DEVELOPING A MULTIDIMENSIONAL GIS FRAMEWORK FOR ARCHAEOLOGICAL EXCAVATIONS

S. Merlo, A. Colin
Shell, Department of Archaeology, University of Cambridge, Downing Site, Cambridge. United Kingdom
sm399@cam.ac.uk,cas4@cam.ac.uk

KEY WORDS: 3D, excavation, exploration, model, GIS

ABSTRACT

Geographical Information Systems (GIS) are currently used in the area of archaeological excavation as mainly a data storage and simple output toolkit. The heterogeneous and complex nature of spatiotemporal datasets involving many variables needs to be addressed if GIS is to be of more use for the understanding and interpretation of excavations. Since archaeological objects are to be treated not only as static but dynamic (when considered as the result of temporally evolving processes), they can be referenced to one, two, three or four dimensions. The traditional 2-D data models are insufficient to support such data description and analysis. The diverging approach from the traditional planar (2-D) data models is multidimensionality, involving 3-D geometry and time. Although existing GIS software lacks the ability to support complex analysis involving space-time and further dimensions (attributes), establishing a framework for supporting analysis in such domains should be a priority of archaeologists wanting to contribute to the advancement of the use of GIS for intra-site analysis. This paper presents the development of a possible framework, where exploratory analysis and simulation using different types of data are used for knowledge discovery and enhancement. The designed concepts are applied to the Hoge Vaart database (collected in the Netherlands in the early 1990s during a rescue archaeology operation, using three-dimensional data collection criteria) and results are discussed in terms of applicability and assessment of the analytical framework.

1. INTRODUCTION

The spatiotemporal nature of processes is an indisputable paradigm in the discipline of archaeology. This paradigm is at the core of archaeological analysis and, due to the particular purpose of GIS (Geographical Information Systems) of handling spatiotemporal data, it lead to a very rapid integration of such systems in archaeological practice. Nonetheless the peculiarity of archaeological spatiotemporal data, as different from geographical data for which GIS was initially designed, has remained almost unexplored. A theoretical discussion of these themes, although certainly due, goes beyond the scope of this paper. The main aim is to present a practical way of approaching the spatiotemporal problem through the design of a framework that that bridges archaeological and computational concepts in order to achieve a platform for knowledge construction.

Proprietary GIS for a long time has remained essentially storage and display systems for two-dimensional spatial data formed from an association of geometric features and attributes. GIS, in particular in the field of intra-site archaeology, has been similarly used. It is difficult to say if this has happened due to software limitations or incapability of archaeologists to see beyond the immediate data storage potentials of the tool. At present, the initial rudimentary analytical functions of the software are developing into increasingly sophisticated tools for spatial modelling and analysis. Moreover, the interest of archaeologists in the application of spatial analysis for intra-site data is increasing. This encourages the incorporation of the available tools in archaeological research. However, the design aim of most widely used GIS packages has not been archaeologically driven. As a result, translation of the unique spatial concepts, relationships and processes typical of archaeology, formed independently of GIS, is not obvious and without misapplications. Therefore, although the current limitation of commercial GIS in providing a dynamic representation of spatiotemporal phenomena is only one obstacle to the development of spatiotemporal modelling, the major challenge is to formulate the conceptual framework and to integrate theories of individual behaviours, interaction and space/time constraints in order to model spatial micro-level dynamics. Another difficulty is to generate and analyse empirical data of micro processes in order to specify and calibrate dynamic micro models.

A framework is proposed that aims to integrate archaeological excavation data in all its phases, from on-site recording to specialists processing and overall interpretation of the considered site. The case study used for the intellectual construction of the framework is limited in terms of chronology and geographical scenarios, but the framework aims at generality and its application is by no means confined to these specific circumstances. Applicability and suitability of the framework as a methodological tool transcends a purely quantitative and case-specific approach towards excavation data integration and analysis. In fact, any system operating under circumstances that could account for both short- and/or long-term phenomena must provide flexible analytical support. The fundamental themes of space and time provide a common thread that runs through the various perspectives that relate to any given spatiotemporal system. These perspectives can be observed as micro-, meso-, and macro-levels of the application both in time and in space. In this way the intellectual effort moves away from direct descriptions of excavations (and in general of archaeological contexts) based on geometric and numeric spatiality (simple models with no analysis potentials) to complex models, with uncertain spatial parameters, varied effects and inputs where full system description might become impossible, but exploration and explanation could find a place.

