GEOGRAPHIC INFORMATION SYSTEM DESIGN FOR ARCHAEOLOGICAL SITE INFORMATION RETRIEVAL

Trevor M. Harris B.A., Ph.D
Countryside Research Unit,
Brighton Polytechnic,
Falmer,
Brighton BN1 9PH

INTRODUCTION

This paper examines the application and implementation of a computer-based information system designed to record, store, retrieve and analyse archaeological site data for the Brighton area based upon techniques currently being developed in the field of Geographic Information Systems (GIS). The paper discusses the need for spatial data handling capabilities in archaeology at the regional level to satisfy archival, educational and research purposes, and importantly, as a decision-making tool in the management of the historic environment.

GEOGRAPHIC INFORMATION SYSTEMS

It is perhaps no coincidence that at the same time the archaeological profession is looking at ways to implement and standardise computer applications in archaeology the House of Commons should, following the recommendation of a House of Lords Select Committee on Science and Technology, be setting up a commission of enquiry to consider the future role of GIS in Britain.(1) As the Select Committee's report indicates there is growing awareness of the need for spatial data handling procedures in industry, commerce, government and academia.(2) The rapidly growing availability of spatially related digital data, not least from the burgeoning field of remote sensing, has brought into being new techniques and software, termed Geographic Information Systems, designed to handle and analyse very large quantities of geographic data.(3)

Archaeologists too have extensive spatial data handling requirements for archaeological phenomena is underpinned by its unique position in euclidean space. Many archaeological site information systems, whether comprising manual card index systems or computer-based information systems, are linked to a series of hand drawn maps usually plotted on Ordnance Survey base maps which augment the information database. This is not surprising for historically the map has proved to be an efficient and effective method of storing and displaying spatial information.(4) The ability to incorporate these spatial plots into the knowledge base of the information system, however, is extremely limited and despite the importance of the spatial component in archaeology such maps are intrinsically outside the domain of the information system. Furthermore, the manual construction of a series of maps storing particular archaeological attributes is time consuming and laborious and considerable problems arise when the maps require updating or amending. The development of computer generated maps would considerably facilitate the production of such plots but the ability to interrogate and link spatial information stored either on maps or in the computer database remains extremely limited. This limitation largely arises from the fundamental problem of linking and integrating data recorded according to different spatial referencing systems, whether by point, line or polygon, within a single spatial database environment. The problems of managing and integrating a range of spatial data sets has been central to the development of GIS. The traditional method of integrating a series of spatial data sets has been to physically construct composite variables by superimposing maps one on another to indicate where phenomena overlap.(5) Such a method, however, is limited at most to a few maps, is a costly and time consuming process, and is antiquated in the computer age. During the last decade, with considerable impetus from North America, GIS of varying sophistication have been designed which provide integrated software for the creation, management, transformation, analysis and display of an infinite range of spatially related data.

Basically, GIS allow the bringing together of numerous spatial data sets to produce composite variables which may be displayed in graphical, tabular or statistical form or undergo further analytical modelling. There are several ways that spatial data may be encoded in preparation for manipulation in a relational database. The two most common methods involve either a polygonal or a grid mesh encoding structure. In a polygonal structure each polygon or zone has a unique shape conforming to the manner in which such phenomena occur on the earth's surface. Thus soils, field systems or administrative boundaries for example form unique polygons and the spatial coordinates of such zones are stored in vector form. Integration through the database for the examination of inter-relationships and the production of boolean type composite variables is achieved through the application of polygon overlay techniques. The second common approach is to superimpose a regular polygonal grid framework over the mapped data and to represent spatial variation in an ij matrix according to the cell in which the phenomena falls. The grid overlay structure may vary in resolution from a few metres to several kilometres and once the data has been encoded it is placed in a file and referenced by its ij geographic coordinates.

Generically GIS consist of several major components concerned with data encoding, data management including storage, manipulation and transformation, data search procedures, and graphical and analytical facilities. The analytical capabilities of GIS are particularly strong especially in the generation of composite variables and attribute files. The ability to generate buffers around points, lines or zones and to perform search procedures and analysis based on these buffers is just one of the advantages of GIS. Also potentially advantageous is the ability to link GIS with Ordnance Survey digital topographic data as currently being pursued with the experimental 1:625000 Ordnance Survey database.

