18

Pattern recognition in sub-surface artefact distributions: expanding the role of computer graphics in the analysis of buried data

Paul Reilly*

Alison M. Locker†

Stephen J. Shennan‡

18.1 The Bronze Age midden at Potterne, Wiltshire

Trial excavation by the Trust for Wessex Archaeology revealed that the make-up of a late Bronze Age midden at Potterne in Wiltshire consisted of a dark grey ash matrix full of well preserved pottery, animal bone and other finds (Gingell & Lawson 1985). Although the top and bottom of the midden were clearly defined, no stratigraphy was visible when the midden was trial excavated. Consequently, the methodology employed during subsequent excavation involved the creation of a system of arbitrary context recording units consisting of one metre squares excavated in 10-centimetre spits. It transpired that when the excavation trench was complete it was possible to discern a certain amount of stratigraphy in the sections.

An understanding of the midden's stratigraphy is considered crucial to the interpretation of both the site chronology and the processes that contributed to the midden's formation. A number of tentative but conflicting interpretations had evolved in the course of the excavations. One theory proposed that the midden was composed of a

* IBM UK Scientific Centre
Athelstan House
St Clement House
Winchester
Hants SO23 9DR

† HBMC
Fortress House
23 Savile Row
London W1X 2HE

‡ Department of Archaeology,
University of Southampton,
Southampton S09 5NH

series of separate dumps, while another postulated a fairly uniform build-up across its full extent.

The graphics-database system called the Winchester Graphics System, or WGS (Heywood *et al.* 1984), which was also used to help analyse the point-provenanced contents of a rubbish pit from Saxon Southampton (Colley *et al.* 1988), is now being applied to the understanding of the composition of the midden at Potterne.

WGS pictures are defined by means of simple primitives such as points, lines, markers, and text strings, all of which may be coloured as required. The primitives are collected into segments, which can be created or deleted independently, or temporarily hidden or re-displayed, in order to produce a series of related pictures. A bridging program reads a table from the database and translates this into a picture segment. Data defining the extent and layout of the excavation, the three-dimensional position of the excavation units, the distribution of various artefact categories, and the section drawings are held in the database as relations.

The graphics considerably speed up the process of identifying aspects of midden structure and composition which could not be directly perceived when the midden was excavated. As in the Hamwic pit project, the graphics database combination is being exploited to pursue a type of analysis which has been styled the 'Investigative Loop'. This means that the graphics are not employed merely to present the data, they are being applied as valuable exploratory or probing tools. Essentially the loop starts with an idea about the data being examined. The idea need not be very complicated. For example, it was hypothesised that the midden built up rapidly, with large volumes of rubbish being tipped onto the midden and the material being mixed up further by scavengers and other post-depositional disturbances. The researcher could test this hypothesis by selecting and displaying those contexts containing material of a certain type or date. For instance, the distribution of refitting sherds might provide valuable clues about site stratigraphy. Such information can be represented and looked at in many different ways, especially when the graphical models being returned can be examined on display units supporting colour and real-time manipulation. For instance, contexts containing refitting sherds from a particular vessel can be displayed using colour codes to show the spread of the material. A careful study has identified a substantial number of pots with refitting sherds spread across several contexts. Another way of expressing this is to join every refit with a connecting line (Fig. 18.1 (B)). Drawing lines to connect refitting sherds from different vertical contexts indicated little vertical displacement, which implies that large amounts of the midden deposits accumulated uniformly without major disturbance (Fig. 18.1 (A)).

During this interaction the user may notice some other feature in the data which may cause a new theory to be formulated or the original question to be modified. The flexibility of the combination of database and graphics positively encourages excavators to examine a variety of data configurations to explain the midden's stratigraphy.

It has been discovered that two particular techniques of graphical representation provide powerful data visualisation tools. The first method is to regard the computer model as being transparent, allowing the analyst to look right *through* the midden matrix. The other method is to treat the midden model as an opaque solid and to remove or clip away successive horizontal or vertical sections through the midden structure, analysing the resulting profiles. This is analogous to drawing a plan or elevation during the excavation: each profile is an arbitrary thin slice of the formation

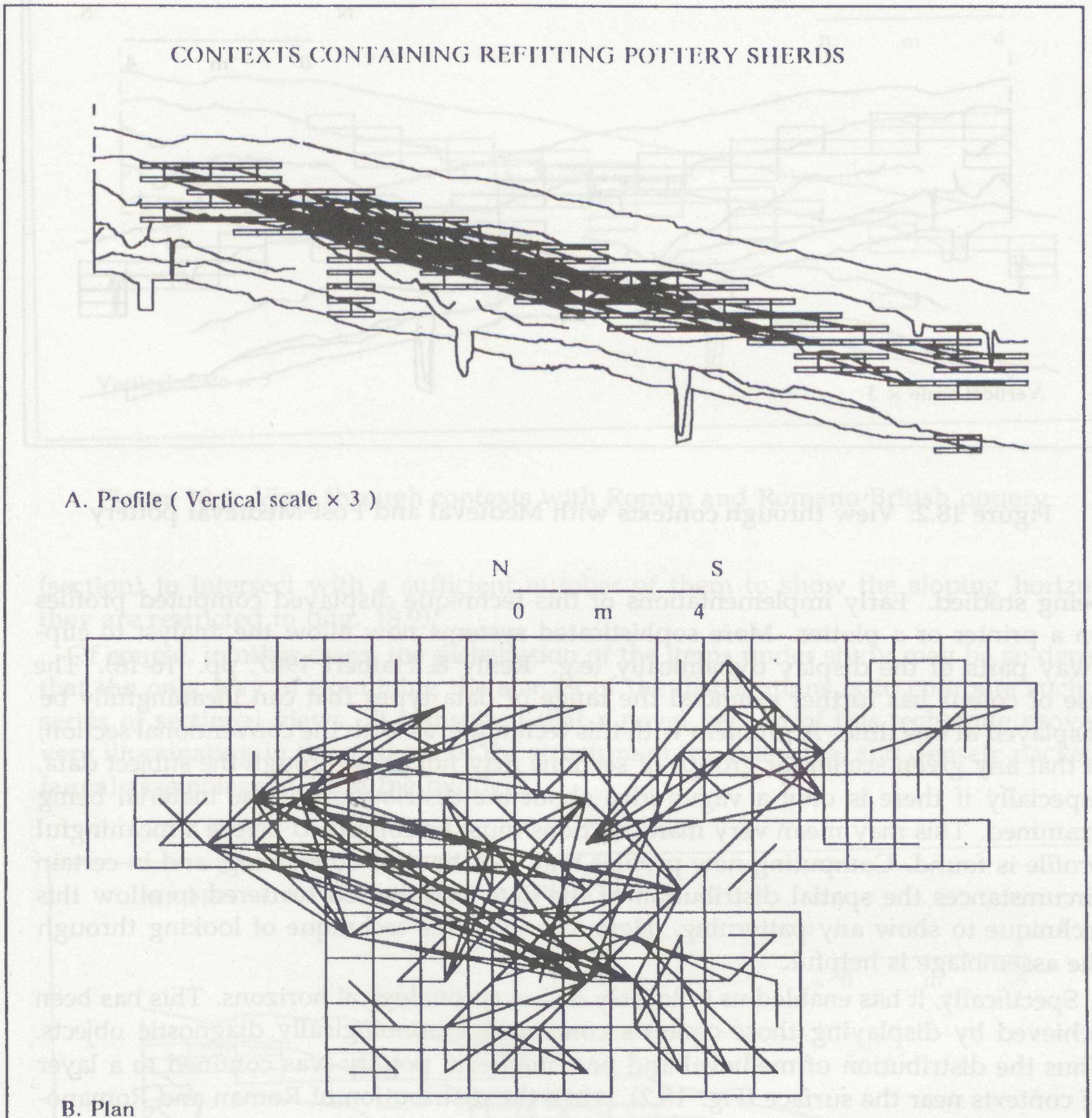

Figure 18.1: Distribution of refitting pottery sherds at Potterne.

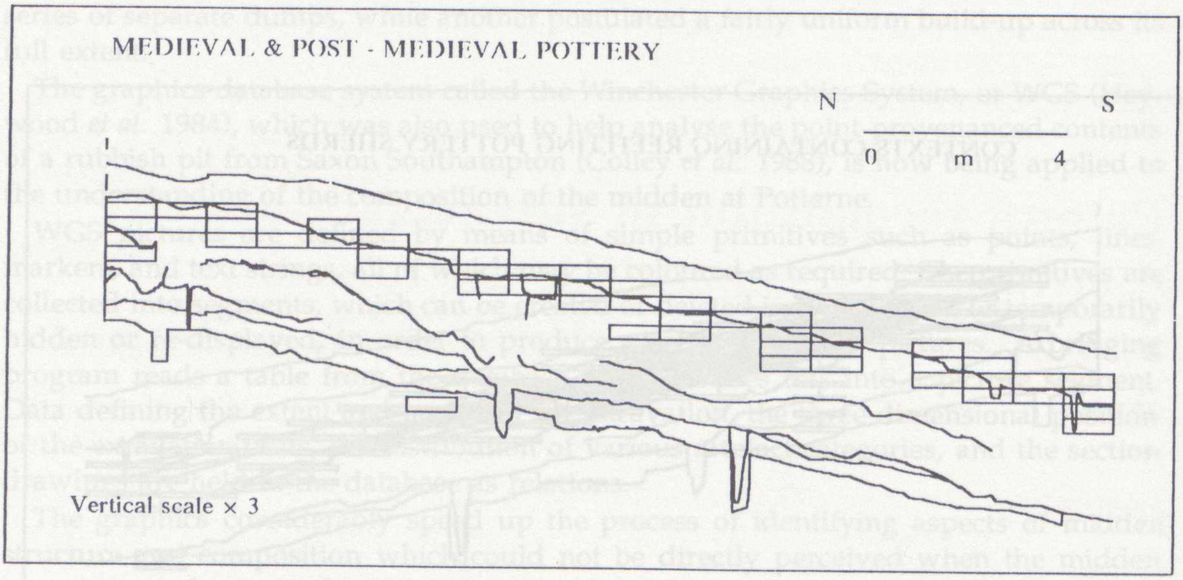

Figure 18.2: View through contexts with Medieval and Post-Medieval pottery

being studied. Early implementations of this technique displayed computed profiles on a printer or a plotter. More sophisticated systems now allow the analyst to clip-away parts of the display dynamically (e.g. Reilly & Halbert 1987, pp. 16-18). The use of colour has further extended the range of data types that can meaningfully be displayed at one time. A problem with this technique, as with the conventional section, is that any given section or groups of sections may not pass through the subject data, especially if there is only a vague idea about the distribution of the material being examined. This may mean very many sections must be computed before a meaningful profile is found. Computing new profiles may also be time consuming, and in certain circumstances the spatial distribution of the data may be too scattered to allow this technique to show any patterning. Here, however, the technique of looking through the assemblage is helpful.

Specifically, it has enabled us to loosely define chronological horizons. This has been achieved by displaying those contexts containing chronologically diagnostic objects. Thus the distribution of medieval and post-medieval pottery was confined to a layer of contexts near the surface (Fig. 18.2), while the distribution of Roman and Romano-British pottery began at a slightly lower level, but seemed to mix to some extent with the later types, suggesting a degree of disturbance and residual matter (Fig. 18.3). It is also notable that neither category was present in a large proportion of the centre of the excavated area (Figs. 18.4 & 18.5).

It is noticeable that the horizons thus defined are not horizontal but mirror the slope of the surface of the midden. Iron technology begins at a slightly lower level (Fig. 18.6). The iron objects did not avoid any areas in particular. Reassuringly, the copper alloys start at the bottom of the midden and are found in virtually all levels (Fig. 18.7).

One or two pottery motifs are also confined to distinct horizons (Fig. 18.8): the layering exhibited by some of the pottery designs is not apparent in any given section as the contexts they appear in are few and too thinly spread to allow a flat plane

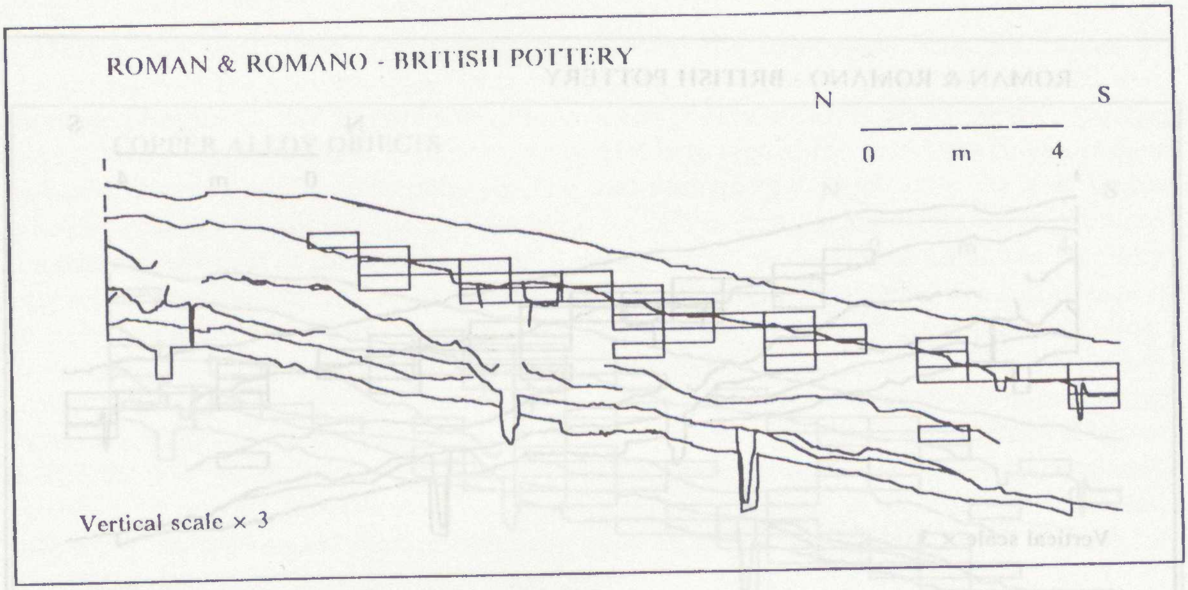

Figure 18.3: View through contexts with Roman and Romano-British pottery

(section) to intersect with a sufficient number of them to show the sloping horizon they are restricted to (Fig. 18.9).

