Sezione 3
Section 3

LA RICOSTRUZIONE SPAZIO-TEMPORALE DEI FENOMENI

SPATIO-TEMPORAL AND CULTURAL DEVELOPMENT STUDIES

Archeologia e Calcolatori
7, 1996, 303-312

INFORMATION SCIENCE IN ARCHAEOLOGY: A SHORT HISTORY AND SOME RECENT TRENDS

1. Introduction

In the first section of this paper, I will try to sketch some of the developments in the use of information science in archaeology, putting them in a more general framework of developments in archaeological theory. I will discuss the shift from "classical" statistical approaches, which concentrate on hypothesis testing, towards more heuristic, pattern-searching methods of analysis. Besides showing my own biases, my emphasis will necessarily be on what happened in the "Anglo-Saxon countries" (Moscati 1990).

In the second section, I will discuss some research I am undertaking presently on the use of Bayesian statistics for solving archaeological problems. In doing so, I want to illustrate, on the one hand, the ease with which rather complicated quantitative analyses can be performed with the help of standard computing tools, and, on the other hand, the risks of carrying out such analyses without a clear, logically sound underpinning.

2. A short overview of the application of 'quantitative methods' in archaeology

2.1 *The beginnings*

When defining 'quantitative methods' as any type of recording or analyzing archaeological materials with the help of numbers, it can be stated safely that the use of such methods dates back to the beginning of the twentieth century. At that time large-scale, detailed excavations and detailed description of the finds and stratigraphy had come to be considered necessary for understanding the cultural history of the different groups of the past (Trigger 1989, chapter 5).

Whether these groups were seen as 'aliens', the case in North America, or as 'predecessors', as in Europe and the Near East, was in this respect unimportant. The detailed description of artifacts, needed for the construction of typochronologies, minimally involved counts, and the early attempts to seriate types by Petrie (1901) can be considered the first steps towards quantitative analysis. Most analysis, however, was restricted to the visual inspection and comparison of frequency tables, and not much changed in this situation for over fifty years.

2.2 *Hypothesis testing and numerical classification*

It is probably justified to say that the more rigorous application of information science to archaeology has most of its roots in the American 'New Archaeology', advocated by Lewis Binford in the early 1960s (e.g. BINFORD 1962, 1965). Under the banner of neo-positivism, the early New Archaeology set as its task the formulation of general laws and theories for human behaviour and the construction of models by which proposed laws and theories could be tested. To construct such models, detailed description, which puts the emphasis on variability, was neither sufficient nor adequate. Instead, the dimensions which represented the underlying general processes had to be made visible by removing all variability that was considered 'random' in relation to those processes (e.g. BINFORD, BINFORD 1966).

The emphasis on the utilization of archaeological data for testing general propositions led to extensive use of existing techniques of 'classical' inferential statistics, often without much attention being paid to the mathematical conditions that had to be fulfilled to validly apply such techniques. Moreover, in those days the available hardware and software often dictated which methods to use, instead of the archaeological problem setting.

Besides the American 'New Archaeology', another trigger for the development of applications of information science in archaeology was numerical classification. This approach, originally developed in biology (SOKAL, SNEATH 1963), appealed to many archaeologists on both sides of the Atlantic Ocean, and led to extensive studies and discussions of the vast gamut of methods that soon became available (e.g. HODSON 1969, 1970). It also triggered (again) the discussion of one of the basic problems in archaeology, the meaning and purpose of classification in general (e.g. DUNNELL 1971). As such, many applications of numerical classification by cluster analysis implicitly accepted the existence of 'natural' or 'objective' classes that could be unveiled by using an 'objective' method.

2.3 *Patterns and variability*

By 1980, the time of 'naive' applications of information science had ended. It had become clear that the variability in the archaeological record could not be directly explained by general laws and theories or, otherwise, safely could be neglected. The 'law and order' archaeology had grown into processual archaeology, partly rooted in the general theory of open systems, an approach which was followed by many, not only in the USA, but all over the world. Around this time, the interest in this type of archaeology also began to increase among classical archaeologists, who study Near-Eastern, Egyptian, Greek, and Roman archaeology. Without going back to a culture-historical approach, variability and the understanding of its causes again became a subject worth to be studied.

In the USA, the work of Michael Schiffer and his followers led to a much better perception of the many influences the archaeological record undergoes before, during, and after its first formation. The archaeological record is *patterned*, and it is the task of the archaeologist to seek, describe, and explain the patterns. Explanations go from the level of understanding the local and general constraints the natural environment puts on the formation and preservation of the archaeological record to the level of understanding the impact, time-bound and spatially restricted, of human decision-making and action.

At this level, the level of 'middle range theory' (Raab, Goodyear 1984), explanation seems to end, however. The next step--trying to find ahistorical and spatially unrestricted explanations, the goal of the early New Archaeology, was not made. This was perhaps not unexpected. Apart from the enormous amount of working and thinking required before this large-scale problem could, and can, be tackled, the unwillingness to address it correlated, in my opinion, with a general change in the outlook of Western society. Interest in investigating what people have in common became more and more replaced by interest in how people divide themselves under labels like 'cultural identity', 'nationality', and 'ethnicity'. In archaeology, this attitude, found its clearest expression in the so-called 'postmodern' approaches, which, fortunately, did and do not gain much influence.

In the area of classification, the search for the 'natural' order of things became replaced by the notion that any classification and ordering is only valid within the context of a well-defined research-design. The definition of classes, the selection of attributes, and the levels of measurement all are explicit decisions which reflect the purpose of the research undertaken (e.g. Whallon, Brown 1982; Cowgill 1990).

2.4 *Informed guesses*

I estimate that around 1985 the application of information science in archaeology entered a new phase, a phase we still find ourselves in. Describing and mapping the variability of the archaeological record in formal and quantitative terms has become familiar, and the construction of formal-mathematical models to develop theories that explain this variability now is one of the main concerns (e.g. Doran 1990). Two approaches are of special interest.