THE BRIGHTON CASE STUDY

The exploratory system discussed here was set up with two main objectives in mind. The first was to create a computer-based inventory of archaeological sites in the Brighton area with associated environmental and archaeological information for retrieval in computer generated map and tabular form. The second objective was to go beyond the relatively 'passive' stage of data collation and storage to design a system which would facilitate advanced analysis by archaeologists and would be capable of supporting ad hoc, yet extensive, 'enquiry-based' requests for archaeological site information. This latter requirement was aimed primarily at satisfying 'third party' enquiries from organisations such as local or county planning departments and would allow the integration of Expert Systems within the structure. Theoretically, these two objectives should not be incompatible though in practice the design and structure of many record systems positively hampers such objectives being achieved.

A raster based archaeological system

The content and structure of this system has thus been customised to the

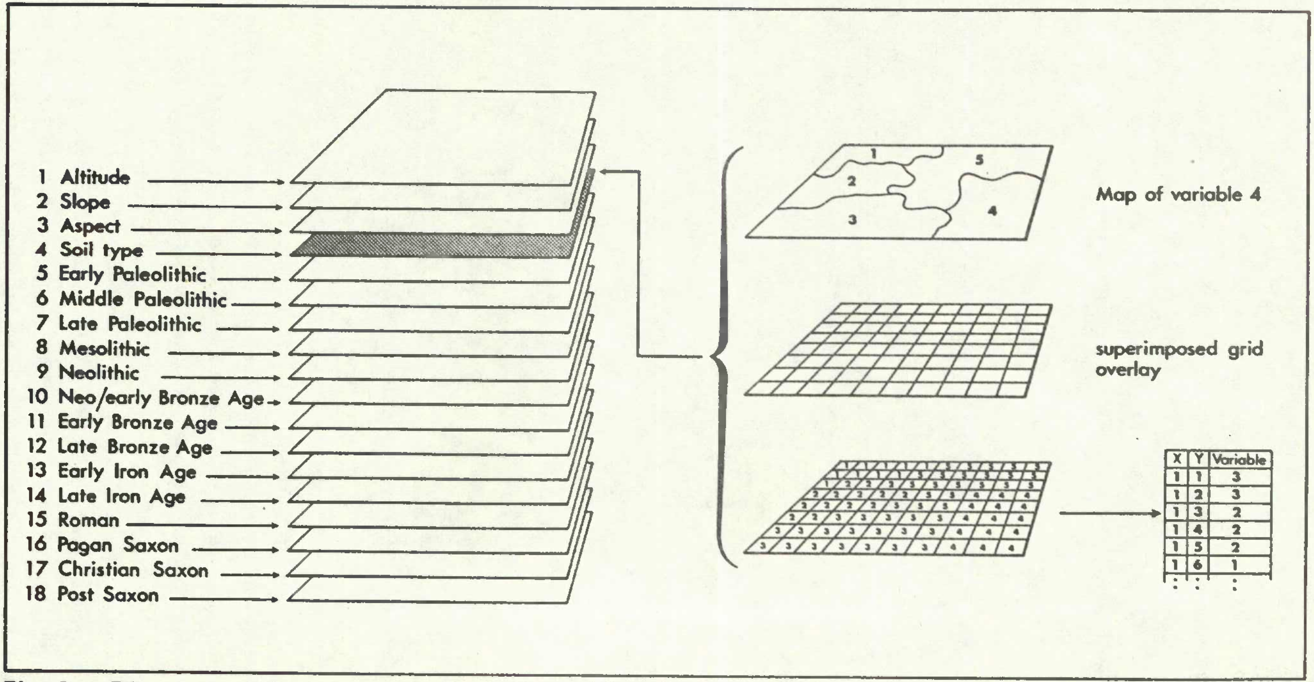


Fig 1 Diagrammatic representation of a grid cell map and file within an 18 level G.I.S. for archaeological site information retrieval