Of course, in other cases, the distribution of the items under study may be so dense that the only way of examining the shapes of the distributions is to compute such a series of sectional views. It transpired that a novel version of this technique proved very illuminating in the analysis of the structure of the extensive, and densely packed, faunal assemblages within the midden.

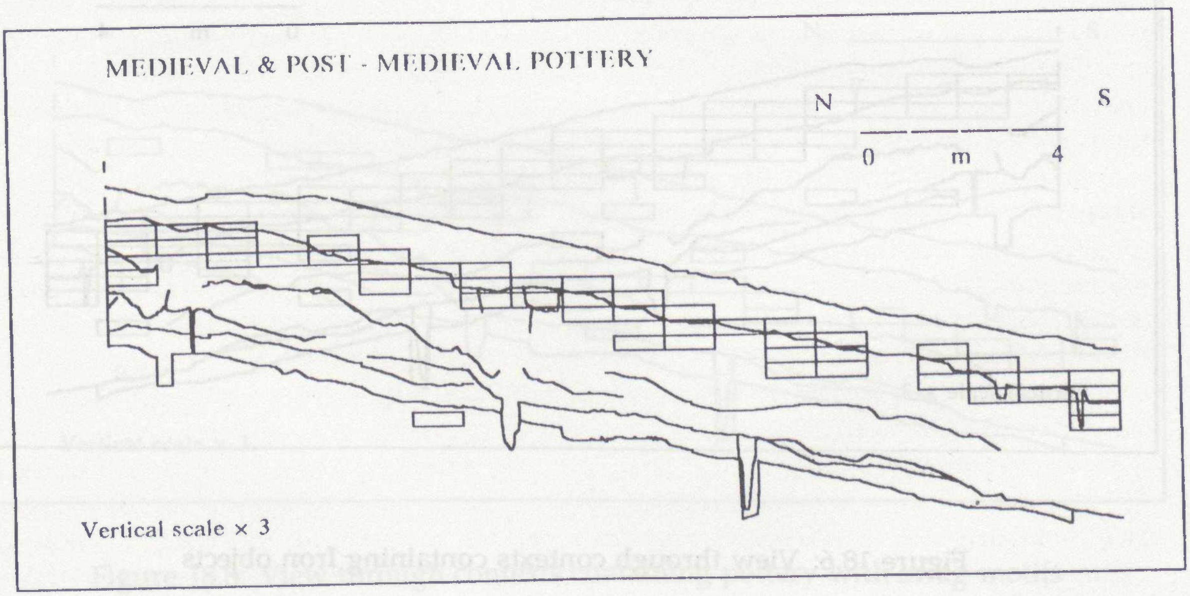

Figure 18.4: Plan showing contexts with Medieval and Post-medieval pottery

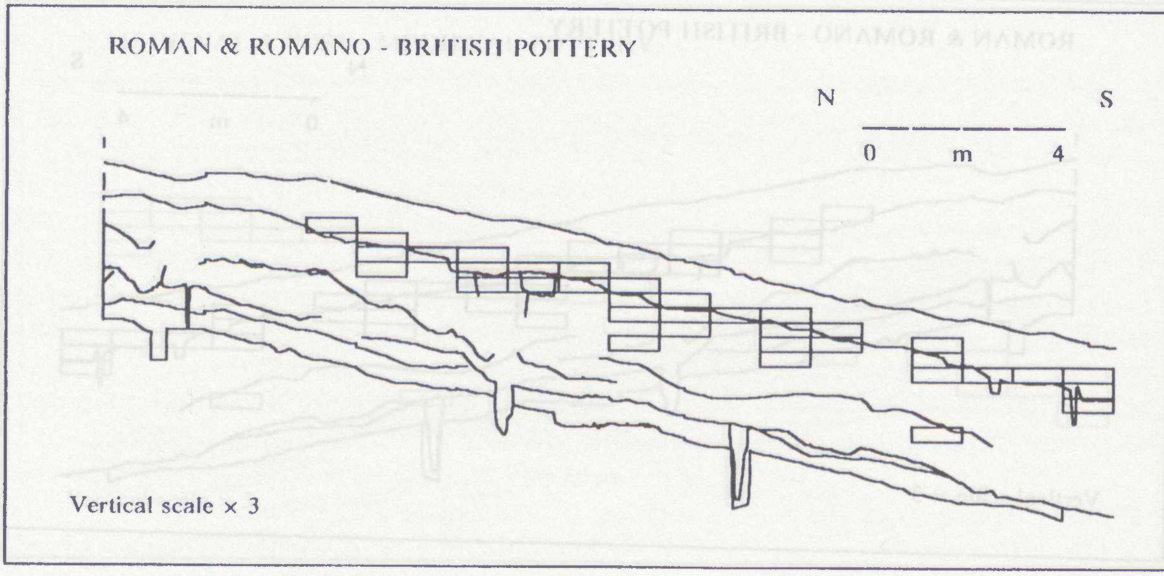

Figure 18.5: Plan showing contexts with Roman and Romano-British pottery

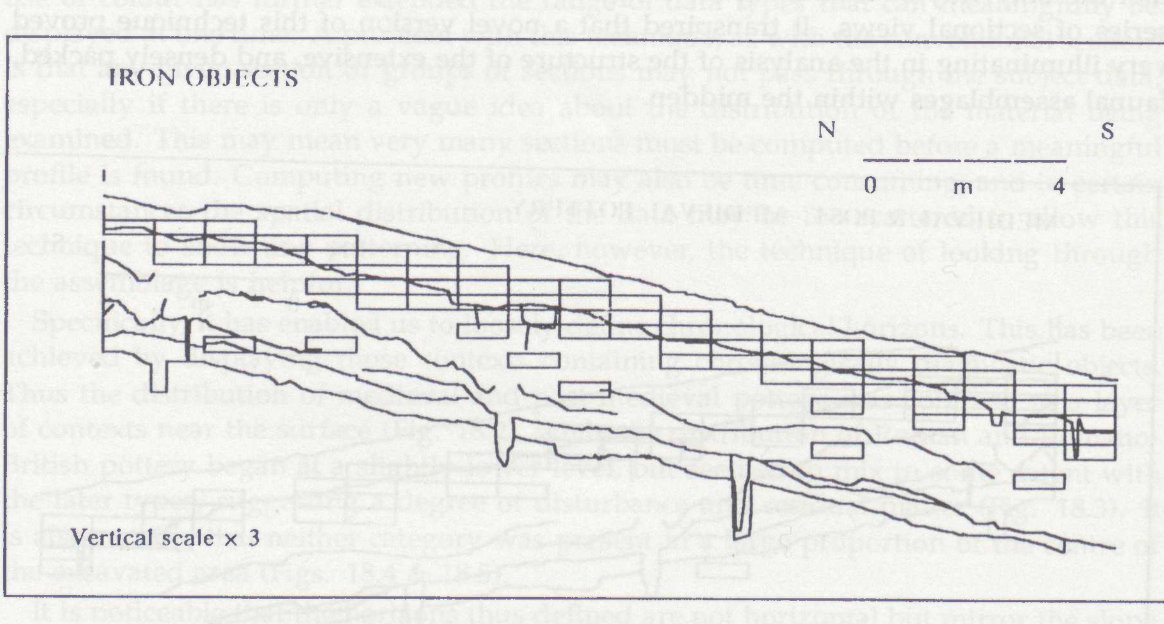

Figure 18.6: View through contexts containing Iron objects

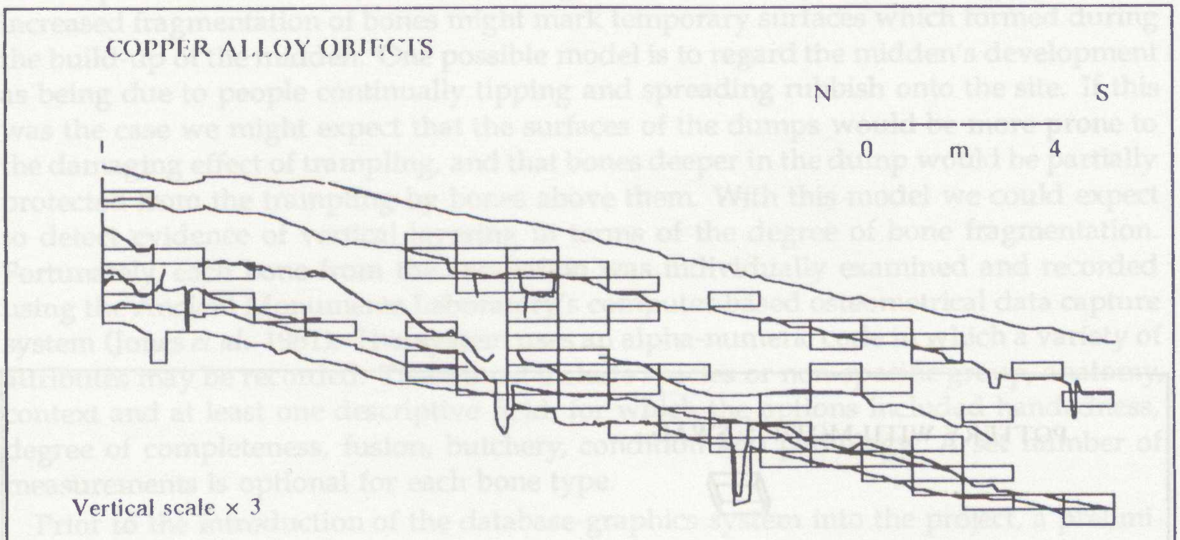

Figure 18.7: View through contexts containing Copper alloy objects

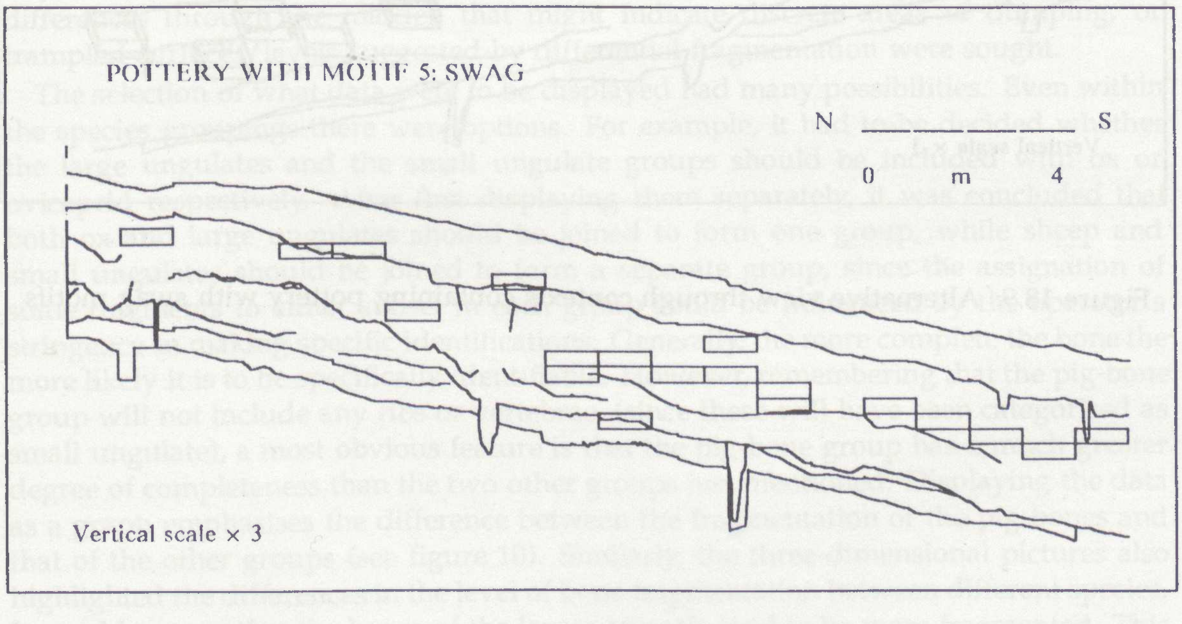

Figure 18.8: View through contexts containing pottery with swag motifs

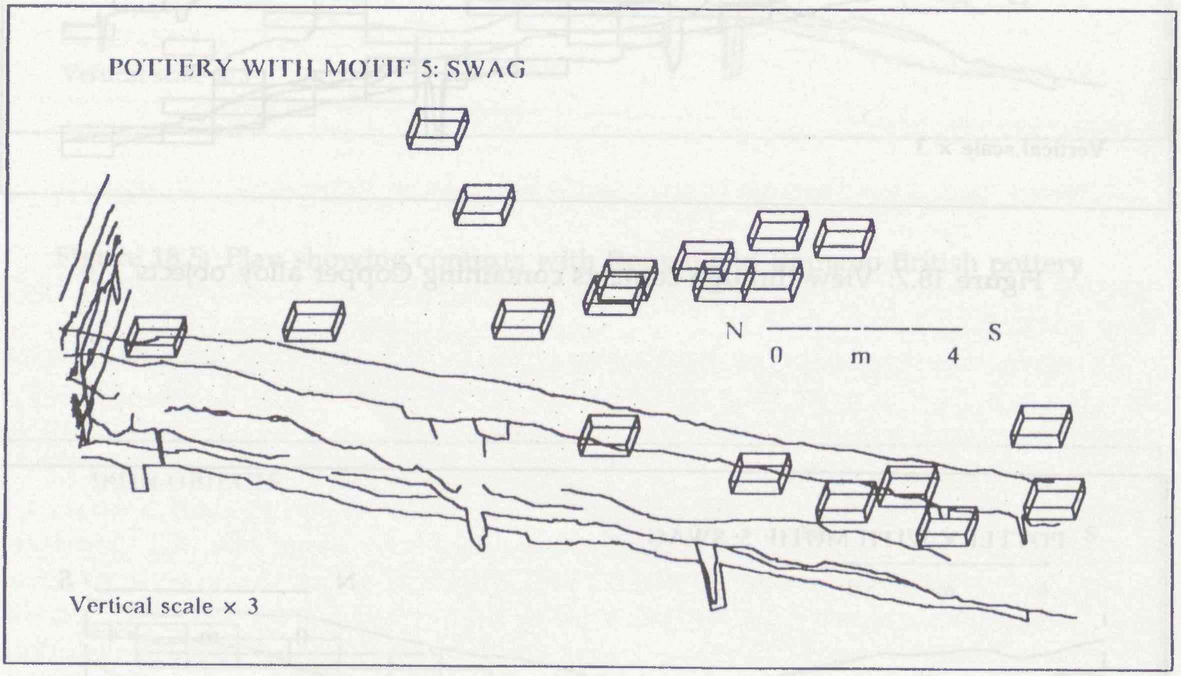

Figure 18.9: Alternative view through contexts containing pottery with swag motifs