The first concentrates on the construction of formal, dynamic models of processes of continuous change. These models are probabilistic in nature and tend to incorporate extensive computer simulations.

The second approach consists of attempts to understand behavioural processes in time and space by modelling human decision making with the help of concepts from artificial intelligence (e.g. Doran, Corcoran 1985; Reynolds 1986). In modelling decision making the Bayesian approach to statistical inference (Howson, Urbach 1993) is gaining more popularity (e.g. Buck 1993). This important development recognizes that, even when designing formal, 'objective' systems and processes, not all trajectories are initially equiprobable, or, even in case they are, will be recognized by the decision makers as such. Also, the Bayesian approach to information processing is, for me at least, intuitively satisfactory--it can serve as a formal model of the way in which the human brain human updates its knowledge and beliefs when new information becomes available.

2.5 Between Archaeology and Information Science: The role of the 'quantitative archaeologist'

To end the general section of this paper, I would like to discuss briefly the role of the archaeologist who has made the application of information science to archaeology his or her specialty. I did the same thing ten years ago, and, while many more archaeologists are now used to computers for storing and retrieving data and texts, I am afraid that on a more abstract level not too much has changed. It still is the case that only a minority of the archaeological community recognizes the potentials offered by information science for better understanding the archaeological record and, by that means, for better understanding humankind.

The metaphor I select to describe those of us who advocate a change in the thinking patterns of archaeologists is that of the *middleman.*«A 'classical' middleman is 'born' inside the culture of archaeology, has learned some of the language and culture of information science, alienates himself more or less from his archaeological culture and then functions as a channel through which information and goods are exchanged between both cultures. A middleman also has his own language. In that language, the concepts from different cultures are worded so that an interaction becomes possible. For applied information science in archaeology, the middleman language consists of mathematical models that are applied to archaeology.

Such models do more than only bridge the gap between archaeology and information science. They try to put concepts from archaeology and information science into a coherent, necessarily more general, framework of thought, thus creating concepts on a higher level of abstraction. Developing such concepts may be of more use to the 'parent' cultures than straightforward translations from one culture to another. In a different terminology: the historically determined dialectic opposition between science and humanities, as represented by old-fashioned mathematics and old-fashioned archaeology resolves itself in a synthesis--applied information science in archaeology.» (Voorrips 1985, chapter 1)

I think that what is expressed by the metaphor of the middleman still holds today. I also think that many of the prevalent attitudes and 'paradigms' in modern, or even later than modern, archaeology makes the work of the middleman extra hard, if not impossible. I hope, however, that it will be the classical archaeologists I referred to above, the archaeologists who wrestle with the rich data set of the Mediterranean and the Near East, rich in all senses of the word, who will first understand the real meaning of information science for archaeology. They have a running start since they do not need to invent or reinvent many wheels--there already exists a vast amount of literature dealing with issues of methodology, including classification, sampling, modelling, simulation, systems theory, etc. Furthermore, the current state of computer technology, of both hardware and software, makes it possible to concentrate on the real issues without having to spend a lot of time circumventing technical problems.

3. An example of a Bayesian approach to (spatial) classification

In the second part of this paper I will describe some of my recent experiments with the application of a Bayesian approach to decision-making in the course of classifying spatial units into clusters on the basis of their attributes. The example is not typically archaeological, although the issue came up in the context of an archaeological problem setting.

In 1994, I spent a trimester at the Museum of Anthropology of the University of Michigan, Ann Arbor, USA, where I taught a graduate seminar on the application of Geographical Information Systems in archaeology. One of the data sets used was derived from the ongoing work of the director of the Museum, Professor John O'Shea, on the role agriculture played for the original inhabitants of the Northern part of the Michigan peninsula (O'Shea, Milner, in prep.).

One of the things my students and I tried to figure out was which types of natural forest-vegetation the inhabitants would have had to cope with in different locations. This is not an easy problem, because that part of the United States was completely deforested in the last half of the nineteenth century, and the vegetation types distinguished among the secondary growth in unexploited areas are supposedly rather different from the original ones.

There is, however, an interesting record with the help of which a reconstruction of the original forest communities could be attempted. This record consists of data collected around the mid of last century by the General Land Office. One of the tasks of this office was to perform a geodetic survey of the State of Michigan in order to construct cadastral maps. To that purpose, the geodesists, or 'chainmen' put markers at the corner points of every square mile and quarter square mile in the area. To be able to find back these markers later, the chainmen were instructed to record the species and the diameter of some nearby trees, as well as their direction and distance from the marker.

In general, four trees were described in this way at each corner point of every square mile, two trees at each corner point of every quarter square mile, and, in addition, at least two trees along each mile-long section (BOURDO 1956). These trees were called 'bearing trees' or 'witness trees'. The chainmen were instructed to select the bearing trees on the basis of size and condition. This, together with a less-than-perfect knowledge of tree-species among the chainmen, and cases of obviously faked data, makes the sample to be found in the records of the General Land Office somewhat suspect. However, as various investigators of these records have reported, altogether the bias seems negligible (BOURDO 1956; HUSHEN *et al.* 1966).

The first step in using this sample of a little over 2500 trees to reconstruct the former forest-types in the region was to plot all tree locations on a map of the region. This map then was digitized, adding the information about the tree-type to each point. The digitized map and additional information were stored using the PC-GIS IDRISI (EASTMAN 1992).