specific requirements of Brighton Museum. It may, however, be extended to be as complex and sophisticated in data content as required including integration with existing county based SMR. In its simplest form each site consists of three major attributes and a number of associated environmental variables. The attributes comprise a spatial reference for each site, the archaeological period of the site (estimated or otherwise) and the archaeological nature of the site. A major weakness of many spatial databases has been the inability to combine and link them with other databases because of the selection of 'odd' spatial keys. The selection of the Ordnance Survey national grid coordinate system considerably reduces this problem. Each site was also linked to a computerised index of Ordnance Survey site numbers thereby providing a potential link with other archaeological databases. The data was stored in an indexed sequential file designed to speed up data retrieval using primary and secondary 'keys'. The considerable advantages of indexed sequential files over sequential files is well known and is particularly apposite given the very large data files generated for use in GIS.(6) The keys under which the information is stored determine how quickly the data can be retrieved and with what ease. Given the desire to produce computer generated maps and that most enquiries envisaged from third parties would be concerned with the presence or absence of archaeological sites in relation to particular locations, the spatial coordinate was adopted as the unique primary key in the indexed sequential file. A series of secondary keys were also set up to accommodate alternative attribute search procedures. Within the system, however, there is the facility to redefine and to select alternative keys at will.

Archaeological site information was recorded for some 130 square kilometres of the Brighton area according to some 14 archaeological time periods. Spatial referencing and data capture in this instance was undertaken according to a raster structure in which a grid overlay of specific dimensions is effectively placed over the study area and site attributes and associated environmental and archaeological information recorded according to the cells in which they fell.(Fig 1) By recording the information in a systematic way the spatial coordinates of each cell can be computer-generated from a single known x-y coordinate. The use of digitisers for automated data entry and the application of vector to raster conversion techniques in GIS are particularly valuable for quick and efficient data encoding. Spatial resolution of varying extent were considered for data capture though the choice must be a compromise between accuracy and cost. Necessarily any increase in the resolution at which data is captured involves a corresponding but geometrical increase in the quantity of information generated. In fact it was found that the question of spatial resolution was of greater concern to the environmental and supplementary data because of the greater spatial variation in these variables. Bearing in mind the spatial accuracy of reported finds and sites and that the actual production of regional site maps could not easily portray greater spatial accuracy, an optimal choice of resolution of 100 metres square, based on Ordnance Survey eastings and northings was adopted. By effectively assigning a site or sites to the centroid of each cell the accuracy of resulting plots and analysis can be estimated to within a maximum error of 50 metres. This in itself overcomes considerable problems associated with the current practice of assigning a single coordinate to the centroid of sometimes quite extensive field systems. If a site 'overflows' a cell into another, in this instance, then that occurrence is recorded accordingly though care is taken not to record it as a separate site.

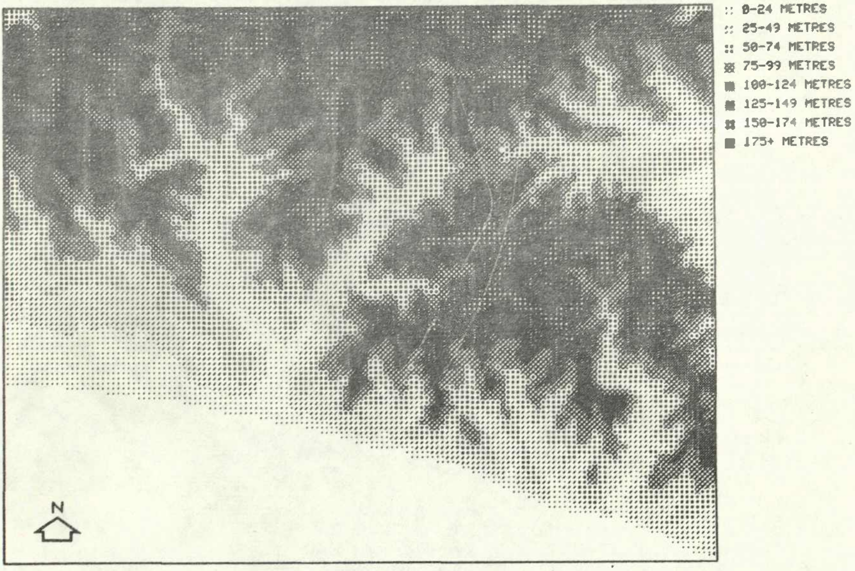


Fig 2 Two-Dimensional altitude plot of Brighton area

The utility of assigning information in such a manner becomes clear when relating archaeological information to environmental and explanatory site information such as soil type, geology, altitude, slope and aspect. The latter two variables were generated from the altitude information during the processing stage thereby avoiding considerable manual effort in their