During the excavation, the excavators received the impression that the bones in certain parts of the midden were more fragmented than others. It was conjectured that increased fragmentation of bones might mark temporary surfaces which formed during the build-up of the midden. One possible model is to regard the midden's development as being due to people continually tipping and spreading rubbish onto the site. If this was the case we might expect that the surfaces of the dumps would be more prone to the damaging effect of trampling, and that bones deeper in the dump would be partially protected from the trampling by bones above them. With this model we could expect to detect evidence of vertical layering in terms of the degree of bone fragmentation. Fortunately, each bone from the excavation was individually examined and recorded using the Ancient Monuments Laboratory's computer-based osteometrical data capture system (Jones *et al.* 1981). This system uses an alpha-numeric code in which a variety of attributes may be recorded. These must include species or non-specific group, anatomy, context and at least one descriptive field, for which the options included handedness, degree of completeness, fusion, butchery, condition and pathology. A set number of measurements is optional for each bone type.

Prior to the introduction of the database-graphics system into the project, a preliminary investigation of the bones by both species and anatomy suggested broad changes from the base to the top of the midden. A compensatory factor, devised by Levitan (in preparation), was subsequently applied to ensure that changes in the degree of fragmentation were not exaggerating changes in distribution. These changes include a general decrease in all body parts for cattle and large ungulates from the bottom to the top of the midden but, conversely, an increase in ovicaprids and small ungulates. Beyond interpreting possible changes in animal husbandry, *i.e.* an increase in the importance of sheep over cattle (the proportion of pigs remains relatively constant), differences through the midden that might indicate discrete areas of dumping, or trampled surfaces/levels suggested by differential fragmentation were sought.

The selection of what data were to be displayed had many possibilities. Even within the species groupings there were options. For example, it had to be decided whether the large ungulates and the small ungulate groups should be included with ox or ovicaprid respectively. After first displaying them separately, it was concluded that both ox and large ungulates should be joined to form one group, while sheep and small ungulates should be joined to form a separate group, since the assignation of some fragments to either subset in each group could be influenced by the operator's stringency in making specific identifications. Generally, the more complete the bone the more likely it is to be specifically identifiable. However, remembering that the pig-bone group will not include any ribs or vertebrae, (since these will have been categorised as small ungulate), a most obvious feature is that the pig-bone group has a much greater degree of completeness than the two other groups just mentioned. Displaying the data as a graph emphasises the difference between the fragmentation of the pig bones and that of the other groups (see figure 10). Similarly, the three-dimensional pictures also highlighted the differences in the level of bone fragmentation between different species. It would appear that the bones of the larger animals tend to be more fragmented. This might simply be a function of size; being larger, the are more likely to incur a breakage.

Attempts were subsequently made, using the graphics-database combination, to determine whether the level of bone fragmentation did in fact point to banding of the type indicated above and, if so, whether there were breaks consistent with the horizons

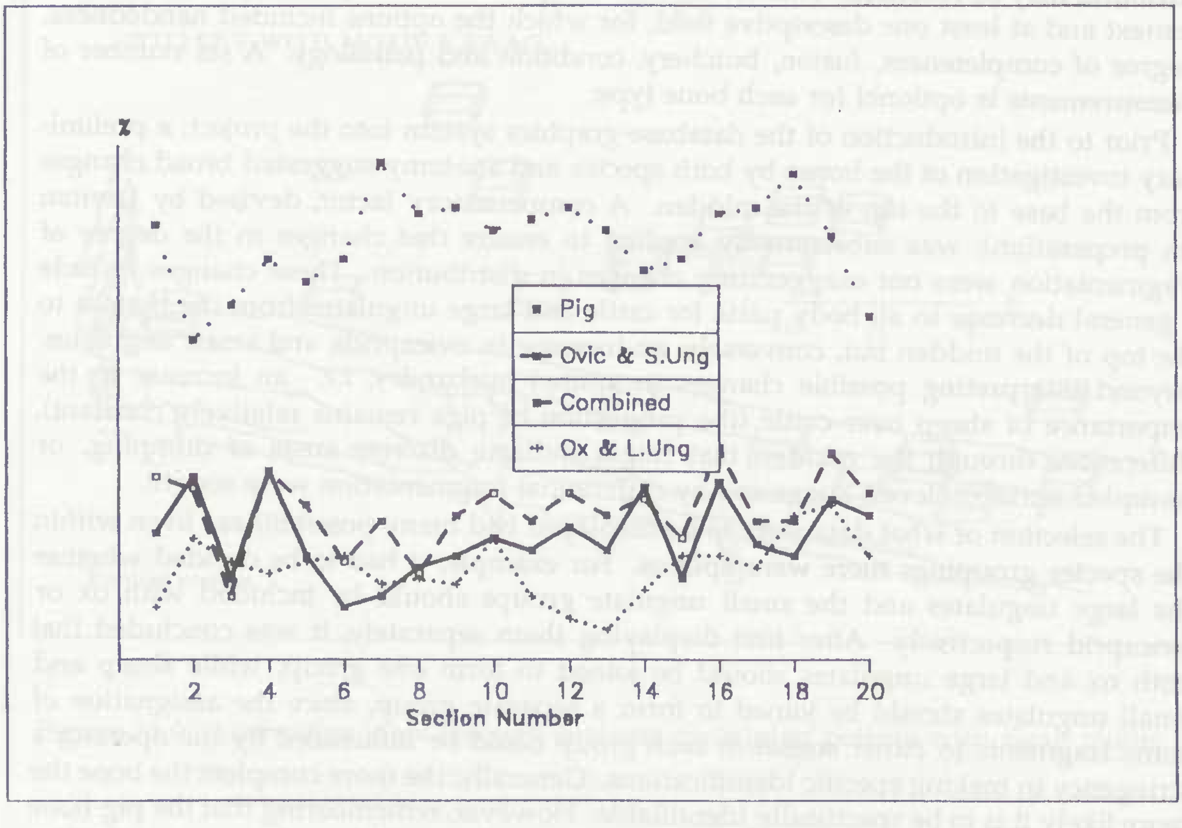

Figure 18.10: Proportion of Bones fragmented to <25% of original whole (by N-S Sect)

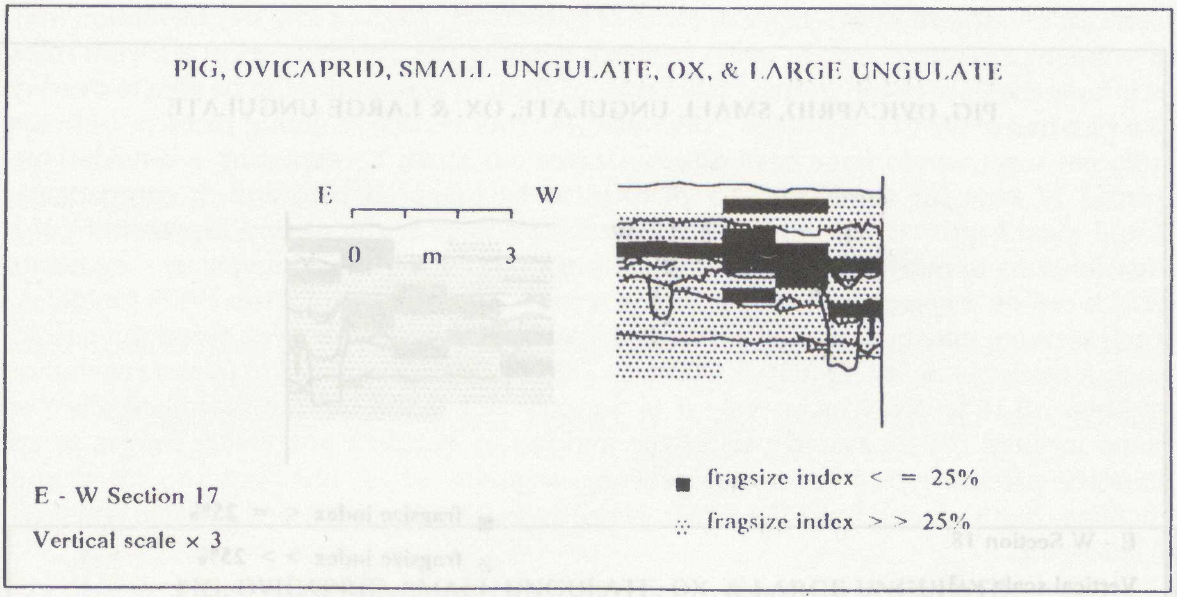

Figure 18.11: E-W section (17) from N showing banding in bone fragmentation.

defined on the basis of the distribution of the chronologically diagnostic material, or whether they represent more localised dumping patterns.

The method adopted to begin exploring these questions was to group each column of one-metre square, and ten-centimetre deep, arbitrary contexts into parallel rows running either north-south or east-west. Next, the bones from these contexts were selected from the data base, on a species basis or anatomy basis for instance. Colour codes were applied to indicate the specific attributes under study. Each row of contexts could then displayed with individual contexts appearing as a separate blocks of different colours.

Searches through the midden for evidence of clustering of associated anatomies, which have provided some information on butchery practises, were inconclusive. Different anatomies were grouped by species, limb and individual bone types, and at a greater level of detail particular aspects of the descriptive fields, such as fusion, condition, etc. were selected separately. It was hoped that by grouping associated anatomies relating to meat quality within species groups, (e.g. scapula, humerus, pelvis and femur, ribs and vertebrae, skull and mandibles, radius, ulna and tibia, phalanges, metacarpal and metatarsal), it might be possible to locate specific dumps of joints or parts of carcasses. However, as the initial displays appeared entirely random and did not suggest any groupings, this approach was abandoned as unproductive.

Initially, seven colours were used to indicate the different fragmentation sizes. Although some sizes did appear more dominant in certain bands, these were not felt to be meaningful. Displaying the data more simplistically on a basis of two fragmentation groupings produced easier-to-understand pictures and did appear to bring out several distinctive bands when the degree of fragmentation was investigated on a species-group basis (e.g. Figs. 18.11 to 18.14).

For any particular row, the presence of *clusters* or *bands* of high or low fragmentation is not always entirely convincing, but when the distributions are echoed in the adjacent

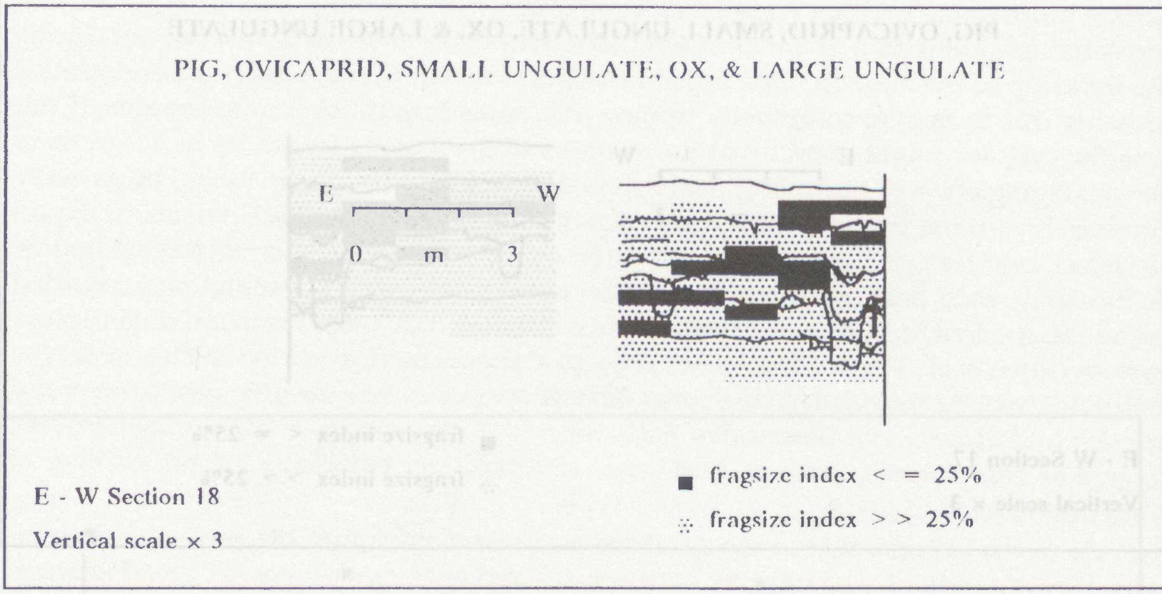

Figure 18.12: E-W section (18) from N showing banding in bone fragmentation.