The paper presents a GIS based framework for exploratory data analysis and dynamic modelling of archaeological processes (deposition and post-deposition). It outlines the application scenario which inspired an alternative mode of data integration and will concentrate on the temporal side of the multidimensional approach. The main objective of the work is to question traditional concepts of space, time, scale and modelling in archaeological theory and practice through the use of a tool that has been introduced to the discipline without much appraisal and evaluation.

2. APPLICATION SCENARIO

The Hoge Vaart project in Holland (Figure 1) researches the Mesolithic and Early Neolithic of the Flevoland polder province (Hogestijn and Peeters 2001, Peeters forthcoming). The investigation was conducted by ROB (Netherlands National Service for the Archaeological Heritage) and it took place between 1994 and 1997 within the framework of the completion of A27 motorway between Blaricum and Almere.

Figure 1. Geographical location of the Flevoland polder area, the Netherlands. The excavation site is located in the municipality of Almere

The project, funded by the Executive Ijsselmeergebied of the Ministry of Transport and Water, can therefore be considered a s rescue archaeology. The main interest of the research team was in the formation processes and landscape dynamics in the transition between the Mesolithic and Early Neolithic, which are barely represented in the regional literature. The project was conducted in two stages. At first, in 1996, exploratory research was carried out. This aimed at collecting material to understand the main principles of landscape development and soil formation of the Holocene coversand, underneath the peat. The area was cored at 61 locations along 12 lines and the cores were then analysed using a combination of geological, pedological, micromorphological and palynological research techniques, completed with a range of C14-dating. to obtain an overview of the geological development of the coversand and of the changes the climate and vegetation of the area underwent in the course of the Holocene. Moreover, the coring allowed the identification of the excavation area, on the base of the charcoal content. The site (approximately 8600 m^2, fig. 2) was located on a coversand ridge along a low area, dissected by various tidal gullies. The excavation revealed a large body of occupation remains and anthropogenic features. Data on paleo-environmental conditions were collected and interpreted in terms of landscape dynamics. Identified features consisted of nearly 120 surface hearts and several hundreds small stake holes. Approximately 100 deep hearth pits (essentially containing charcoal) were also recognised.

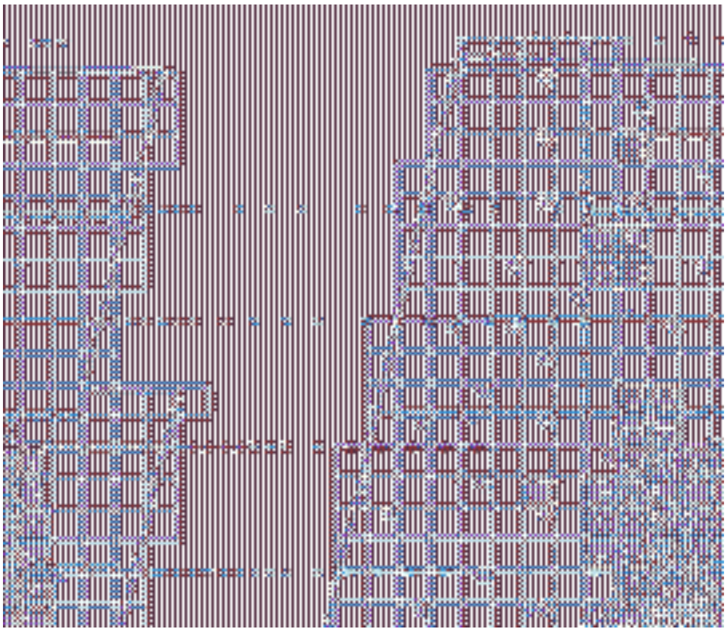

Figure 2. Hoge Vaart, Almere. Compisite map showing the extent of the excavation site, the gridding systems of 2x2 m and the finer resolution excavation areas (50x50 cm). The map presents also the location of the bore holes of the 1996 campaign and the features recognised during excavation