Next, a decision had to be made on the definition of the spatial units to be used in the further analysis. One option was to put a grid over the area, and use the counts of the different species per grid-cell. This approach, which is technically the easiest one, has the drawback that it dissects the data, so that some trees located close to each other end up in different grid-cells. Another option, in my opinion the best one, would be to define a circle with a radius derived from the average distance of the bearing trees to the markers, to 'move' this circle over the map, and to establish non-overlapping spatial units whenever the number of points in the circle was over some threshold value. Unfortunately, no GIS-package or other computer program known to me at the time was able to do this (the paper by M. BAXTER and C.C. BEARDAH in this volume seems to provide a method to do this, however).

A third approach, and the one I decided to take, was to do a k-means cluster analysis of the locations, using their coordinates as variables. The number of clusters was set to 250, so that each cluster would contain approximately ten points. The coordinate-data were transferred from IDRISI to SPSS/PC+, and put through the k-means procedure this package provides (NORUSIS 1988). The k-means procedure of SPSS is rather primitive compared to the one found in Kintigh's package 'The Archaeologist's Analytical Toolkit' (KINTIGH 1988). One manually has to repeat the analysis until a stable solution has been reached, and there is no provision to compare the solution with solutions based on randomized distributions of the data, as available in Kintigh's package. However, the distance of each point to the centre of the cluster it belongs to, data needed later in the analysis, is included in the standard output.

The coordinates of the centres of the 250 clusters found were transferred back to IDRISI and translated into a point map. Using one of the IDRISI-routines, I constructed Thiessen polygons around the cluster centres and used these polygons as the spatial units in the rest of the analysis. Next, I wrote a small Fortran program that aggregated the information on the tree-type of each point and the spatial cluster it belonged to into a table of 250 rows, the spatial clusters, and 33 columns, the number of different tree-types. The cells of this table contained the counts for the tree-types, and, as can be imagined, most of them were empty. The table I transferred to a database management package for PC, in this case Microsoft Access.

At this point in the analysis it was necessary to decide on the method to use to group the spatial units into something that might represent the former forest-types. An obvious candidate for clustering the spatial units was again the k-means clustering procedure using the different tree-types as variables. But first, the problem of the empty cells had to be tackled. Without solving this problem, the choice would be either obtaining clusters composed of locations that shared the *absence* of many tree-types, or using a similarity coefficient with questionable mathematical properties, such as Jaccard's coefficient. After looking at the frequency distribution, I first removed all tree-types that occurred less then 40 times, ending up with 2333 points divided over 12 types, which was, on average, still less then one per cell. I then used Kintigh's k-means procedure, computing the solutions for two to ten clusters, and checking the validity with the randomization test.

The results were not convincing since there was only a weak patterning in the data. As for the number of clusters to be distinguished, five seemed to be the best. I added the results of the five-cluster solution to the table in the database.

It was clear that the matrix was too sparse for cluster analysis of some kind or another, and I therefore decided to try a different approach, inspired by the ideas of, among others, C.D. Litton (BUCK, LITTON 1993). I first assumed that there indeed were five forest-types in the region. Given that forest-types had been distinguished elsewhere by botanists trying to reconstruct the old forest vegetation in other parts of Michigan, I then assumed that it would be possible to decide on botanical and pedological grounds which five of those types could be expected to have existed in the study-area, may be assisted by the results of the k-means cluster analysis, weak as they were. If the probabilities of the occurrence of the different tree-types in those five forest-types were known, then, by applying Bayes theorem I could estimate the probability that a spatial unit belonged to one of the vegetation types, looking at the tree-types the unit contained. In that manner I would use only the 'real' information in the data and not similarities based on the empty cells in the matrix.

Unfortunately, it soon became clear that more or less reliable estimates for the probability of a tree-type to occur in a vegetation-type were not in the literature. The only possibility for estimating these probabilities was to make one more assumption, which was that the results of the k-means cluster analysis could be taken as a rough approximation of the composition of the former forest-types in the region. Under that assumption, I could compute the probabilities with which each tree-type occurred in each of the five forest-types.

Using the database I aggregated the data necessary, after which I had to write a small Fortran program to perform the actual computations.

I decided to use only those points in each unit which were at less than average distance from the cluster centre, assuming that in most cases these would be enough to reach an unambiguous result. By doing so, the compactness in space of the points used would be high, hopefully leading to better defined vegetation units. To calculate average distances and select the points obeying the criterium, I created a database table containing the identifiers of the spatial units and the distances of each point toward the centre of the unit it belonged to, using the output from the SPSS locational clustering. A simple query of this table produced the selection.

The outcomes of the analysis were interesting. A large number of the spatial units had a final probability of over 0.9 of belonging to a specific forest-type, and in a majority of the cases a specific forest-type had a probability of over 0.7. I decided to use 0.7 as cut-off value so that units with a lower probability for any of the forest-types were considered unclassified. To date, this is as far as the analysis has gone.

What can be done with the unclassified units? Here, again, a Bayesian approach may be useful. New sets of posterior probabilities can be calculated for them using, one at a time, the final probabilities of their neighbours (classified or not) as new information. This will, depending on the neighbour selected, in most cases result in one of the five posterior probabilities reaching a value of over 0.7. Identification then will be into the forest-type with the highest posterior probability. If no posterior probability reaches a value of over 0.7 the unit can be considered to represent a transition zone between two forest types.

The main reason for this rather detailed description of a small part of a project that is still under construction and, probably, of very restricted scientific meaning is to show how a Bayesian approach can offer alternatives for solving practical problems. A second reason is to show that the process is far from automatic--almost at every stage decisions by the investigator on how to proceed are necessary.

Finally, I wanted to show the relative ease with which an investigation like this one can be done using a few computer tools. Without a GIS, the calculation of Thiessen polygons is a nightmare, without a package like SPSS/PC+ or Kintigh's Toolkit, k-means cluster analysis is impossible, and without a relational database management package many manipulations of the data are cumbersome, to say the least. Notwithstanding all the ready-made computer tools, however, it was still necessary to write a few simple computer programs myself, to perform a number of needed computations not included in the available packages.