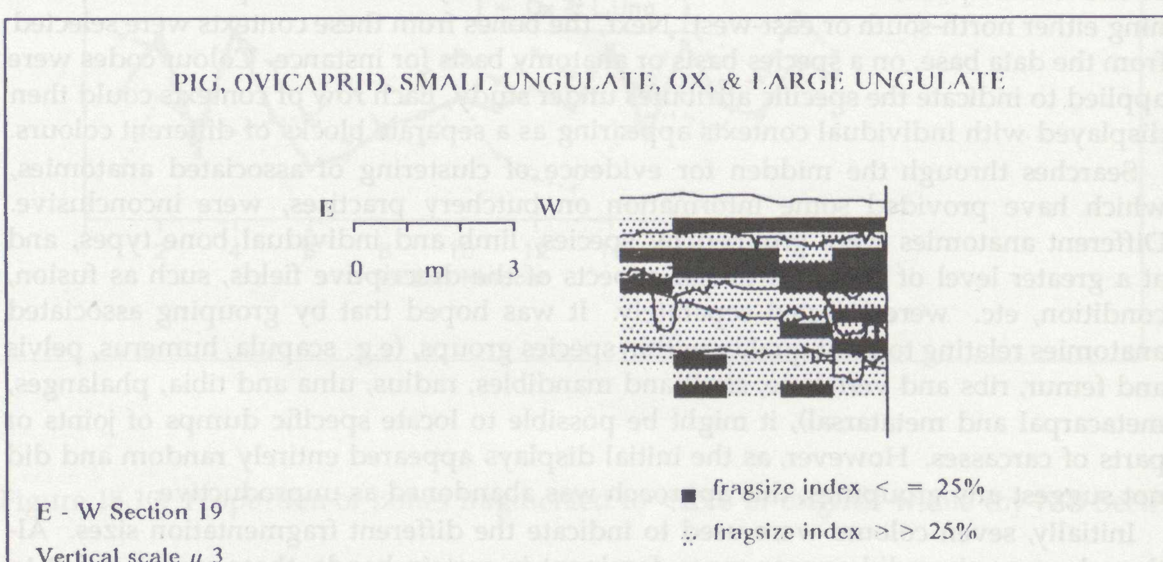

Figure 18.13: E-W section (19) from N showing banding in bone fragmentation.

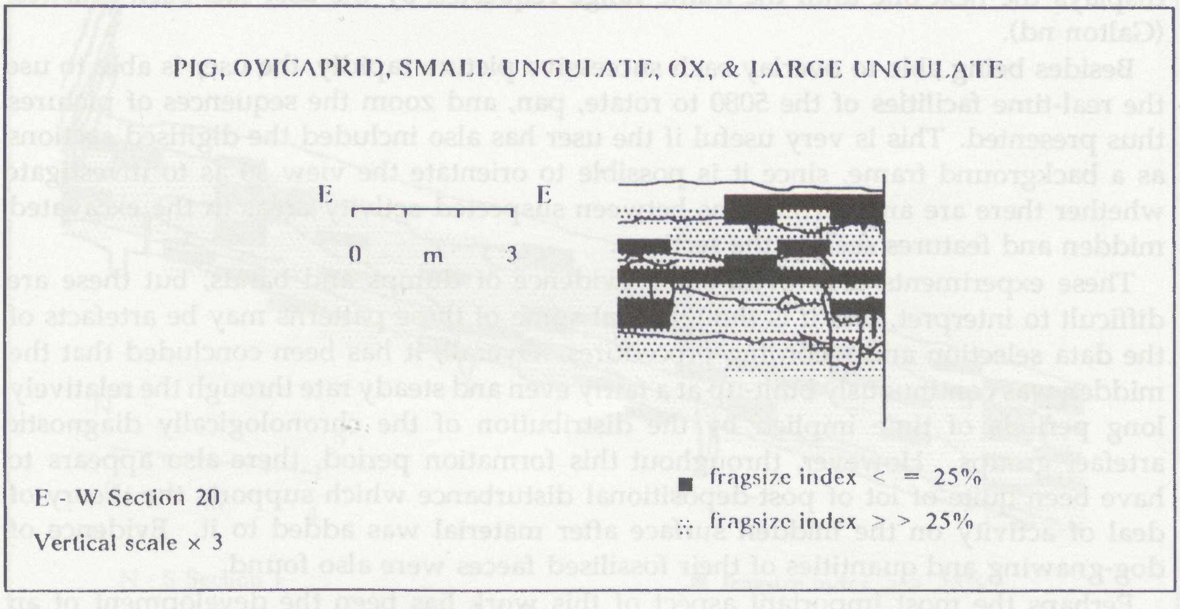

Figure 18.14: E-W section (20) from N showing banding in bone fragmentation.

rows an identification of a layer or locus of particularly high or low fragmentation may become more convincing. In order to check whether the distributions in adjacent rows reinforce the impression derived from each of them, it is important to be able to overlay the pictures of the distributions. Previously, the process of overlaying pictures from the adjacent rows would have been quite a drawn out affair, necessitating a considerable period of time for the hard-copy pictures to be prepared before such comparisons could even begin. However, in WGS, the components of the picture representing the distributions in individual rows of contexts can be stored as separate picture segments, which can be displayed or hidden on the screen using the appropriate WGS functions. Each segment may be regarded as the equivalent of a frame in a movie. Several hundred such frames can be manipulated within an IBM 5080 work station (10 *frames* are shown in Figs. 18.15 to 18.24 inclusive). It is possible to produce an animated sequence of steps through the excavated part of the midden by building successive frames in an iterative process which hides any previous segment which has been processed and displays the next one until the frame range requested by the user has been satisfied (Galton nd).

Besides being able to overlay each successive picture rapidly, the user is able to use the real-time facilities of the 5080 to rotate, pan, and zoom the sequences of pictures thus presented. This is very useful if the user has also included the digitised sections as a background frame, since it is possible to orientate the view so as to investigate whether there are any correlations between suspected activity areas in the excavated midden and features seen in the sections.

These experiments did reveal some evidence of dumps and bands, but these are difficult to interpret, and it is thought that some of these patterns may be artefacts of the data selection and recording procedures. Overall, it has been concluded that the midden was continuously built-up at a fairly even and steady rate through the relatively long periods of time implied by the distribution of the chronologically diagnostic artefact groups. However, throughout this formation period, there also appears to have been quite of lot of post-depositional disturbance which supports the theory of deal of activity on the midden surface after material was added to it. Evidence of dog-gnawing and quantities of their fossilised faeces were also found.

Perhaps the most important aspect of this work has been the development of an interesting and very powerful new way of looking at buried data. Clearly, the analytical power of what may be termed an interactive movie is great. The other positive effect, of course, has been to focus attention back on the strategies used in the data collection and recording procedures.

18.2 An early Bronze Age settlement site, St. Veit-Klinglberg, Austria

Similar methodologies are being applied in the current investigation of the site of St. Veit-Klinglberg[1], about 70 km south of Salzburg in the Austrian Alps. The site is a settlement of the later Early Bronze Age in one of the major European copper-mining

[1]Financial support for the Klinglberg project has been provided by the Department of Archaeology, University of Southampton, the University of Southampton Research Fund, the British Academy, the National Geographic Society, the Society of Antiquaries, and the local councils of St. Veit and St. Johann, to all of whom SJS is extremely grateful. Thanks must also go to Professor A. Lippert, University of Innsbruck, and Dr. F. Berg of the Austrian Bundesdenkmalamt for their assistance.