Data were collected using the so-called arbitrary excavation (Lucas 2001, 163) based on a system of 50x50 cm grid-cells, as it is usual in the Netherlands for Neolithic sites (Hans Peeters, p.c.). The underlying idea was the consideration that three-dimensional recording of individual objects is time consuming and total stations are expensive. At the same time the deposit presented ecologically very fragile material. As sieving recovers ecological material in good conditions and gridded spatial research is appropriate for finding structures where the limits of units are invisible, a combination of grid and spit systems was used to gather the data. Methods of assessment and adjustment of vertical grid resolution were constantly used and led to the decision of using a vertical resolution of 4 cm. No use was made of the Harris Matrix diagram. As the majority of the features was located within 15/20 cm and the character of the sediment was chaotic, this would have required an enormous investment of time. Although the principles of arbitrary excavation were employed during the excavation, when particular features were recognised, these were recorded by line drawing (mainly in plan but also in section) and a description and interpretation was given. The research strategy gave a fundamental importance to the use of automation in order to improve data quality, efficiency and monitor errors in real time. During the fieldwork an effort was made to collect, when possible, all data digitally. To achieve this, two parallel systems were used:

- 'automatic' 3D location definition by means of a tacheometer (total station) with a radio link to a hand-held computer and by means of altitude measurements with laser theodolite or water level in parallel with the assigning of work pit and section numbers. The find numbers were recorded by means of a bar code scanner, the feature data input through a menu on the hand terminal.
- A part of the field record was written and drawn in analogous form. This applied particularly to the recording and descriptions of the anthropogenic features. The field drawings were made in the

traditional manner 'by hand', without recourse to digital measuring apparatus. Nonetheless they were digitised and stored in dxf vector format at a later stage.

The dataset is characterised by an enormous potential in terms of quantity and quality of multidimensional (as intended in section 3.1) data. It presents a combination of data sources collected at different scales. It was therefore the ideal candidate for testing the ability of GIS to respond to the needs of complex archaeological datasets.

3. THE CONCEPTUAL FRAMEWORK

Archaeological excavation records are characterised by large heterogeneous spatiotemporal data sets involving many variables. The spatial data infrastructure is the combination of a subject-oriented framework and field data used for a variety of purposes (rescue archaeology, research oriented to answer specific questions, palaeo-landscape studies). Moreover, inhomogeneity of space and time and of time represented by space (underlying the principles of stratigraphy) must be taken into account too. Very often, for example, periods of intense activity are of greater relevance to the analysis than periods of relative inactivity. Similar requirements for context dependent spatio-temporal rules and user-defined semantics for the manipulation of variables are also to be noted.

In these circumstances, a computer framework is used to underpin spatiotemporal research in an exploratory and intuitive manner. Among the requirements of such framework Lee and Kemp (1998) list:

- *The user to deal with natural representations of the phenomena and relationship of interest*
- *The application of statistical functions to the raw dataset so that flexible partitioning of the problem space is achieved*
- *The intuitive and interactive visualization of selected subsets to enhance exploration and analysis*

3.1. Multidimensionality explained

The definition of multidimensional data is here used to express a concept that goes beyond the simple representation of space and time in a Cartesian world. The major limitations of intra-site archaeological analysis have been linked to the inability of current GIS systems to represent a three-dimensional space, in particular in terms of a vector structure where a triplet of coordinates represents the entity point (Harris and Lock, 1996). In fact, even when attempts are made to achieve such representation, the complexity of archaeological data and the richness of the aspatial attributes is frequently lost. In particular the direct connection often made between stratigraphic excavation, single contexts' identification, their recording and representation as surfaces is a dangerous oversimplification of other aspects of the excavation process and the excavated record. The Hoge Vaart excavation presented in the previous section is but one example of circumstances in which these methods are not only inapplicable but most probably misleading. During excavation practice the recording is often carried out under extremely difficult circumstances that require immediate and pragmatic reaction. Features may or may not be visible, time management may or may not allow to carry out a painstaking operation of single context planning. Data may be collected in gridded bulk samples and at random locations that are identified to represent the processes characteristic of the site. These data can not and are not to be represented exclusively by surfaces, rather by interpolated and extrapolated volumes with attached attributes. To conclude, the complexity of the data collected at sites such as Hoge Vaart can only be accounted for by a concept of multidimensionality that incorporates space as volumetric geometry, multiple spatiotemporal resolution and linked attributes and takes into account data inhomogeneity arising from the excavation pragmatic decisions on the way the site is excavated and recorded. The issue here concerns the accurate reconstruction of various phenomena, where the complexity occurs in all the dimensions involved in spatiotemporal analysis. Time, location, weight, densities are examples of what a dimension can represent.