Albertus Voorrips
Faculty of Environmental Sciences
University of Amsterdam

Acknowledgements

First of all, I want to thank the organizers, in particular Dr. Paola Moscati, for inviting me to participate in the *III Convegno Internazionale di Archeologia e Informatica*, in Rome, November 1995. I am very grateful to the Museum of Anthropology, University of Michigan, which made it possible for me to spend a fruitful trimester in Ann Arbor in the fall of 1994, during which the research described in the second part of this paper was initiated. Special thanks go to the students who attended the seminar on 'GIS in Archaeology' I taught at that time. Their enthusiasm and eagerness provided me not only with a lot of tediously computerized raw data, but with much excellent food for thought as well. A final word of thanks goes to Susan Holstrom who corrected the English.

BIBLIOGRAPHY

BINFORD L.R. 1962, *Archaeology as Anthropology*, «American Antiquity», 28, 217-225.

BINFORD L.R. 1965, *Archaeological Systematics and the Study of Culture Process*, «American Antiquity», 31, 203-210.

BINFORD L.R., BINFORD S.R. 1966, *A preliminary analysis of Functional Variability in the Mousterian of Levallois France*, in J.D. CLARK, F.C. HOWELL (eds.), *Recent Studies in Paleoanthropology*, «American Anthropologist», 68, 238-295.

BOURDO E.A. 1956, *A Review of the General Land Office Survey and of its Use in Quantitative Studies of Former Forests*, «Ecology», 37(4).

BUCK C.E. 1993, *The provenance of archaeological ceramics: a Bayesian approach*, in J. ANDRESEN, T. MADSEN, I. SCOLLAR (eds.), *Computing the Past: Computer Applications and Quantitative Methods in Archaeology, CAA92*, Aarhus, Aarhus University Press, 293-301.

BUCK C.E., LITTON C.D. 1993, *Application of the Bayesian paradigm to archaeological data analysis*, in J. PAVÚK (ed.), *Actes du XII^e Congrès International des Sciences Préhistoriques et Protohistoriques 1*, Nitra, Institut archéologique de l'Académie Slovaques des Sciences à Nitra, 367-374.

COWGILL G.L. 1990, *Artifact classification and archaeological purposes*, in A. VOORRIPS (ed.), *Mathematics and Information Science in Archaeology: a Flexible Framework*, Studies in Modern Archaeology 3, Bonn, Holos Verlag, 61-78.

DORAN J., CORCORAN G. 1985, *A computational Model of Production Exchange and Trade*, in A. VOORRIPS, S.H. LOVING (eds.), *To Pattern the Past*, PACT 11, Strasbourg, Council of Europe, 349-359.

DORAN J.E. 1990, *Computer-based simulation and formal modelling in Archaeology: a review*, in A. VOORRIPS (ed.), *Mathematics and Information Science in Archaeology: a Flexible Framework*, Studies in Modern Archaeology 3, Bonn, Holos Verlag, 93-114.

DUNNELL R.C. 1971, *Systematics in Prehistory*, New York, The Free Press.

EASTMAN J.R. 1992, *IDRISI, User Guide and Technical Reference Manual*, Worcester, Clark University.

HODSON F.R. 1969, *Searching for Structure within Multivariate Archaeological Data*, «World Archaeology», 1, 90-105.

HODSON, F.R. 1970, *Cluster Analysis and Archaeology: some new developments and applications*, «World Archaeology», 1, 299-320.

HOWSON C., URBACH P. 1993, *Scientific Reasoning: the Bayesian Approach*, second edition, Chicago and La Salle, Illinois, Open Court.

HUSHEN T.W., KAPP R.O, BOGUE R.D., WORTHINGTON J. 1966, *Presettlement Forest Patterns in Montcalm County, Michigan*, «Michigan Botanist», 5, 192-211.

KINTIGH K. 1988, *The Archaeologist's Analytical Toolkit (C) 1985-1988*, Tempe, Kintigh.

MOSCATI P. 1990, *Indirizzi e sviluppi dell'Archeologia Quantitativa*, in P. MOSCATI (ed.), *Trattamento di Dati negli Studi Archeologici e Storici*, Roma, Bulzoni, 1-54.

NORUSIS M.J., SPSS Inc, *SPSS/PC+, Advanced Statistics V2.0*, Chicago, SPSS Inc.

O'SHEA J.M., MILNER C.M. in prep., *Late Prehistoric Economy and Ecology in Northeastern Lower Michigan*, «Memoirs of the Museum of Anthropology», Ann Arbor, University of Michigan.

PETRIE W.M.F. 1901, *Diospolis Parva*, London, Egypt Exploration Fund.

RAAB L.M., GOODYEAR A.C. 1984, *Middle range theory in archaeology: a critical review of origins and applications*, «American Antiquity», 49 (2), 255-268.

REYNOLDS R.G. 1986, *An adaptive computer model for the evolution of plant collecting and early agriculture in the eastern Valley of Oaxaca*, in K.V. FLANNERY (ed.) *Guila Naquitz: Archaic Foraging and Early Agriculture in Oaxaca, Mexico*, Orlando, Academic Press, 439-500.

SOKAL R., SNEATH P. 1963, *Principles of Numerical Taxonomy*, San Francisco, Freeman.

TRIGGER B.G. 1989, *A History of Archaeological Thought*, Cambridge, Cambridge University Press.

VOORRIPS A. 1985, *Chancy Choices: mathematical models in archaeology*, PhD thesis, Amsterdam, University of Amsterdam.

WHALLON R., BROWN J.A. (eds.) 1982, *Essays on Archaeological Typology*, Evanston, Illinois, Centre for American Archaeology Press.

