and their Significance in Relation to the Vegetation and Climate Conditions in Prehistoric Times. *Journal of Ecology* 28. pp.310-325.

Santley RS, 1993 The structure of the Aztec transport network. IN Trombold CD(ed) *Ancient road networks and settlement hierarchies in the New World.* Cambridge: Cambridge University Press. pp.198-210.

Savage SH, 1990 GIS in archaeological research. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology*. London: Taylor & Francis. pp.22-32.

Schofield J, 1988 The Interpretation of Surface Lithic Collections: Case Studies from Southern England. University of Southampton PhD Thesis.

Sebastian L & Judge WJ, 1988 Predicting the past: correlation, explanation and the use of archaeological models IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past* Denver: US Department of the Interior Bureau of Land Management. pp.1-18 .

Shanks M & Tilley C, 1982 Ideology, Symbolic Power and Ritual Communication: a Reinterpretation of Neolithic Mortuary Practices. IN Hodder, I (ed) *Symbolic and Structural Archaeology*. Cambridge: Cambridge University Press. pp.129-154.

Shellard HC, 1976 Wind. IN Chandler TJ and Gregory S (eds) *The Climate of the British Isles.* London: Longman.

Shennan SJ, 1980 Meeting the plough damage problem: a sampling approach to area-intensive fieldwork. IN Highcliffe J & Schadler Hall RT (eds) *The past under the plough.* Department of the Environment Occasional Paper No. 3. pp.125-133.

Shennan SJ, 1985 Experiments in the collection and analysis of archaeological survey data: the East Hampshire Survey. Sheffield: Sheffield University Press.

Shennan SJ, 1988 Quantifying Archaeology. Edinburgh: Edinburgh University Press.

Shennan SJ, 1993 After social evolution: a new archaeological agenda? IN Yoffee N & Sherratt A (eds) *Archaeological theory: who sets the agenda?* Cambridge: Cambridge University Press. pp.53-59.

Simmons IG and Tooley MJ, 1981 The Environment in British Prehistory. London: Duckworth.

Smith AG, 1970 The Influence of Mesolithic and Neolithic Man on British Vegetation: A Discussion. IN Walker D & West RG (eds) *Studies in the Vegetational History of the British Isles.* Cambridge: Cambridge University Press. pp.81-95.

Smith CJ, 1980 The Ecology of the English Chalk. London: Academic Press.

Smith IF, 1965 Windmill Hill and Avebury, Excavations by Alexander Keiller 1925-1939. Oxford: Clarendon Press.

Smith IF, 1971 Causewayed Enclosures. IN Simpson DDA (ed) *Economy and Settlement in Neolithic and Early Bronze Age Britian and Europe.* Leicester: Leicester University Press. pp.89-112.

Smith IF, 1974 The Neolithic. IN Renfrew C (ed) *British Prehistory: A New Outline.* London: Duckworth. pp.100-136.

Smith RW, 1984 The Ecology of Neolithic Farming Systems as Exemplified by the Avebury Region of Wiltshire. *Proceedings of the Prehistoric Society* 50. pp.99-120.

Smith RW, 1985 Prehistoric Human Ecology in the Wessex Chalklands with Special Reference to Evidence From Valleys. University of Southampton PhD Thesis.

Star J & Estes J, 1990 Geographic information systems: an Introduction. London: Prentice Hall.

Stone JFS & Gray Hill N, 1938 A Middle Bronze Age Site at Stockbridge, Hampshire *Proceedings of the Prehistoric Society* 4. pp.249-257.

Thomas DH, 1975 Non-Site Sampling in Archaeology: Up the Creek without a Site. IN Mueller JW (ed) *Sampling in Archaeology*. Arizona: University of Arizona Press.

Thoms AV, 1988 A survey of predictive locational models: examples from the late 1970s and early 1980s. IN Judge WJ & Sebastian L (eds) *Quantifying the present, predicting the past*. Denver: US Department of the Interior Bureau of Land Management. pp.581-645.

Thomas J, 1984 A tale of Two Polities: Kinship, Authority and Exchange in the Neolithic of South Dorset and North Wiltshire. IN Bradley R & Gardiner J (eds) *Neolithic Studies: A Review of Current Work*. Oxford: British Archaeological Reports British Series 133. pp.161-175.

Thomas J, 1991 Rethinking the Neolithic. Cambridge: Cambridge University Press.

Thomas J & Whittle A, 1986 Anatomy of a tomb – West Kennet revisited. *Oxford Journal of Archaeology* 5(2) pp.129-156.