3.2. Rationale of the system: exploratory analysis and knowledge discovery

Archaeological analysis and interpretation happens at different levels, touches several sub-disciplines and is carried out by a variety of experts. Nonetheless it is characterised by a common thread: the search for interesting patterns, which, in computing terminology corresponds to the concept of knowledge discovery. The main concern is to provide the archaeologist with a general capability for specifying the patterns and aggregates that may be required across several dimensions to produce such knowledge. Generally, different analytical techniques and base models are needed of a particular spatiotemporal dataset. Moreover, views of such dataset are often particular and different. In addition to this, in a multidimensional problem space the constraints and conditions that apply to each dimension can only be expressed in the context of the specific analytical process carried out. The aim of designing a prototype as a framework for archaeological information system is to inform the choice of data collection and storage strategies and of data structures to be used and integrated for an informed exploration and knowledge discovery within GIS terminology and capabilities. It provides a set of approaches to the design of patterns discovery in the dataset, allowing the user to make quantitative and qualitative considerations on the methods used and results obtained. It is a framework for reasoning within a reasonable position between the data and its existence within a computerised system. Conceptual design becomes here an exercise for alternative interpretation.

The conceptual framework is based upon an exploratory analysis approach, strictly connected with the two principles exposed by Kemp (1993), as follows:

1. *it is both desirable and possible to separate the mathematical operations which will be performed on data about spatial phenomena from the form of discretisation used to represent those phenomena in the computer*
2. *this separation allows issues about the implementation and manipulation of these digital representations to be dealt with automatically, without external control, in such a way that they can be considered extraneous to the modelling task.*

The aim is to provide the archaeologist with a general capability required for identifying and specifying the patterns and aggregates that occur across several dimensions and at different levels of analysis. This is an important issue that has been neglected up to now in the development of GIS applications, with the consequence of preventing a methodologically correct and practically useful use of GIS. Given the importance of building a framework of study, the fundamental requirements of the conceptual framework are identified and explained in the following section.

3.3. Underlying principles

Generally, different views of a particular spatiotemporal dataset are expressed by different requirements. Moreover, in a multidimensional platform the constraints and conditions to apply each dimension can only be expressed in the context of the analytical process being carried out. Therefore different ranges of requirements can be summarised for the spatial and temporal dimension. Despite the seeming difficulty in identifying connections across dimensions, two properties can be recognised:

- Multiple resolution (issues of scale in time)
- Multiple characterisation (issues of aggregation of space and time in units different from the ones used for recording)

The two identified properties are inspired by the Analytical Abstraction Layer (AAL) discussed by Lee and Kemp (1998), used amongst others for the management of marine fisheries.