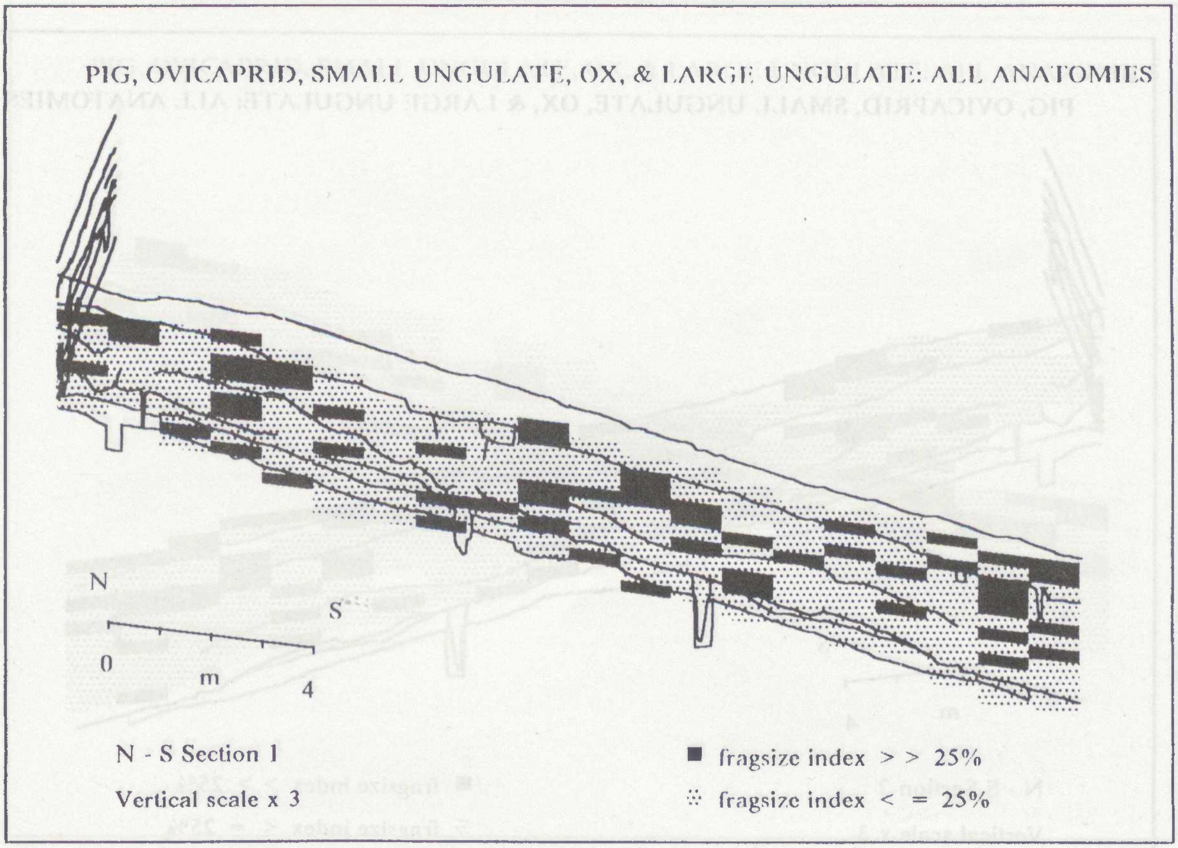

Figure 18.15: Changes in the level of fragmentation in N-S section 1

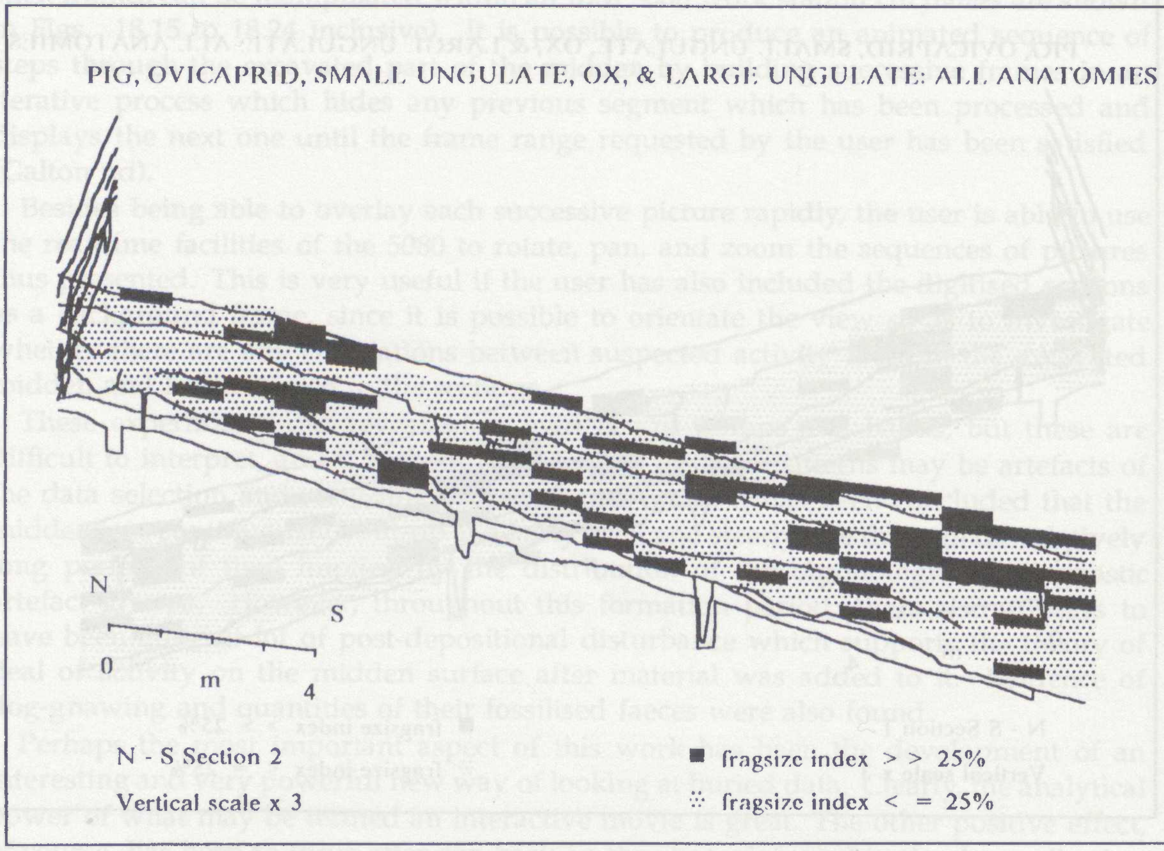

Figure 18.16: Changes in the level of fragmentation in N-S section 2

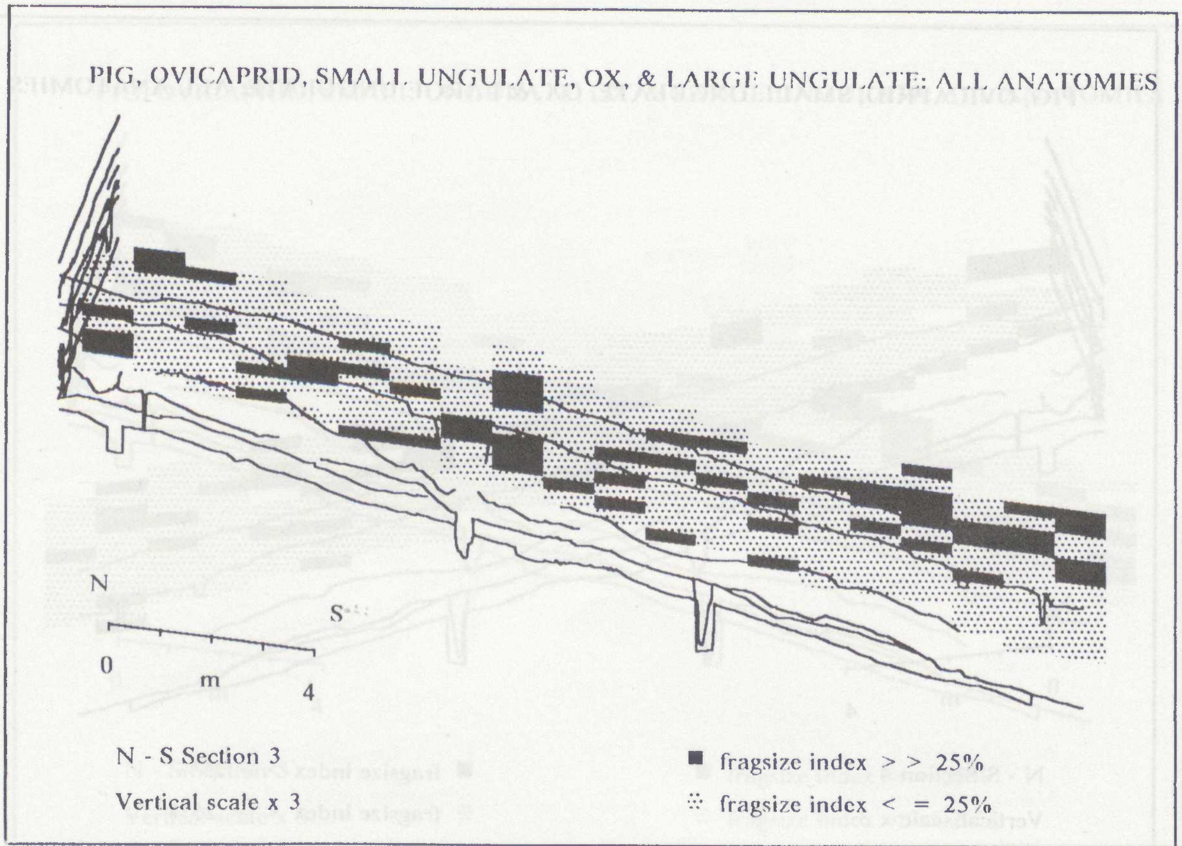

Figure 18.17: Changes in the level of fragmentation in N-S section 3

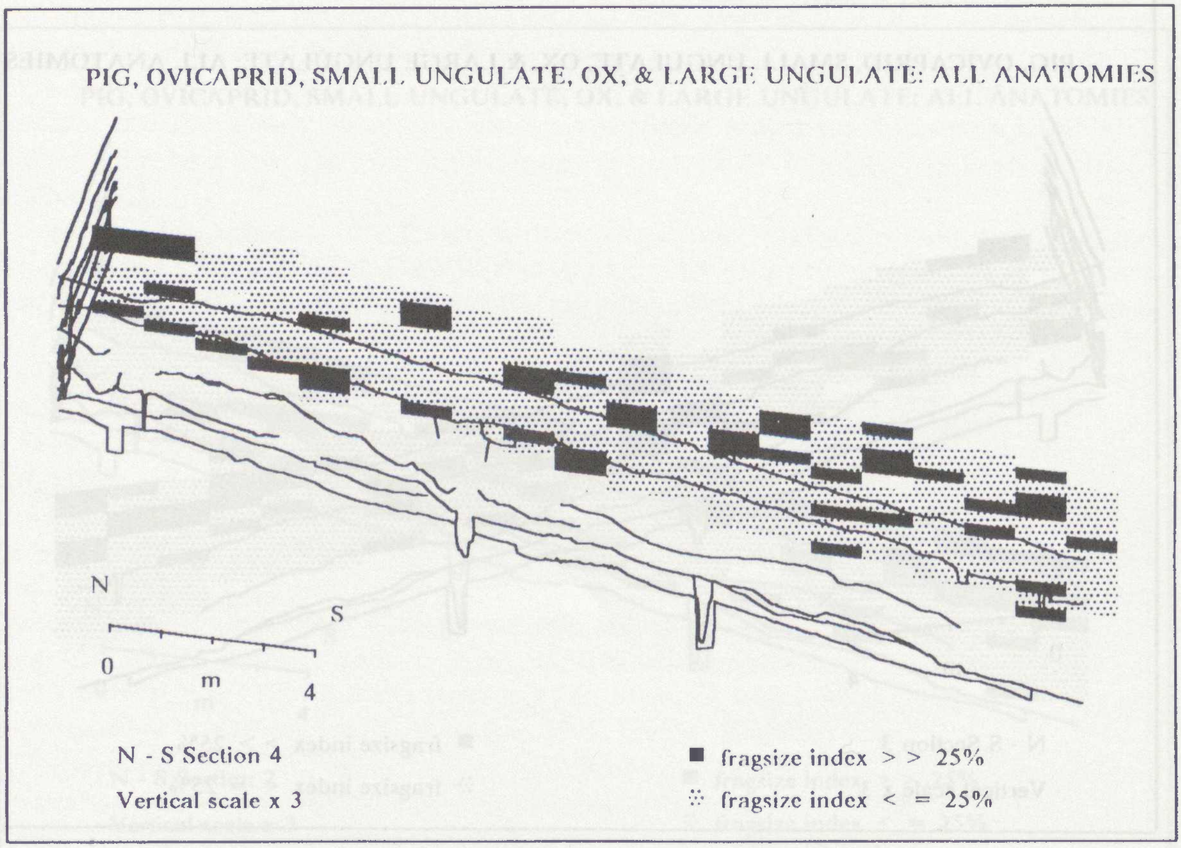

Figure 18.18: Changes in the level of fragmentation in N-S section 4

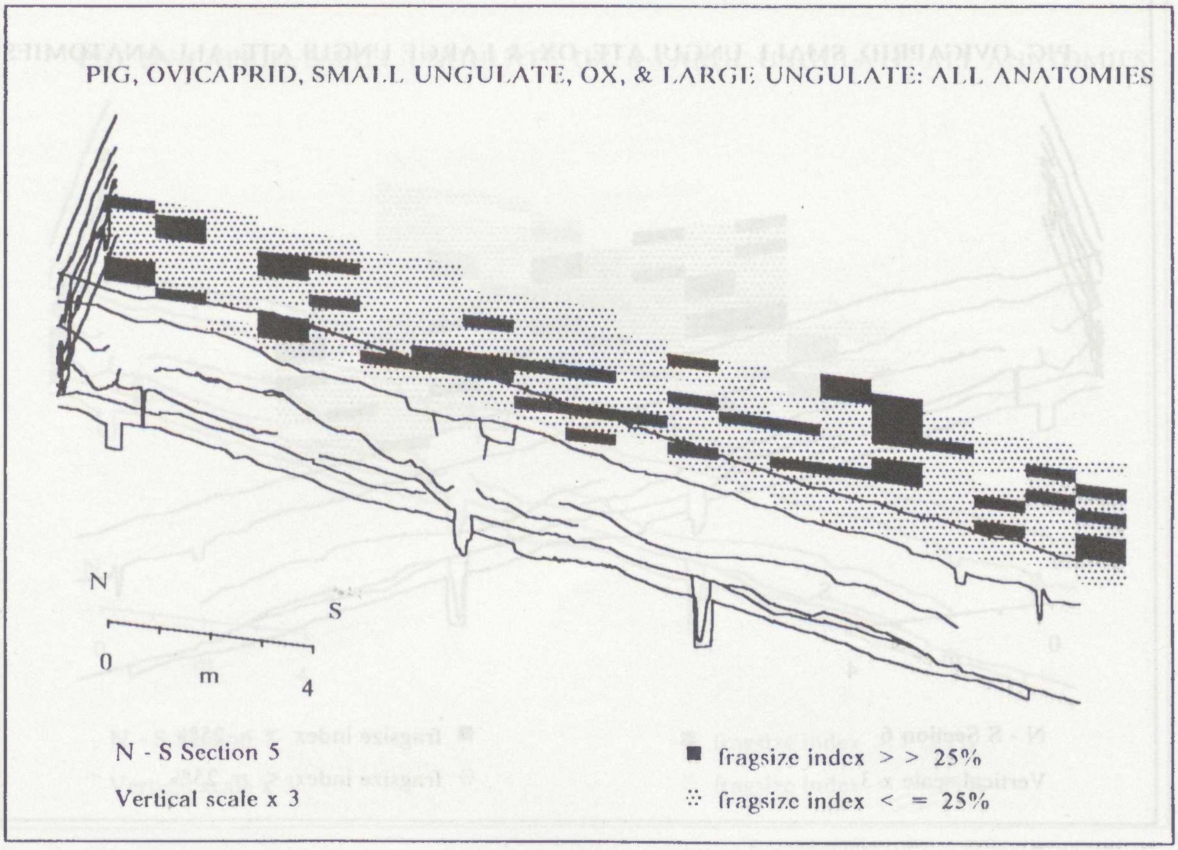

Figure 18.19: Changes in the level of fragmentation in N-S section 5

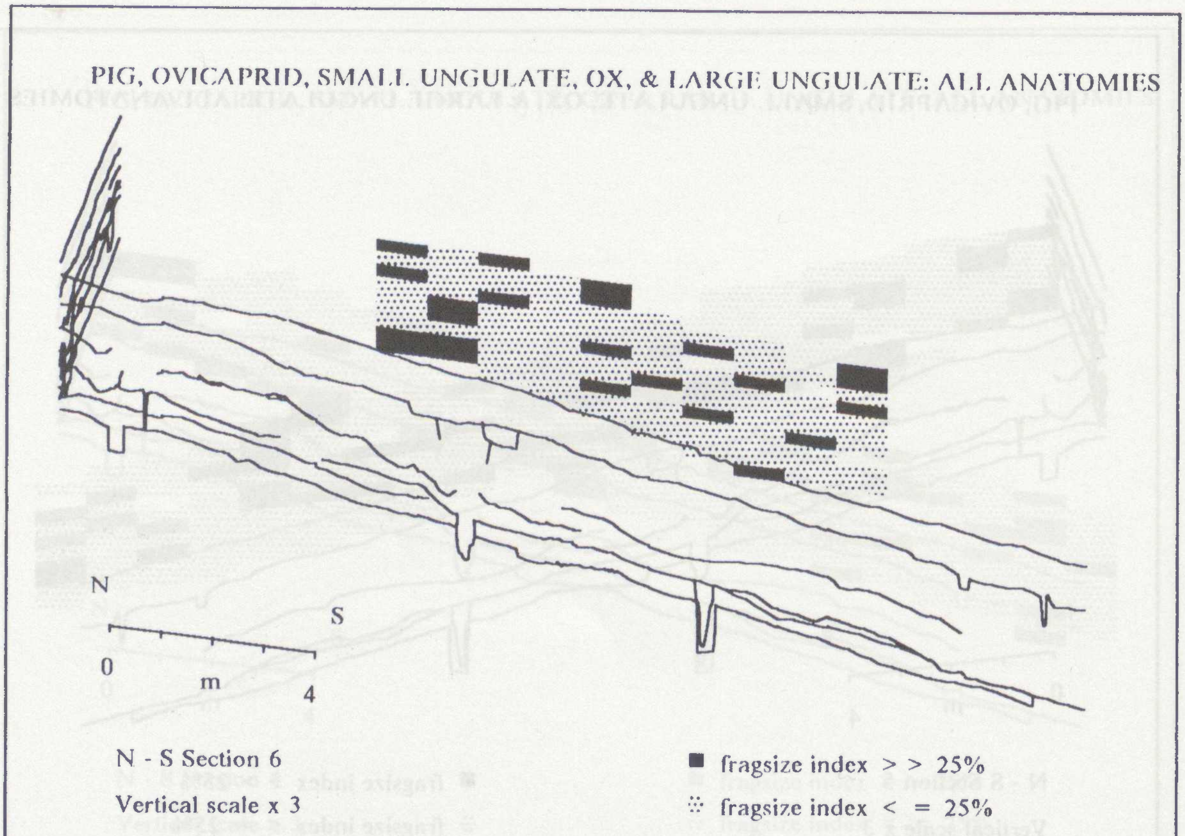

Figure 18.20: Changes in the level of fragmentation in N-S section 6

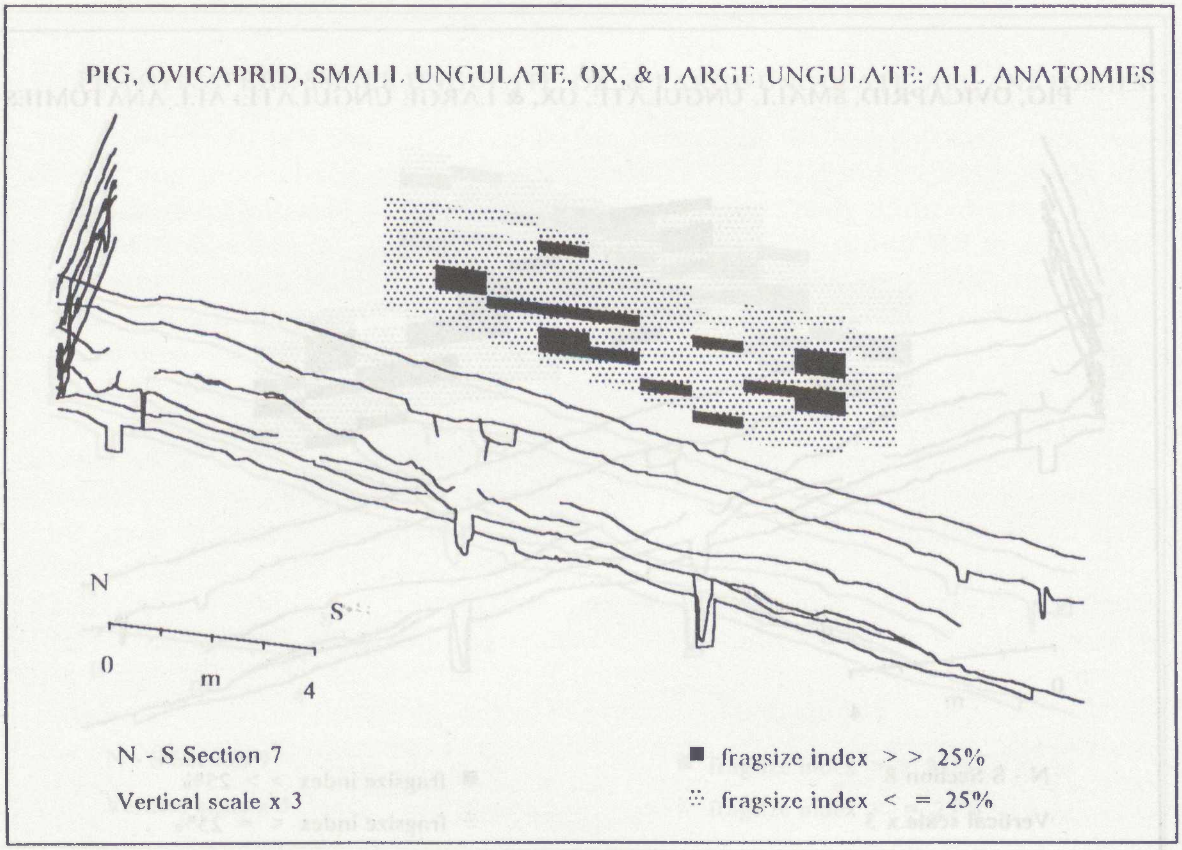