Thomas K, 1982 Neolithic Enclosures and Woodland Habitats on the South Downs in Sussex, England. IN Bell M & Limbrey S (eds) *Archaeological Aspects of Woodland Ecology*. Oxford: British Archaeological Reports International Series 146. pp.147-170.

Thorley A, 1981 Pollen Analytical Evidence Relating to the Vegetation History of the Chalk. *Journal of Biogeography* 8 pp.93-106.

Thorpe IJ, 1984 Ritual Power and Ideology: A Reconstruction of Earlier Neolithic Rituals in Wessex. IN Bradley R & Gardiner J (eds) *Neolithic Studies: A Review of Current Work*. Oxford: British Archaeological Reports British Series 133. pp.41-60.

Tivy J & O'Hare G, 1981 The Human Impact on the Ecosystem. Edinburgh: Oliver & Boyd.

Troels-Smith 1960 Ivy, Mistletoe and Elm: Climatic Indicators – Fodder Plants: a Contribution to the Interpretation of the Pollen Zone Border VII-VIII. *Danmarks Geologiske Undersøgelse* 4(4). pp.1-32.

Upham S, 1979 Computer graphics in archaeology: statistical cartographic applications to spatial analysis in archaeological contexts. Arizona: Arizona State University Anthropological Research Papers No 15.

US Corps of Engineers, 1993 GRASS 4.1 Users Reference Manual. United States Army Corps of Engineers Construction Engineering Research Laboratory, Illinois. Distributed in digital format.

van-Leusen PM, 1993 Cartographic modelling in a cell-based GIS. IN Andresen *et al.* (eds) *Computing the Past – CAA92*. Aarhus: Aarhus University Press. pp.105-123.

Vatcher F de M, 1961 The Excavation of the Long Mortuary Enclosure on Normanton Down, Wiltshire. *Proceedings of the Prehistoric Society* 27. pp.160-163.

Vita-Finzi C & Higgs ES, 1970 Prehistoric economy in the Mount Carmel area of Palestine: Site Catchment Analysis. *Proceedings of the Prehistoric Society* 36. pp.1-37.

Wainwright G, 1979 Mount Pleasant, Dorset: Excavations 1970-1971. Report of the Research Committee of the Society of Antiquaries. London: Thames and Hudson.

Wansleeben M, 1988 Applications of Geographical Information Systems in Archaeological Research. IN Rahtz SPQ (ed) *Computer and Quantitative Methods in Archaeology 1987*. British Archaeological Reports International Series 446. Oxford: British Archaeological Reports. pp.435-451.

Wagstaff JM (ed), 1987 Landscape and Culture. Oxford: Blackwell.

Warren RE, 1990a Predictive Modelling of Archaeological Site Location: a Case Study in the Midwest. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology*. London: Taylor & Francis. pp.90-111.

Warren RE, 1990b Predictive Modelling of Archaeological Site Location: a Primer. IN Allen KMS, Green SW & Zubrow EBW Interpreting Space: GIS and Archaeology. London: Taylor & Francis. pp.201-215.

Waton PV, 1983 A Palynological Study of the Impact of Man on the Landscape of Central Southern England, with Special Reference to the Chalklands. University of Southampton PhD Thesis.

Wheatley DW, 1993 Going over old ground: GIS, archaeological theory and the act of perception. IN Andresen *et al.* (eds) *Computing the Past – CAA92*. Aarhus: Aarhus University Press. pp.133-138.

Wheeler REM, 1943 Maiden Castle Dorset. Oxford: Report of the Research Committee of the Society of Antiquaries. No.12.

White R, 1969 Mental maps of Britan. *New Society* 327.

Whittle AWR, 1978 The Earlier Neolithic of Southern England and its Continental Background. Oxford: British Archaeological Reports International Series 35.

Whittle AWR, 1985 Neolithic Europe: A Survey. Cambridge: Cambridge University Press.

Whittle AWR, 1990 A Model for the Neolithic-Mesolithic Transition in the Upper Kennett Valley, North Wiltshire. *Proceedings of the Prehistoric Society* 56. pp.101-110.

Yoffee N, 1993 Too many chiefs? (or safe texts for the '90s). IN Yoffee N & Sherratt A (eds) *Archaeological theory: who sets the agenda.* Cambridge: Cambridge University Press. pp.60-78.

Zubrow EBW, 1990a Contemplating space: a commentary on theory. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology.* London: Taylor & Francis. pp.68-72.

Zubrow EBW, 1990b Modelling and prediction with geographic information systems: a demographic example from prehistoric and historic New York. IN Allen KMS, Green SW & Zubrow EBW (eds) *Interpreting Space: GIS and Archaeology.* London: Taylor & Francis. pp.307-318.