- Multiple resolution is a concept that encapsulates issues of scale in space and time and is one of the parameters that determines the outcome of the analysis. When a dataset is explored at various spatial scales, corresponding degrees of detail are revealed; similarly, temporal information can also be presented at different levels of resolution (e.g. decennia, centuries, millennia), with each level showing different patterns and trends. When data is examined at a coarser spatio-temporal resolution than it was recorded, it needs to be aggregated. The same operations can also be used to generalise information that is considered less important for a specific level of analysis. If spatial resolution needed to be finer than the level at which it was recorded, decomposition and disaggregation could be achieved using, for example kriging.
- Multiple characterisation stems from the need of archaeologists to classify and categorise data in a meaningful manner. In addition to this, characterisation enhances the dynamism and expressivity of such operations by allowing conceptual mobility and multiple belonging. Data sets may belong to one or more characterisation classes or groups of classes. Each characterisation level or group can have many members. Categorisation classes allow the preparation of higher level data for further analysis. Multiple characterisation also allows the same data set to be portrayed in different graphical forms that emphasise different aspects of data. Moreover they allow the representation and compatibility of data collected with different criteria (single contexts transformed in digital vectors as opposed to voxel interpolated samples). Certain categorisation methods will reduce the number of dimensions or the complexity of a data set. Data sets from different classes can only be compared subject to a mapping function, used to reconcile not directly compatible datasets. Characterisation involves using as aggregation, inter- and extrapolation, classification, generalisation, categorisation and partitioning to interactively elicit patterns and anomalies alike.

3.4. The framework illustrated

Taking into consideration the issues discussed in the previous sections, the framework proposed (figure 3) is a system in which multiple resolution and characterisation are adopted to create micro-, meso- and macro-base models of the excavation data.

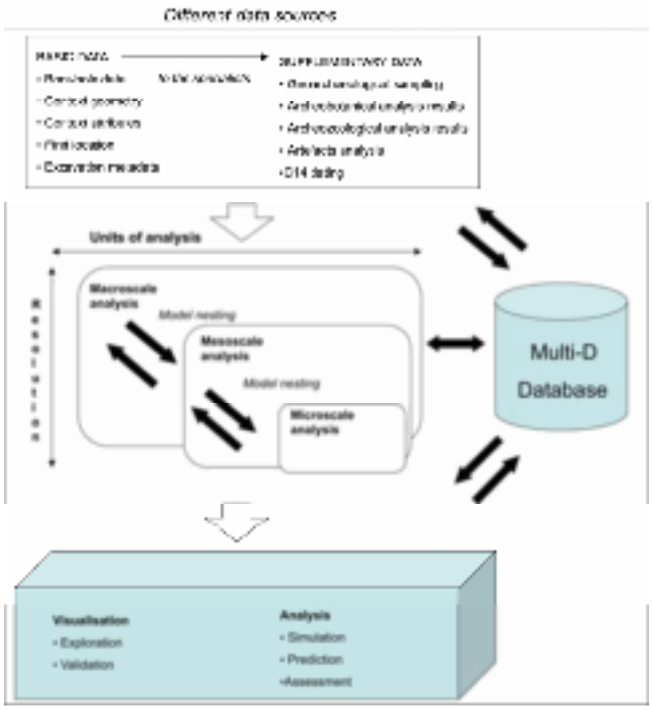

Figure 3. The spatiotemporal framework. Conceptual scheme of an integrated archaeological information and modelling system

This combination of raw data and their manipulation in terms of scale and characterisation implies possibilities of aggregation, classification, generalisation, specification and partitioning to enable pattern and anomalies to be elicited. The principle is that of a nested hierarchical structure where analytical operations (map algebra, extrapolation, interpolation, etc.) are used to produced new datasets suited to proceed to further phases of the analysis and visualisation. The elements of the system are connected by a double direction flow, to avoid the determinism typical of some hierarchical structures. As indicated above, the framework considers data multidimensionally: time, location, aspatial attributes, inheritance from previous levels of interpretation are examples of what a dimension can represent. Data as such are stored in the multidimensional database. The dimension of interest can be one or many at a time and data are explored using the pathways of multiple resolution and multiple characterisation, as shown in figure 3. These properties, common to all archaeological data, are the foundations of the system and between them they provide support for a wider range of analytical procedures. Multiple resolution and multiple characterisation become the axes of the base model building. Changing them can change the interpretation of the spatiotemporal phenomenon and can create different interpretative scenarios. These operations are used both for time and space to suppress or enhance detail, differentiate or generalise the components both for display and analysis.

The described framework is not limited to pure intellectual exercise. It is the conceptual thinking behind a prototype which exploits the capabilities of GRASS 5.3 raster engine for volume analysis and GRASS 6.0 new topological 2D/3D vector engine for vector network analysis. GRASS is a free Software/Open Source released under GNU General Public License (GPL).