Figure 18.21: Changes in the level of fragmentation in N-S section 7

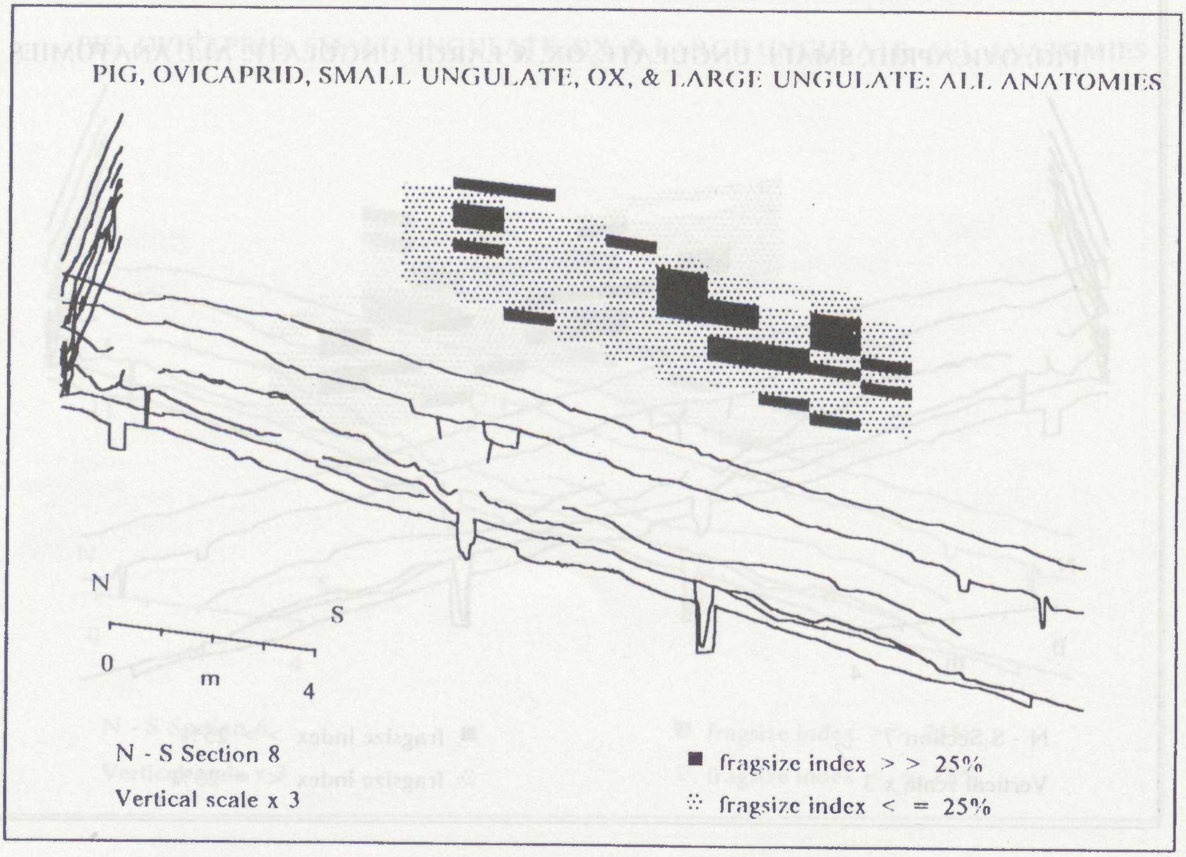

Figure 18.22: Changes in the level of fragmentation in N-S section 8

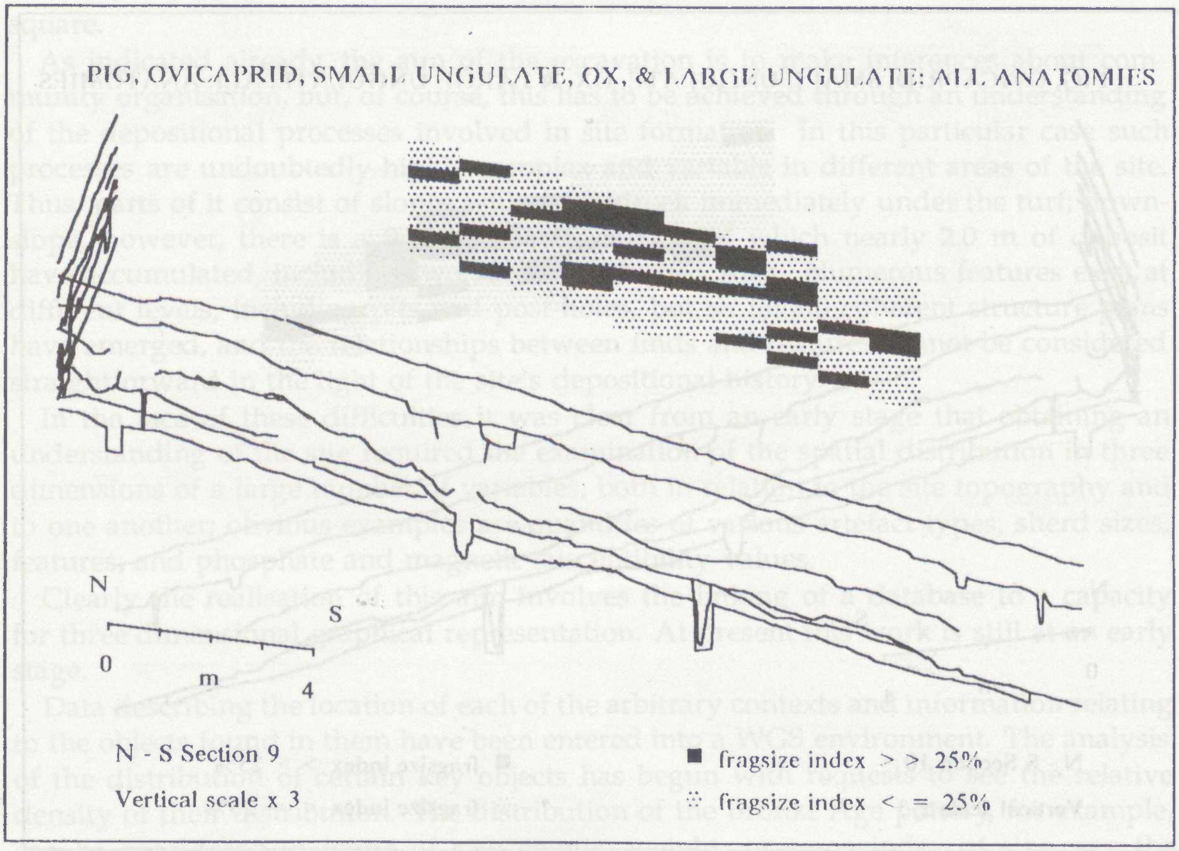

Figure 18.23: Changes in the level of fragmentation in N-S section 9

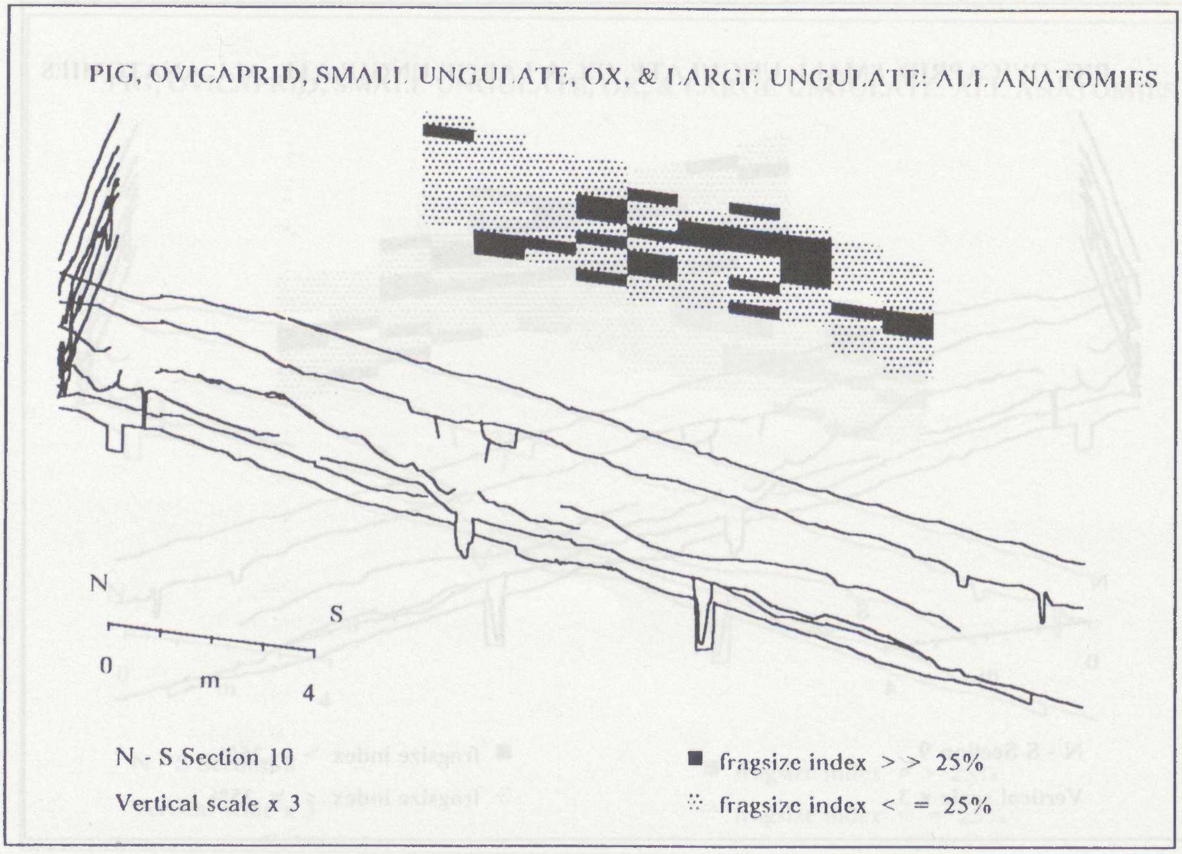

Figure 18.24: Changes in the level of fragmentation in N-S section 10

areas and the project's aim is to understand the organisation of a community involved in the production and distribution of one of the major exchanged commodities of Bronze Age Europe. It is the first such settlement to be excavated on any scale using modern methods.