ABSTRACT

In the first section of this paper, some of the developments in the use of information science in archaeology are discussed, putting them in a more general framework of developments in archaeological theory. It shows the shift from "classical" statistical approaches, which concentrate on hypothesis testing, towards more heuristic, pattern-searching methods of analysis. In the second section, some research is presented on the use of Bayesian statistics for solving archaeological problems. It illustrates, on the one hand, the ease with which rather complicated quantitative analyses can be performed with the help of standard computing tools, and, on the other hand, the risks of carrying out such analyses without a clear, logically sound underpinning.

Archeologia e Calcolatori
7, 1996, 313-326

A CRITIQUE OF G.I.S. IN ARCHAEOLOGY. FROM VISUAL SEDUCTION TO SPATIAL ANALYSIS

1. Spatial analysis in archaeology and GIS software

Most archaeological information is *spatial* in nature, because it deals with the placement of archaeological finds, contiguity and neighbourhood relationships between archaeological entities.

Since archaeologists realised the potential benefits of studying relationships between behaviour and the *spatial* distribution of material culture, new ways of representing, visualise and analyse archaeological findings have been proposed. Mapping, or the map-based approach has represented the earliest phase of spatial analysis in archaeology insofar as it has been for centuries an extremely efficient storage medium for condensing a large amount of spatial data and associated attribute variables into a single sheet of information. In this respect, until the earlier 70s archaeological spatial analysis had relied on descriptive informal methods based upon map inspection and almost intuitive impressionistic interpretation. This visual analysis benefited from the dissolution of composite maps into overlay plans showing selected features and categories, to examine their degree of correspondence and to make some subjective judgements about the strengths of the relationships between them. Although in the initial and exploratory stages of many types of archaeological spatial analysis such methods are still valuable, the human eye-brain system is not always a very precise instrument to assess the strengths of spatial relationships.

This fact lead to the adoption of formal quantitative methods borrowed from other disciplines such as plant ecology and geography in order to get more objectivity in data recording and analysis. This quantitative movement got off on the wrong direction because it made too often the straightforward assumption that archaeological problems were easy to solve and that statistics and archaeology shared the same problem solving logic. This thoughts yielded a mechanical non-reflective application of mathematical techniques to solve badly defined archaeological problems, without considering the analogy between mathematical models and real problems. For more criticisms of quantitative spatial analysis approach see, among others Aldenderfer 1987; Lock 1992; Ammerman 1992; Barceló *et al.* 1994.

In such an atmosphere of disappointment about the performance of formal techniques as they have been applied over the last two decades one may understand the rapid growth and expansion of Geographical Information Systems during the 80s. Currently, GIS software has been proposed as the best solution for nearly all archaeological problems, because it has the ability to store not only the locational and attribute data for each archaeological entity but also the *spatial relationships* between them. Nevertheless, despite widespread recognition that spatial analysis is central to the purpose of GIS, most applications to date have only shown their power to input, storage and manipulate spatial data, in order to elaborate computer mapping. In fact many of the current archaeological GIS projects are only a database containing a discrete representation of archaeological data in a static two-dimensional space, with a functionality largely limited to primitive geometrical operations to compute simple relationships between points in the space, and to simple query and summary descriptions (GOODCHILD *et al.* 1992). It seems as if the only goal was to insert the maximum quantity of information into a map. This indiscriminate map production is related with a lack of prior theory or hypotheses about the kind of problems archaeologists want to solve or about the expected relationships between spatial data, believing that mapping equals spatial analysis (GAFFNEY, VAN LAUSEN 1995). This tendency can explain why many recent applications of GIS, particularly in regional settlement location studies, reflect an environmentally deterministic approach to archaeological explanation (as have been pointed out by GAFFNEY, VAN LAUSEN 1995, VERHAGEN *et al.* 1995; HARRIS, LOCK 1995). Archaeologists are mainly working with environmental variables that are amenable to cartography describing the topography, lithology and hydrology of an area and that are relatively simple to map, forgetting the importance of social interaction in the analysis.

The research challenge lies in building on these principles a new theory of spatial relationships and using a new set of spatial analytical tools. There has been some work in that direction (BIRKIN *et al.* 1987; GOODCHILD 1987; BATTY 1988; CLARKE 1990; OPENSHAW 1990; FISHER, NIJKAMP 1992; FOTHERINGHAM, ROGERSON 1993; ANSELIN, GETIS 1992), but most of these references have not found their way into archaeology.

2. THE NEED OF A SOCIO-SPATIAL THEORY

2.1 *The concept of social space*

Even nowadays most archaeologist still do not recognise that space is a very general concept used in many different contexts to denote different things. There are abstract mathematical spaces (structures made up of arbitrary elements according to a set of axioms), psychological spaces, economical spaces, physical spaces or the "real" space in which human activity evolves, etc. (HERNANDEZ 1994).

The word "Space" seems to denote a "set" of entities to which may be attached associated attributes or properties, together with a relation or relationships, defined on that set (GATRELL 1983, 1991; PEUQUET 1988). *Space* is then the network of all underlying *spatial* relationships between different entities. A *spatial variable* is, consequently, any quantitative or qualitative property of data varying *spatially*, and which contribute to explain the dependency relationships between the locations of those entities (CRESSIE 1991). Given these features, we can define *social space* as any network of spatial relationships linking any set of *social units*.

2.2 *Social activity areas*

These social units are not always sets of people (families, local groups, villages, tribes, political territories, etc.), but any set of *social actions* (productive, reproductive) that have been performed in a single location. We may call these units *Social Activity Areas.* Given the fact that there are many different possible associations between social actions performed at a single location, there are many different ways to describe those social units, and thus many levels of representations. In this respect some isolated finds may define a low-level social activity area, whereas the spatial relationships among those low-level units may contribute to the definition of higher-level units. Here, *low* or *high* level refers not to any measure of relevance, but to the degree of structural complexity.