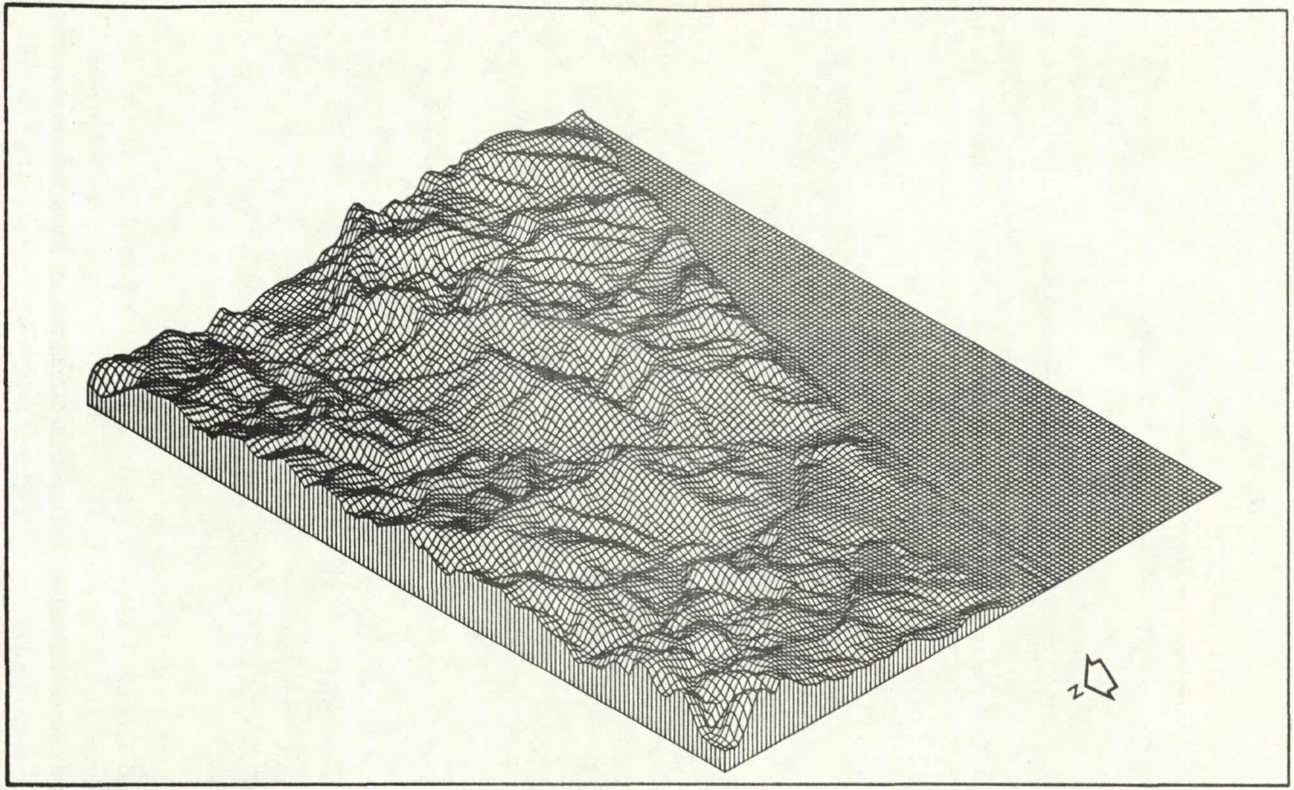


Fig 3 Three-Dimensional topographic plot of the Brighton area

production. In fact an environmental database comprising complete coverage of soil type and altitude at 100 metre square resolution was established and combined with the archaeological database. This further enabled both two and three dimensional plots of relief and sites to be combined with the archaeological information.(Fig 2 and 3) In connection with this latter aspect work is currently under way by the author to apply solid modelling techniques to archaeological site data whereby sites may be 'placed' graphically upon a three dimensional solid model of relief with obvious benefits for interpretation and presentation.

The archaeological description of each site was recorded according to a coding menu deliberatly chosen to provide the level of information required by Brighton Museum. Sites were classified typologically according to primary categories of ceremonial, domestic, defensive or funerary structures, cremations or inhumations, and field systems and a series of secondary categories under each primary category. Whilst these codes and map plots represent summary details of sites they can be extended to cover thousands of potential occurrences. A land-use classification currently in use by the author utilises a modified hierarchical coding menu extending over several thousand land-use categories. The temporal dimension of archaeological sites and the occurrence of multiple sites in close proximity entails an extensive coding menu to cover such eventualities. Once the information has been encoded and stored in an indexed sequential file, the resulting database may be interrogated by the database management system to produce information in a number of forms according to specific criteria laid down by the user.