The settlement is a characteristic location on a spur of the mountainside on the south-facing side of the valley of the river Salzach, about 50 m above the valley bottom. Area excavation techniques are being used in the main part of the site. In the upper levels, where stratigraphic layers are not detectable, the basic recording unit has been arbitrary spits 2.0 x 2.0 x 0.15m. Where layers and features have been found they have been excavated stratigraphically in the usual way, although in the case of layers of large extent spatial control has been maintained down to at least the level of the 2.0 x 2.0 m square.

As indicated already, the aim of the excavation is to make inferences about community organisation, but, of course, this has to be achieved through an understanding of the depositional processes involved in site formation. In this particular case such processes are undoubtedly highly complex and variable in different areas of the site. Thus, parts of it consist of slopes with the bedrock immediately under the turf; down-slope, however, there is a flattened shelf in parts of which nearly 2.0 m of deposit have accumulated, including two successive buried soils. Numerous features exist at different levels, including pits and post-holes, but so far no coherent structure plans have emerged, and the relationships between finds and features cannot be considered straightforward in the light of the site's depositional history.

In the face of these difficulties it was clear from an early stage that obtaining an understanding of the site required the examination of the spatial distribution in three dimensions of a large number of variables, both in relation to the site topography and to one another; obvious examples are quantities of various artefact types, sherd sizes, features, and phosphate and magnetic susceptibility values.

Clearly the realisation of this aim involves the linking of a database to a capacity for three dimensional graphical representation. At present this work is still at an early stage.

Data describing the location of each of the arbitrary contexts and information relating to the objects found in them have been entered into a WGS environment. The analysis of the distribution of certain key objects has begun with requests to see the relative density of their distribution. The distribution of the Bronze Age pottery, for example, can be considered in terms of raw counts, weight, or some index of size, say the weight of pottery in a context, divided by the number of pieces, so that variation in fragmentation, and hence deposition patterns may be defined.

Spectral colour sequences can be used to show the relative values. The convention used by us is to regard red as representing the high and blue the low end of the range. An alternative is to use grey-scale or dot-density values.

The distributions of different categories of object can be compared. These can then be compared with other distributions, pottery with copper objects for example (see Figs. 18.25 to 18.32). However, the scope of the graphics work planned is more ambitious. It is intended that individual archaeological features and layers will also be incorporated into the system.

As will be clear already, these graphical methods have the potential to revolutionise the process of analysing and interpreting the results of archaeological excavations, but

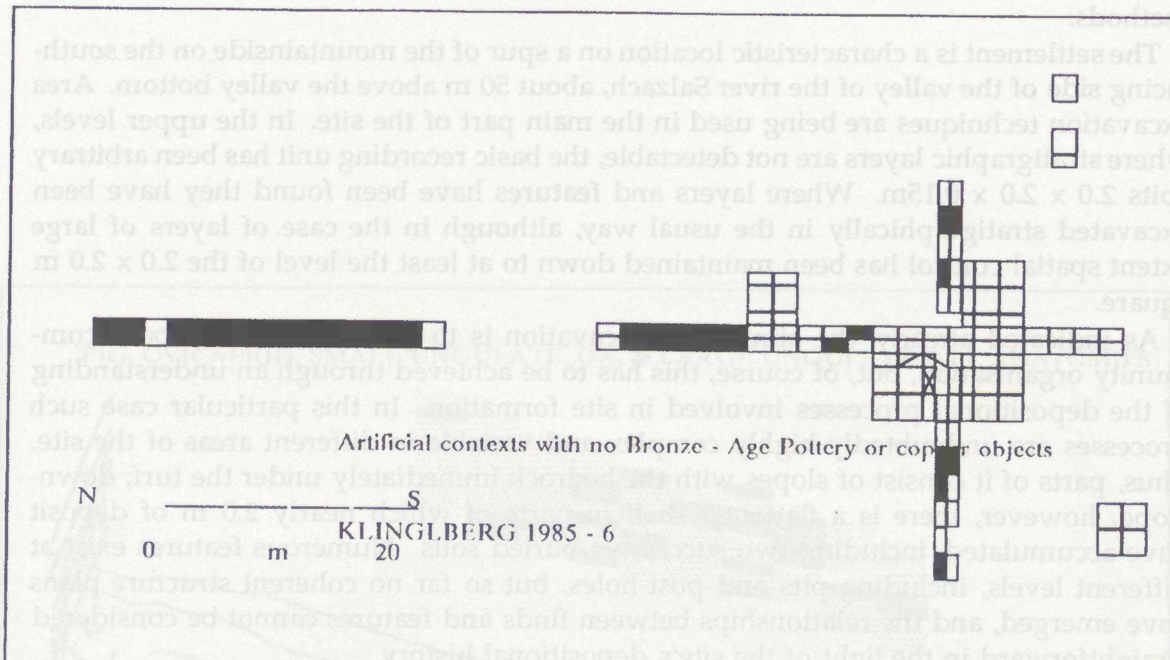

Figure 18.25: Distribution of contexts with neither Copper Objects nor Bronze-Age Pottery (Plan)

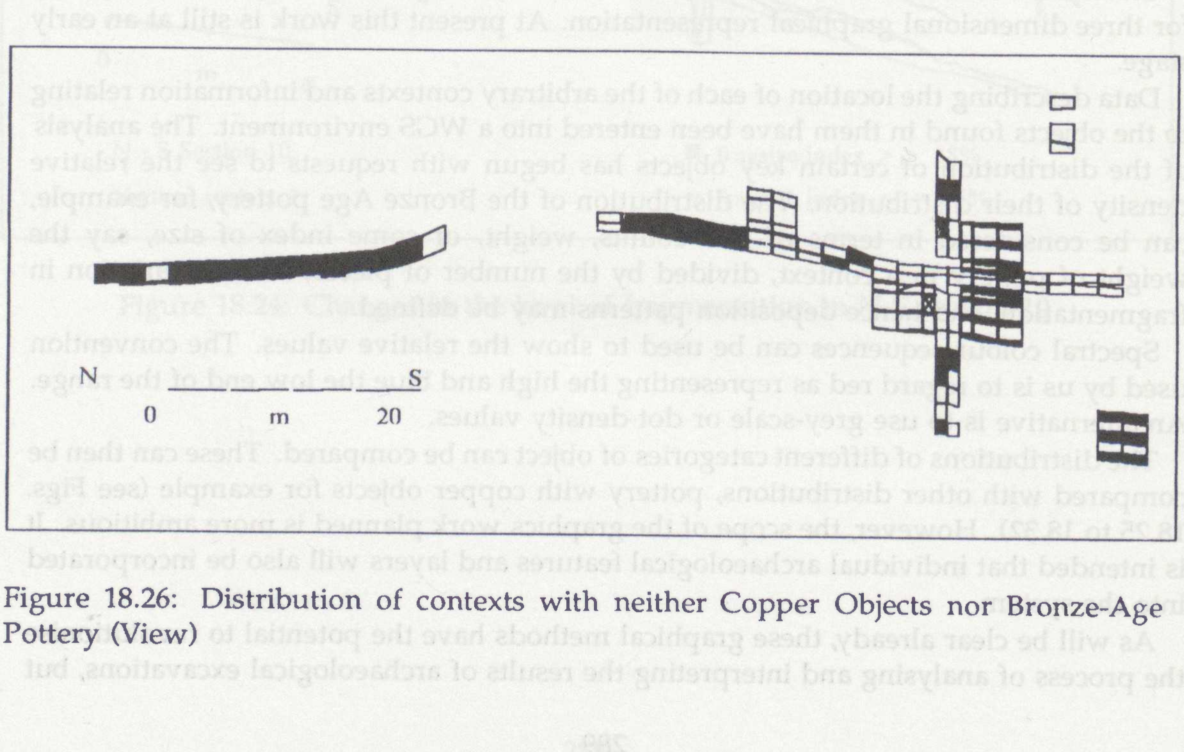

Figure 18.26: Distribution of contexts with neither Copper Objects nor Bronze-Age Pottery (View)

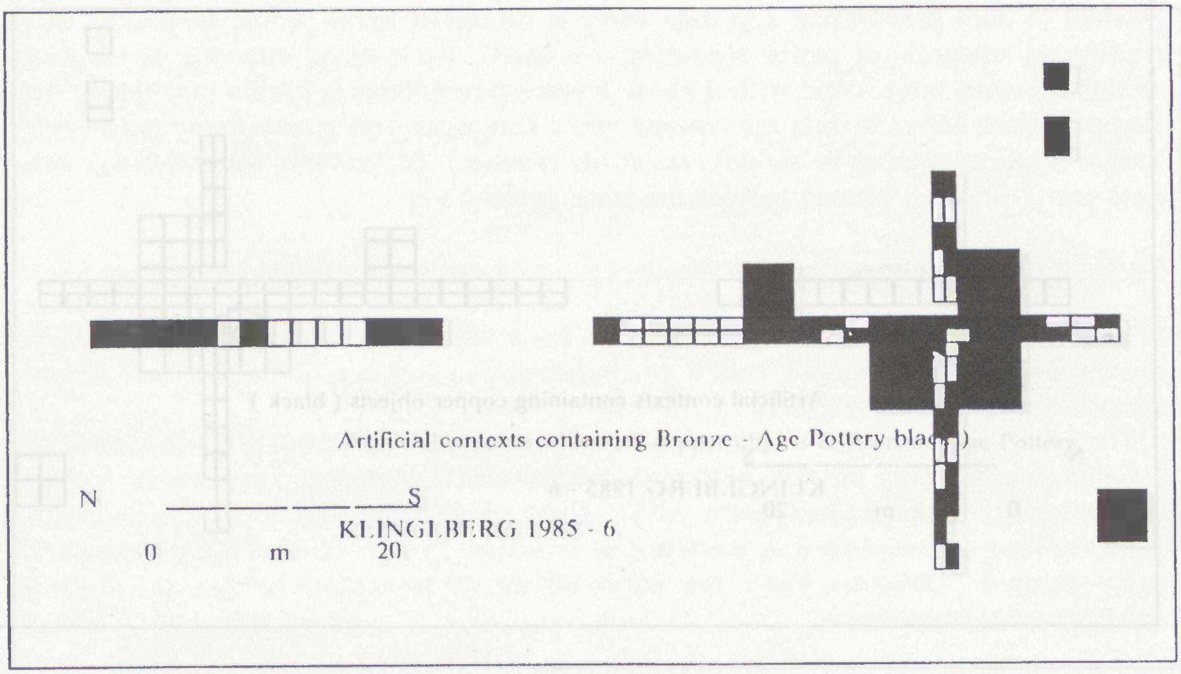

Figure 18.27: Distribution of contexts containing Bronze-Age Pottery, but no Copper Objects (Plan)

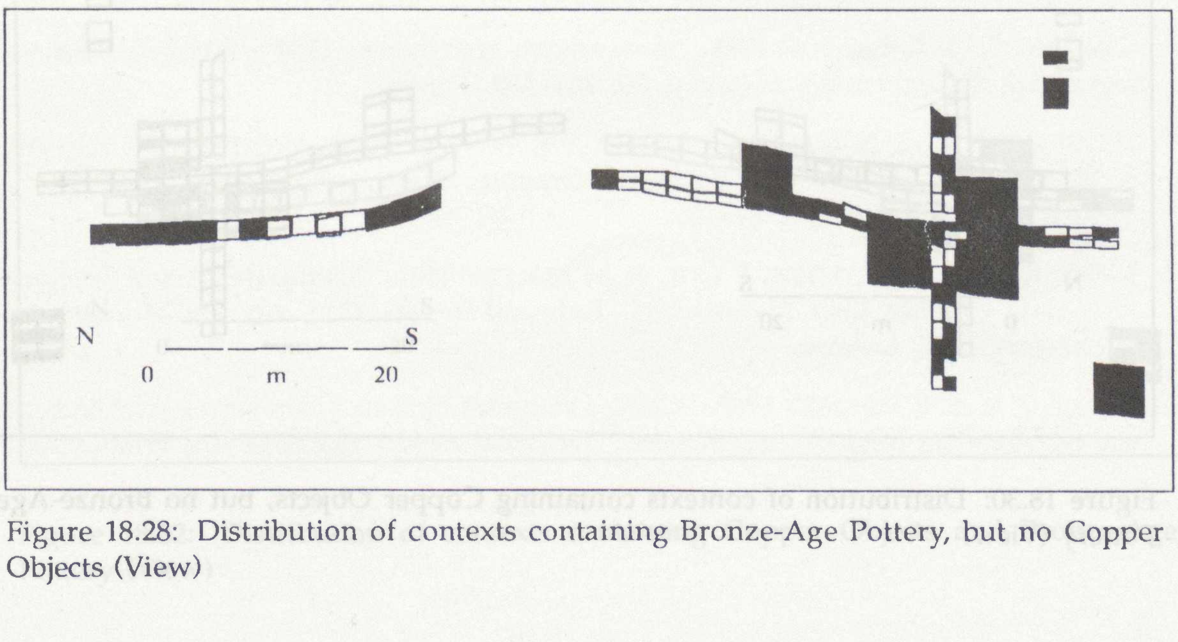

Figure 18.28: Distribution of contexts containing Bronze-Age Pottery, but no Copper Objects (View)