Some examples of social activity areas may be:

- an empty area, for instance, a ceremonial area after ritual cleaning
- an isolated grindstone
- a hearth
- a garbage pit or dumping area
- a storage pit
- any activity area within the settlement (butchering area, tool manufacturing area, food preparing area, metallurgical or artesanal activity area, etc.)
- a house
- a village (a set of houses, storage and garbage pits, activity areas, etc.)
- a grave (the set of rituals and ceremonies performed in that location, not the people buried there)
- a burial area (a set of graves and ritual structures)
- a territory (for instance all its village components and the set of social interaction networks among them)

Social activity areas are spheres of social interaction without any fixed boundaries; they are characteristically 'fuzzy' (GOODCHILD 1987; BURROUGH 1990; CASTLEFORD 1992; USERY 1993). Moreover, their features and topological characteristics change/vary according to time (LANGRAN 1989), because the units of social space are not only multidimensional, but *dynamic.* Consequently, not any single partition method of geographical space results in a model of the social space in use.

Archaeologists used to mechanically define social activity areas from discovered artefact concentrations. However, there is not a direct correspondence between the observed properties of archaeological contexts and social activities because of the enormous variety of transformational processes with different dimensions and temporal rhythms that can have acted upon the archaeological record. Consequently, in order to divide physical space in spatial areas where some social actions were performed, we have to discover *temporally dynamic, multi-dimensional* and *fuzzy* sets of spatial associations among archaeological finds, not only tool-kits, but also any kind of elements or features useful to diagnose a social activity.

In the same way, as it was shown by early etnoarchaeological studies, activity areas do not need to be spatially dispersed but there are different alternative models which embrace:

- Dispersed or segregated activity areas that assumes that different objects and features are partitioned into spatially distinct units, each corresponding to a single activity or group or related activities.
- Agglomerated or multifuncional activity areas, characterised by the overlapping of different social activities.

Given the effects due to *time*, we have to distinguish between more or less stable (fixed) areas and dynamic (temporally modified) ones. This temporal distance may be present in a single continuous episodic occupation but it is stronger as a product of different processes of reoccupation and reuse. Insofar as the archaeological record is the product of repeated depositional events over different periods of time, we must also take into account this change through time.

2.3 *Relationships between social activity areas*

Social activity areas are not spatially undifferentiated or isolated. They are in a intrinsically better or worse location for some purpose because of their position relative to some other meaningful unit: social groups build their own space because they appropriate some biophysical spatial areas and are able to defend them against the members from other groups (TRICOT, RAFFESTEIN 1979). Consequently, *space* is not a property of distinct areas existing outside and prior to society, but it is socially constructed (LEFEBVRE 1974; SOJA 1980, 1985; COUCELIS 1988; VERHAGEN *et al.* 1995; PALLARÉS 1993; BARCELÓ 1995).

The fundamental premise of the socio-spatial dialectic, already pointed out by Lefebvre is that social and spatial relationships are dialectically interreactive, interdependent; that social relations of production are both space-forming and space-contingent (SOJA 1980). In this respect, as spatiality is simultaneously the medium and outcome of social action and relationship, it is not only a product but also a producer and reproducer of the relations of production and reproduction (LEFEBVRE 1974).

As a result, in order to study *social* spaces we should discover the *spa-*

tial properties of all relationships linking those areas defined by the set of associated social actions performed there. Social Space is not only a partition of geographical space into "social" areas, but a network of interactions between those units. These interactions are built upon difference/similarity relationships between social activity areas, and configure a complex structure of *social* and *physical distances.* For instance:
- distance produced by the spatial proximity between each area
- distance produced by the diversity on resources in each area
- distance produced by the diversity of production activities in each area
- distance produced by the differences in volume of production in each area
- distance produced by the diversity of consumption activities in each area
- distance produced by the differences in volume of consumption in each area
- distance produced by the differences on quantity (density) of social agents in each area
- distance produced by the differences on the nature of social agents in each area
- distance produced by the differences on quantity of social interactions (contact) in each area
- distance produced by the diversity of interactions (contact) in each area

If we arrive to integrate in a GIS environment different layers showing fuzzy social activity areas at different levels of complexity and temporal modification, and we extract similarity relationships between any kind of units to define a multi-dimensional distance metric, we will be in the position to describe the structure of a social space.

However, current commercially GISs do not allow the representation and analysis of social space models, because they are still generally designed around basic raster and vector models which place primary importance on locations of geographic phenomena, sacrificing the rich analysis capabilities provided by structuring entities on the basis of classification attribution and interrelationships (GOODCHILD 1987; USERY 1993). This inappropriate language of representation is a consequence of the very fact that GIS systems lack a coherent body of theory and organising principles by which real-world archaeological entities can be represented in a *social* space.

New models must be developed to fully support spatial analysis and to relate particular social processes to particular spatial associations of objects, elements and relationships.

3. INTEGRATING SOCIAL SPACE THEORY AND SPATIAL ANALYSIS TECHNIQUES IN GIS PROJECTS

Some basic features of social spaces can be determined automatically through a combination of analytical and statistical processing, using set theory. For example, several social activity areas are defined by the presence/absence of specific archaeological finds. In those cases we can build a production rule associating the presence of archaeological data with their interpretation as a distinct social activity area. An expert system may be programmed to do this job, provided we have the right knowledge-base with ethnoarchaeological and experimental information (Barceló 1996).

Nevertheless, understanding the complexity of spatial processes and therefore how relational patterns are produced, controlled, and reproduced is not an easy task. Most social space categories are fuzzy, because it is not possible to specify a rule that identifies all of its members and only its members. The solution comes from a multi-dimensional approach that accepts the existence of social activity areas at different resolution levels. The goal of the analysis is to assess how low-level areas are organised in higher-level units, and how relationships between low-level areas contribute to explain the multi-dimensional structure of social space.