Applications

A number of specific benefits of the system can be identified. Firstly, the resultant site inventory enables spatially related archaeological information to be linked and integrated throughout the database no matter what the original spatial referencing schema. In this way site information may be retrieved according to time-period, location, or attribute.(Figs 4-6) Archaeological sites may be linked with extensive environmental data including remotely sensed data. The system allows any part of the database to be updated or amended, retrieved, related or combined at will based on logical search procedures. It permits up to the minute statistical summaries, listings and, particularly, computer generated maps to be compiled at little effort to the user. The database may be linked with other databases for the exchange of information. A land-use, planning designation and environmental

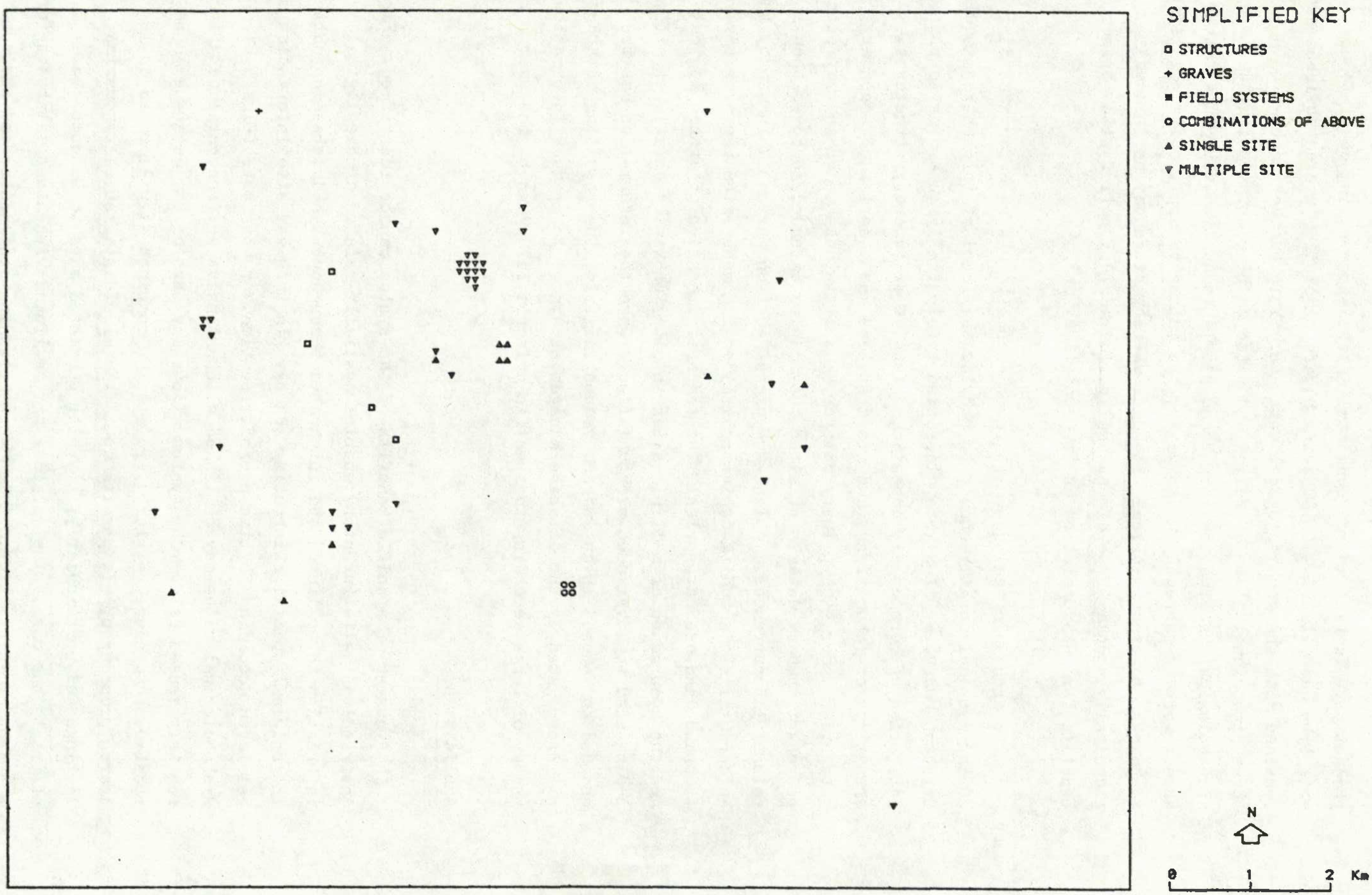
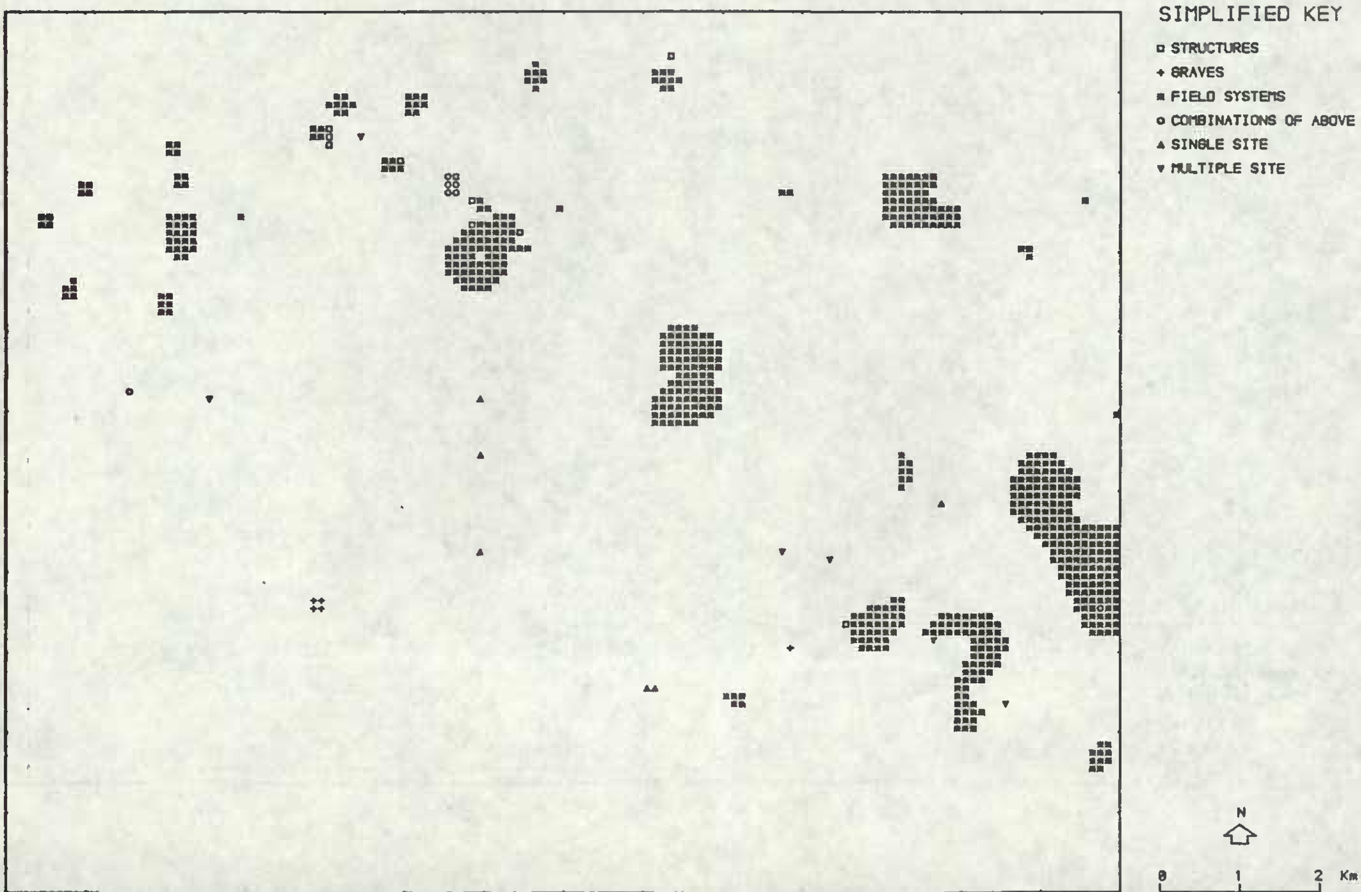