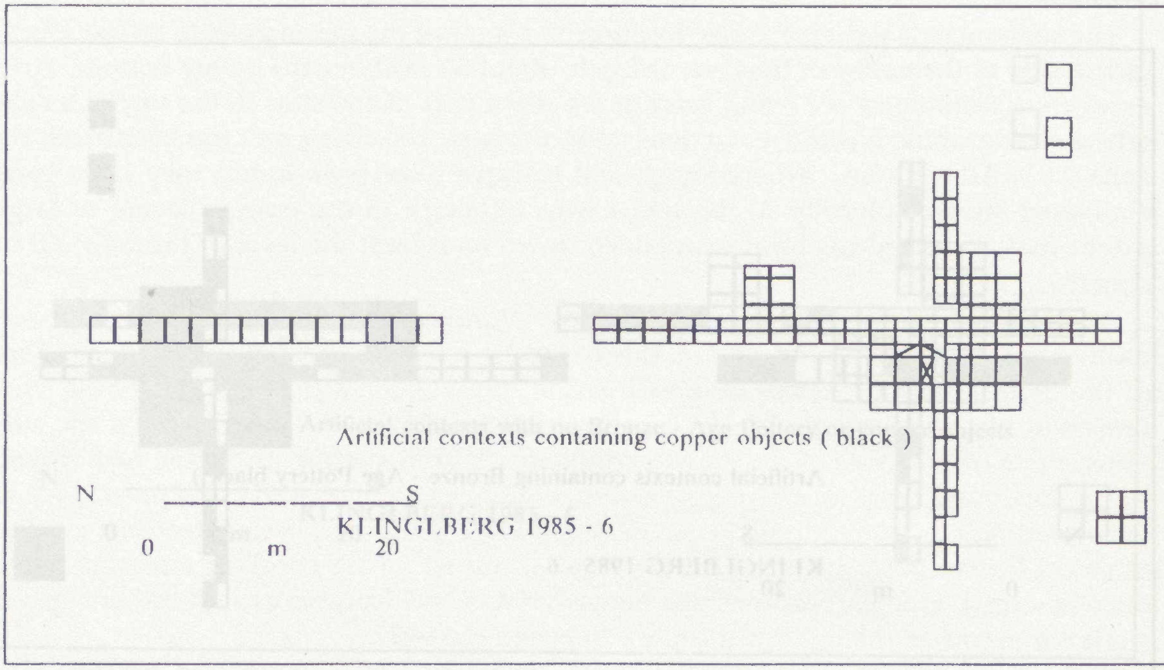

Figure 18.29: Distribution of contexts with containing Copper Objects, but no Bronze-Age Pottery (Plan)

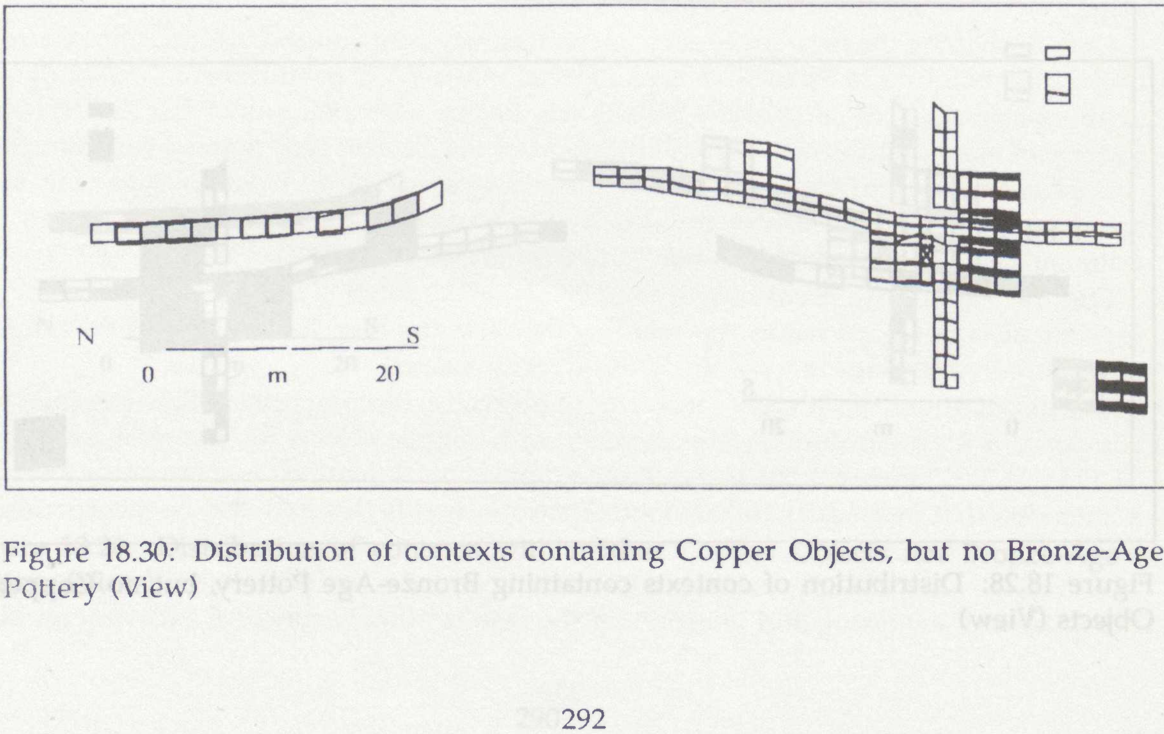

Figure 18.30: Distribution of contexts containing Copper Objects, but no Bronze-Age Pottery (View)

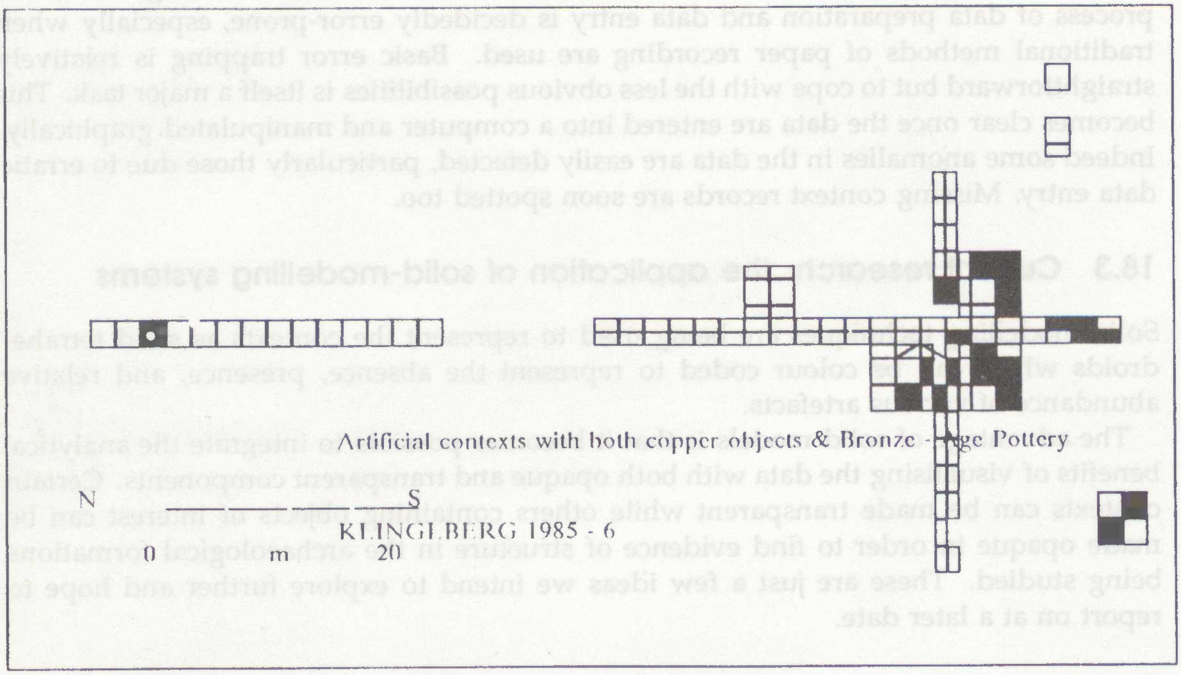

Figure 18.31: Distribution of contexts containing both Copper Objects and Bronze-Age Pottery (Plan)

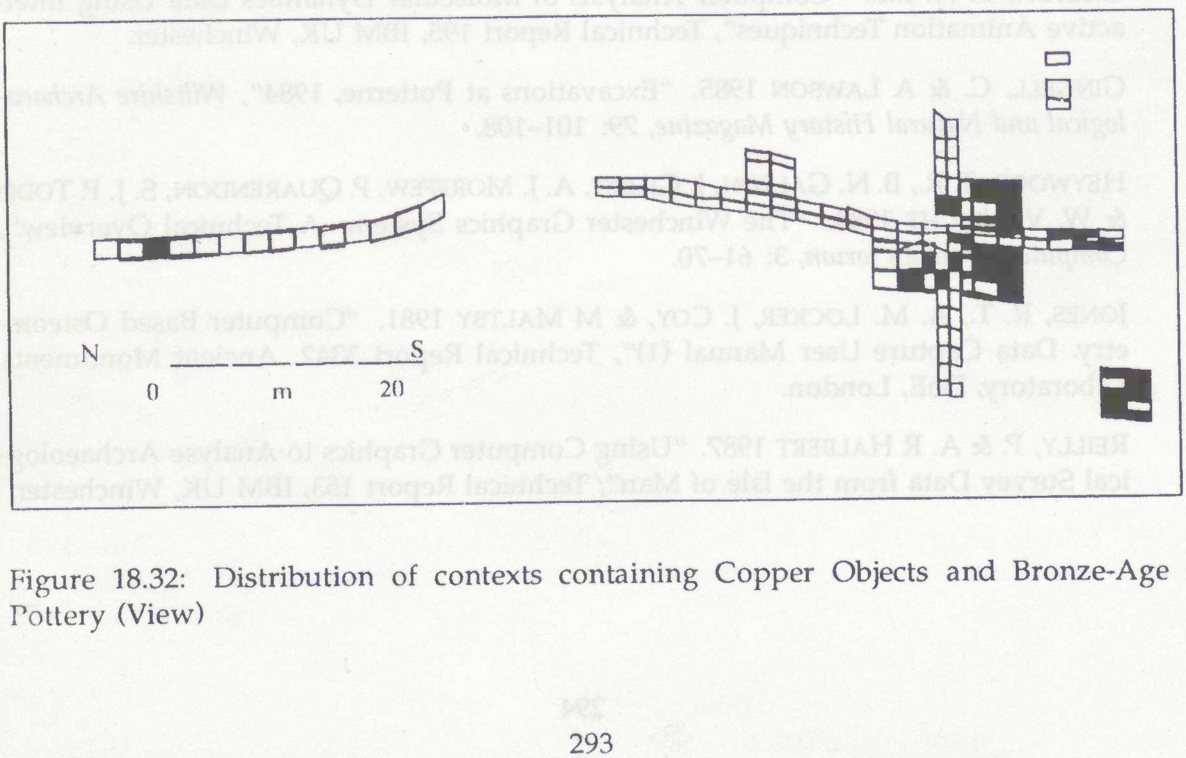

Figure 18.32: Distribution of contexts containing Copper Objects and Bronze-Age Pottery (View)

it is important to appreciate that they also make demands which must be taken into account at the stage of carrying out the work in the field and preparing the data for analysis. In a 3-D model every point in excavated space must be accounted for. Achieving adequate standards in the field is difficult and time-consuming, while the process of data preparation and data entry is decidedly error-prone, especially when traditional methods of paper recording are used. Basic error trapping is relatively straightforward but to cope with the less obvious possibilities is itself a major task. This becomes clear once the data are entered into a computer and manipulated graphically. Indeed some anomalies in the data are easily detected, particularly those due to erratic data entry. Missing context records are soon spotted too.

18.3 Current research: the application of solid-modelling systems

Solid-modelling techniques are being used to represent the contexts as solid tetrahedroids which can be colour coded to represent the absence, presence, and relative abundance of various artefacts.

The advantage of solid models is that it becomes possible to integrate the analytical benefits of visualising the data with both opaque and transparent components. Certain contexts can be made transparent while others containing objects of interest can be made opaque in order to find evidence of structure in the archaeological formations being studied. These are just a few ideas we intend to explore further and hope to report on at a later date.

References

COLLEY, S. M., S. J. P. TODD, & N. R CAMPLING 1988. "Three-Dimensional Computer Graphics for Archaeological Data Exploration: An Example from Saxon Southampton", *Journal of Archaeological Science*, 15: 99–106.

GALTON, B. N. n.d. "Computer Analysis of Molecular Dynamics Data Using Interactive Animation Techniques", Technical Report 195, IBM UK, Winchester.

GINGELL, C. & A LAWSON 1985. "Excavations at Potterne, 1984", *Wiltshire Archaeological and Natural History Magazine*, 79: 101–108.

HEYWOOD, T. R., B. N. GALTON, J. GILLET, A. J. MORFFEW, P. QUARENDON, S. J. P. TODD, & W. V WRIGHT 1984. "The Winchester Graphics System: A Technical Overview", *Computer Graphics forum*, 3: 61–70.

JONES, R. T., A. M. LOCKER, J. COY, & M MALTBY 1981. "Computer Based Osteometry. Data Capture User Manual (1)", Technical Report 3342, Ancient Monuments Laboratory, DoE, London.

REILLY, P. & A. R HALBERT 1987. "Using Computer Graphics to Analyse Archaeological Survey Data from the Isle of Man", Technical Report 153, IBM UK, Winchester.