In the following sections we introduce a multi-dimensional approach which may be easily integrated into a GIS project. The framework can be summarised as follows:

- Geostatistical analysis of archaeological data (artefact concentrations, waste, ecofacts, architectonic remains, etc.) to discover discrete units ("partitions") in the geographical space that can be classified as distinct social activity areas.
- Ordering of social activity areas according to their complexity level (in terms of associations between different kinds of archaeological evidence). Social activity areas of the same level are included in a single layer. Different layers contain social activity areas at different complexity levels.
- Comparison of social activity areas through Boolean analysis between layers.
- Analysis of neighbourhood relations (spatial *dependence* model) between social activity areas, both within the same layer and between layers of different complexity level.
- Analysis of similarity relationships (*distance* model) between social activity areas, both in the same layer or between layers of different complexity level.

The result of all these techniques is not a "visualisation" of social space, but a *model* of social interactions between social units. We obtain not only a list and a description of social activities performed at different locations, but also the *dependence* between those locations produced by the similarities and differences between social actions. Given that *social interaction* is the formation process of social spaces, we describe "archaeological space" as a structure defined by the network of dependencies between social activity areas. In this way *social space* appears as something constituted, reproduced and changed by social relations, and in turn constraining the unfolding of such relations (Coucelis 1988).

3.1 *From artefact distribution to social activity areas*

Because artefact concentrations do not correspond necessarily with social spaces, we defend the use of several methods in a complementary fashion in order to define social activity areas using archaeological evidence as an indicator of their presence. All these methods are related with the partition of continuous physical space in discrete social units. To translate physical features into social structures, we should select material evidence of social actions, and the spatial distribution of those data. But the most important aspect of social partitioning is their multi-dimensional nature, that is, the need of associating disparate kinds of evidences (human manufactured with nature produced, foods with tools, luxury items with rubbish, etc.).

There is a fertile literature concerned with different techniques to identifying spatial patterning isolating discrete areas. This task may be performed by means of different clustering methods, that try to partition objects and features into groups based on observed similarities or differences. The methods of pattern recognition that at present seem to have a better performance are *Pure Locational Clustering, Unconstrained Clustering, Presab* (presence/absence method) and *Correspondence Analysis* (KINTIGH, AMMERMANN 1982; KINTIGH 1991; WHALLON 1984; BLANKHOLM 1991; GREGG *et al.* 1991; PALLARÉS 1993).

Image processing techniques provide an alternative approach to clustering. *Image segmentation* is the process of dividing an image into regions or parts of uniform appearance that have a strong correlation with objects or areas of the real world contained in the image (SONKA *et al.* 1993). In archaeology this can be used to locate areas where archaeological sites are likely to be found. Edge detection and image enhancement are techniques used to identify specific cultural features. Current applications in archaeology have focused mainly on field survey data, and regional intersite analysis, in order to locate sites and features, define physiographic regions, soil zones, etc. Nevertheless, most of these techniques can also be applied with any spatially distributed data in two or three dimensions at the intrasite scale, to investigate spatial organisation and thus establish social activity areas (VORION CANICIO 1993; LANG 1992). The purpose of image processing is not to see images, but to analyse information contained in an image, searching for unknown structure by removing the effects of noise or blurring, or to find a relation between an input image and an archaeological model.

Another possibility implies the use of unsupervised learning neural networks based on algorithms which implement some kind of "competitive" learning rules allowing clustering of input data solely on the basis of the intrinsic statistical properties of the set of inputs (CAUDILL, BUTLER 1992). Given a raster model of physical space in a GIS layer, the neural network allows the *classification* of different *discrete* input into a model of *continuous* space. For instance, let us imagine we have discovered two hearths and several postholes in a site. We know that each hearth constitutes a social activity area, but we are interested also in knowing if there are some activity areas at a higher complexity level associating the hearths with the postholes (for instance, one or two dwellings each centred around each hearth). An unsupervised neural network can be trained to calculate the degree by which the input activation from the postholes is assigned to each one of the hearths.

It is beyond the scope of this paper to detail technical questions or to make a comparative evaluation of the shortcomings and resolution power of each method and technique. The only thing we have to keep in mind is that in order to classify discrete partitions of physical (archaeological) space as social activity areas, we should include as many social data as possible, in the form of different kind of evidences (productive process evidences, consumption remains, residential structures, natural resources, etc.). This is the principal restriction of some statistical methods that are specially sensitive to differences in scale or that impose metric constraints on raw data. We simple advocate the use of those methods which have been developed to solve archaeological problems and operate under as few constraints as possible. If we use them in a complementary fashion, working both with continuous/binary data, and coordinate/gridded data, thus they may be useful to isolate discrete activity areas.

3.2 *Comparing social activity areas at different complexity levels*

Once we have defined some discrete units by means of different complementary methods, the next stage of the analysis consist of comparing these social activity areas at different complexity levels. The objective is to integrate in a GIS environment several layers with different information concerning location, morphology, size, archaeological content and contextual information of every discrete unit in order to build similarity relationships between social activity areas at different structural levels.

This process must be done by means of a formal GIS language with a defined syntax and vocabulary specific to map analysis which can define any model of spatial interrelationships. In this respect, it is needed a map algebra that defines not a simple arithmetic combination of map layers but integrates some more complex spatial operators.

Specially relevant for us are the possibilities to compute mathematical and Boolean operators with points and clusters of points. This paradigm is based on the formalised system for expressing GIS functions developed by C. Dana Tomlin (Tomlin 1990; Mills 1994). The representation language MapAlgebra, described by Mills (Mills 1994), seems one of the most powerful modelling paradigms because it works with different operations that seem able to induce any kind of associative principles, connections and relationships between the variables of interest to describe social spaces.