4. RELATED ISSUES

The research presented in this paper is aimed at combining areas of GIS, spatial database systems, modelling, simulation, visualisation. The work should be viewed in the context of the complete framework for modelling and knowledge-discovery. By their nature, these activities are iterative and evolve in stages. The research conducted so far identifies other areas of investigation that necessarily follow the previous presented ones.Visualisation, closely linked to multidimensional approaches, is one of them. The system should provide a presentation mechanism that enables users to determine the combination of display and visualisation required – cartographic, graphic and tabular.

Another related issue concerns the capture and persistent storage of any derived multidimensional partition. The problem is obviously of importance in the area of archaeological data where datasets are large and distributed, and analysis progresses in stages.

5. CONCLUSIONS

This paper is the presentation of work in progress. It aimed at showing the problems and intricacies of analysis in the spatiotemporal domain by an account of the problematic and specific nature of such data in archaeology and proposes a framework. An important element of the framework is the modularity of the system, which provides better management and understanding of spatiotemporal problems. Data in the nested (level) system are multidimensional, where this multidimensionality is not only a measure of space and time components but also multiplicity of attributes and characterisation of layers of analysis. This means that we can view data at various levels of detail and in a multitude of representations and elaborations. In turn, the information we retrieve can be used for complex analysis and for constructing higher or lower order patterns of spatiotemporal processes.

With the focus on spatiotemporal information systems as opposed to spatiotemporal databases, we have taken a step towards realising the enormous potential of such systems. The framework is particularly effective at exploratory analysis whereby different types of data are used interactively and collaboratively by archaeologists from different areas of expertise and wanting to write a story of the excavation site taken into consideration. Although analytical techniques have been traditionally associated with statistics, the complex and diverse nature of archaeological data calls for an approach to analysis resembling a process of knowledge discovery. Knowledge discovery is fundamentally the search for interesting patterns. However, in order to support knowledge discovery, it is important for different data types to be able to exchange information in a standardised (meaningful) manner. The user-mediated mechanisms for organising the search process ensures that any discovered knowledge during analysis is consistent with the user's requirements. This also helps constrain the search process with domain knowledge not easily available within the system. We believe this framework will play an important role in helping to develop the next generation of GISs tailored for excavation and the thinking behind the application of GIS to intra-site analysis.

REFERENCES

Harris, T. H. & Lock, G. R. 1996. Multi-dimensional GIS: exploratory approaches to spatial and temporal relationships within archaeological stratigraphy, Analecta Praehistorica Leidensia, 28, pp. 307-316.

Hogestijn, J. W. H. and Peeters J.H.M 2001. *De Mesolitische en Vroeg-Neolitische vindplaats Hoge-Vaart A-27 (Flevoland).* Almere: RAM 79.

Lee, H. & Kemp, Z. 1998. Supporting complex spatiotemporal analysis in GIS. *Innovations in GIS 5*, pp. 151-161

Lucas, G. 2001. Critical approaches to fielwork. Contemporary and historical archaeological practice. Routledge: London.

Kemp, K. 1993. Environmental modeling with GIS: a strategy for dealing with spatial continuity. NCGIA Technical rep. 93-3 http://www.ncgia.ucsb.edu/Publications/Tech_Reports/93/93-3.PDF (accessed 13 June 2005)

Kemp, Z. and Lee, H. 2000. A Multidimensional Model for Exploratory Spatiotemporal Analysis. Geocomputation 2000 http://www.geocomputation.org/2000/GC033/Gc033.htm (accessed 01 June 2005)

Peeters, J.H.M., forthcoming. The Mesolithic to Early Neolithic occupation at the Hoge Vaart (Almere, the Netherlands): forager land-use dynamics in the context of a drowning landscape. In *Acts of the XIVth UISPP Congress (Liège, 2001).* Oxford (BAR International Series), Archaeopress.

ACKNOWLEDGEMENTS

Stefania Merlo is funded by the University of Milan, the University of Cambridge and Newnham College. Infinite thanks to Hans Peeters of ROB in the Netherlands for providing the case study data and giving never ending support and advice.