Fig 4 Neolithic Sites


SIMPLIFIED KEY
▫ STRUCTURES
+ GRAVES
▪ FIELD SYSTEMS
○ COMBINATIONS OF ABOVE
▲ SINGLE SITE
▼ MULTIPLE SITE
N
0
1
2 Km

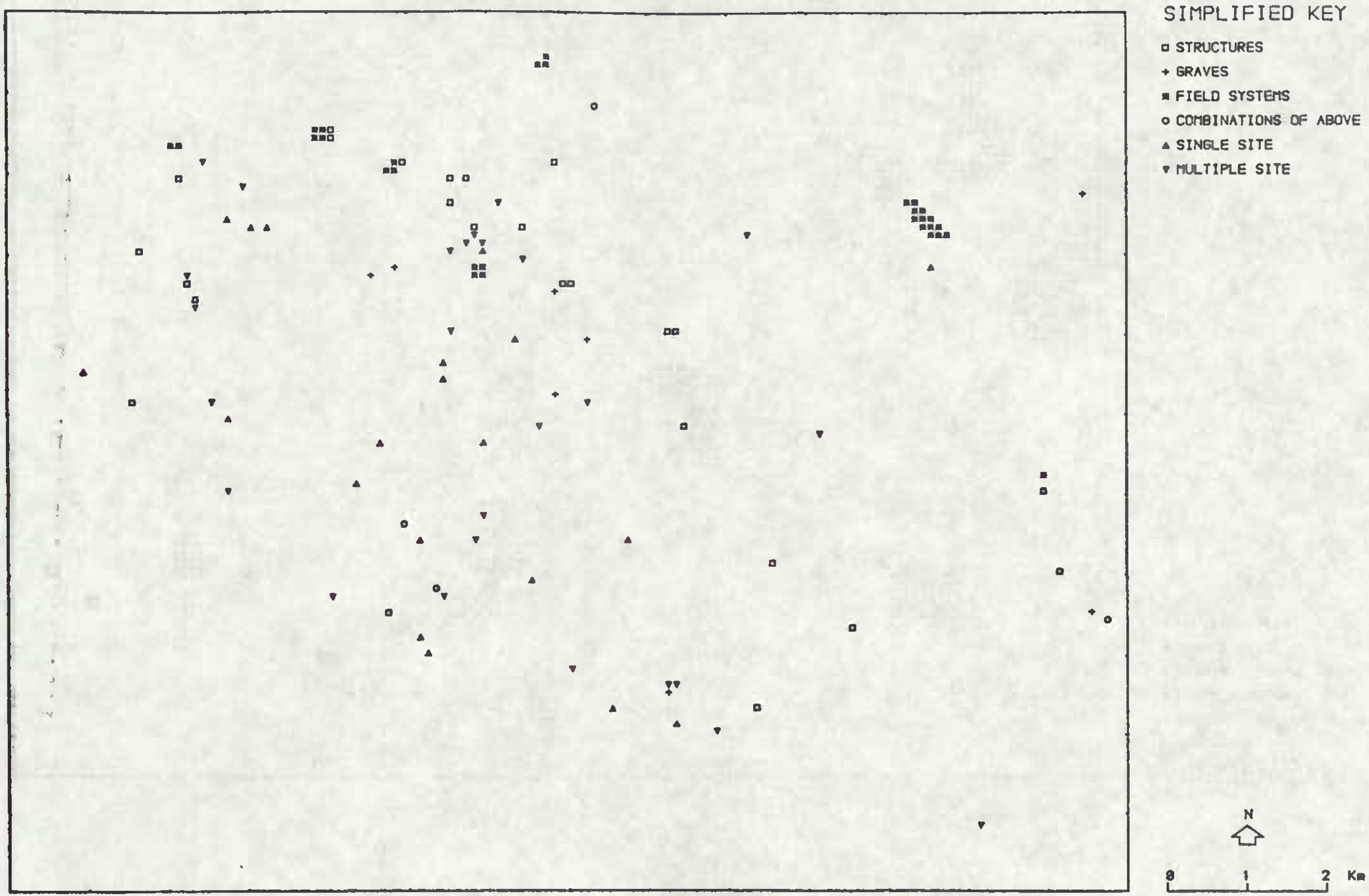


Fig 6 Roman Sites

database for example comprising some 100,000 records for East and West Sussex, an ecological database of West Sussex, and a geotechnical database of bore-hole data may be linked with this archaeological site database to enable the free exchange of information. Thus the linkage of archaeology with ecological conservation for example would be possible with this system and would provide a powerful tool for incorporation into the planning system. As indicated earlier the potential also exists to integrate GIS with Ordnance Survey topographic digital data.

Secondly, the data integration possible within the database permits advanced spatial and temporal analysis of archaeological phenomena to be undertaken with considerable ease. The ability to seek and examine relationships between sites and environmental variables is a particularly valuable aspect of GIS. Such relationships may be examined statistically through confirmatory or exploratory data analysis or through graphical output. The site data may be correlated with associated environmental data and extensive modelling of the information undertaken, as for example in the estimation of soil loss over time. Time-series analysis of archaeological sites and the examination of diffusion processes are further areas of analysis possible in the GIS context whereby spatial and temporal data are integrally linked. Other advanced statistical analysis is equally easy to perform. Thus the system could provide a basis for predicting archaeologically sensitive areas based upon multivariate relationships between known sites and environmental factors. Such forecasting methods may be invaluable in aiding the delineation of other, as yet undiscovered, archaeological sites and meaningfully reducing the extent of 'white' areas on archaeological site maps, particularly when responding to queries concerning development proposals. The results of a linear or polynomial regression in this instance may go some way in aiding such decisions and furthermore may be output in map form.

Thirdly, and importantly, the flexible graphical and mapping procedures possible within GIS allow the output of maps or graphics using a variety of output devices. Neccessarily the generation of composite variables via map overlay techniques in GIS is vital to such work and immediate cartographic output in hard or soft copy is possible. Output to aid interpretation, for publication or display can be undertaken in a variety of quality modes from low quality working maps to high quality high resolution colour maps eminently suitable for publication or display without the need for costly redrawing. Indeed with the current range of low cost high quality plotters perhaps the greatest problem facing workers in this field is the need to convert colour plots to monochrome for publication purposes.

Finally, data retrieval from the information system is quick and efficient and can provide the basis of an enquiry-based system for 'ad hoc' requests for information. In this respect the information system can act as a primary decision-making tool in the management of the historic environment. The system can respond to requests for information concerning the archaeological sensitivity of specified areas, as for example areas proposed for development, using the buffer facility. Thus requests for information concerning an area of say 200 metres either side of several proposed road routes can be provided graphically or in list form very quickly. The system can also be usefully incorporated into planning systems to meet the new European Community Environmental Impact Assessment (EIA) requirements for GIS have been used in EIA in the United States for many years.

CONCLUSION

The accurate recording of archaeological site information is a vital part of the archaeologist's responsibilities. The compilation and efficient storage of an exhaustive record of archaeological sites is not only a vital prerequisite for archaeological research but is particularly important as sites come under increasing pressure from urban and transport development and contemporary agricultural practices. One of the important aspects to arise from the current debate concerning computer applications in archaeology is the concern with how best to organise and structure archaeological data for the variety of uses required of it, not only now but in the mid to long term. This paper has indicated the potential of GIS in archaeological work concerned with site information retrieval. GIS obviously extends beyond that of a straight forward cataloguing system and has the potential to considerably increase the potential for the computer analysis of spatially related archaeological data. The ability to quickly output up to the minute archaeological data in many forms, not least in that of high quality computer generated maps, is a particularly powerful feature of GIS.

Notes

1. House of Lords (1984) **Remote sensing and digital mapping**, HMSO, London

2. ibid.

3. See Marble, D.F., Calkins, H.W. and Peuquet, D.J. (eds) (1984) Basic readings in geographic information systems, SPAD Publications

4. Steinitz, C. (1977) Hand drawn overlays: their history and prospective use, Landscape Architecture, 67, 444-55

5. McHarg, I.L. (1969) Design with nature, Natural History Press, New York

6. Deen, S.M. (1982) Fundamentals of database systems, Macmillan; Martin, J. (1976) Principles of database management, McGraw Hill