3.3 *Geographical distance between social activity areas*

Our next task is to measure the degree of *spatial* variation in each layer of social activity areas, that is, variation in social space due to neighbourhood distances. What we are looking is if what happens in one social activity area is related (depends on) with what happens in the mean of neighbour social activity areas. If the observed spatial variability between social areas has not any known source (time, function, ethnicity, culture, economy, society, etc.), then we shall not expect any spatial association. Spatial heterogeneity occurs when there is a lack of spatial uniformity in relationships between the variables under study. When the variation is not wholly erratic, and there is some regularity, we say that there is a certain degree of *spatial dependence* between spatial units (OLIVER, WEBSTER 1990). The analysis then pretends to examine if the characteristics in one location have anything to do with characteristics in a neighbour location through the definition of a general model of the space. There are many different techniques to compute if there is some degree of spatial dependence (CRESSIE 1991; GETIS, ORD 1992; TRICOT 1987).

It may also be useful to calculate a *model* of the dependency structure discovered. Surface interpolation is the most usual technique to perform this task. Polynomial surfaces of various orders may be fitted to the maps containing social activity areas. The goal is to obtain a geometrical surface generalising the observed distribution of data to portray their overall patterns of location (SCHIEPATTI 1985; OLIVER, WEBSTER 1990; CRESSIE 1991; VOIRON CANICIO 1993).

Once we know whether neighbouring social space units are similar or not, we have to explain why the location of social activity areas shows that level of spatial homogeneity or heterogeneity. This can be done using contextual classification methods, that is, the assignation of conceptual labels (explanations) to our social space dependence model, relating different social actions (manufacture, cooking, residence, ritual, etc.) with the discontinuities measured on our model of social space. The problem is that most actual GIS projects confound geographical variables (soil productivity, soil erosion, vegetation, etc.) with contextual data, ignoring the specific nature of contextual information. Contextual data can be defined as those that are relevant to identifying some archaeological observation or pattern, or to interpreting some facts, excluding the data that are used by the one model to make the identification/ interpretation (CARR 1991).

3.4 *Other distance measures between social activity areas*

There is not any single method to analyse the effects of non geographical distance on the variability between social activity areas. One of the most common approaches in recent archaeological studies consist of applying multidimensional scaling, or correspondence analysis, in order to *classify* archaeological data in different social activity areas.

We can also use unsupervised learning Neural-Network to "translate" feature input vectors into neighbourhood functions using similarity relationships between input vectors (KOHONEN 1988). The purpose of the system is to evolve localised response patterns to input vectors. When two input vectors are similar, they evoke similar localised response patterns. Consequently patterns of high dimension (distinctive features of social activity areas) are transformed into a two-dimensional pattern, preserving the ordering of the input patterns. Distances between points will not be preserved but their topology will -that is, input vectors that were adjacent to each other will still be adjacent to each other. Given that distance metric is not preserved on the output layer, the result is not a representation of physical space, but of social space *excluding* distance between social activity areas. The location of areas in the Kohonen layer has not any sense, but the degree of social partitioning measured can serve as an evidence for the *complexity* and *differentiation* level of social space.

Another way to study non-geographical distance models on social space would be through graph theory (HERNANDEZ 1994). Mathematically speaking we can translate the locations of social activity areas into vertices which are connected by edges to points with the same value on a spatial variable (similarity). We call it the *neighbourhood structure*, because we suppose that the closer the points, the closer the measures of spatial space.

The analysis of similarities and differences between social activity areas does not end here. We can use many other methods derived from Classification Theory and Machine Learning (BARCELÓ 1996). Some of these methods can be integrated into a GIS framework (specially induction methods and genetic algorithms), but some other (fuzzy cognitive maps, for instance) are difficult to integrate. Much more work is needed in this area of spatial analysis.

3.5 *From spatial analysis to visualisation*

In our view current approaches to GIS applications in Archaeology are based on the wrong assumption that inference proceeds from *visualisation* to spatial analysis, as if visualisation was a tool for spatial analysis. To us, the best procedure is to use visualisation tools to support results from spatial analysis.

Visualisation is a way of explanation: it transforms the symbolic into the geometric, enabling researchers to observe their simulations and computations. Data visualisation is a means whereby much more multi-dimensional data can be brought within the range of human experience and cognition. It should be stressed that the graphical aspect of solid modelling systems is not necessarily their prime function. Realism in solid modelling is achieved not only by modelling natural lighting effects, but by trying to embody some element of the social context and function of the subject in the visualisation and thereby bring the visualisation to life.

Geographic visualisation will be defined here as the use of concrete visual representations-whether on paper or through computer displays or other media- to make spatial contexts and problems visible. Visualisation in a GIS environment then has to be a part in the process of scientific discovery, helping archaeologists to detect regularities in the patterns of spatial relationships among social contexts. It may be integrated into a GIS platform as a means of presenting a series of hypothetical scenarios which are relevant to our understanding of human/environmental relationships. This can be done principally through the construction of a 'territorial' model designed to articulate a set of semi-autonomous activity spheres which are said to be implicated in the reproduction and organisation of a specific archaeological locus, or settlement (FLETCHER, SPICER 1992; LOCK 1992; HARRIS, LOCK 1995; REILLY 1992; VERHAGEN *et al.* 1995).

4. CONCLUSIONS

In recent years GIS has emerged as the best solution for nearly all archaeological problems due to its ability to manage large amounts of georeferenced data and integrate different kinds of spatial relationships. Here it has been argueed that one of GISs major restrictions as they have been commonly applied is their lack of analytic capacities. The wrong assumption that visualisation equates Spatial Analyis has led archaeologists to reduce problem solving to the making of pretty but sometimes meaningless pictures.

In our view GIS software can be of great utility in Spatial Analysis but only if we use it in the frame of a well reasoned theory, posing the appropiate questions to explain historical phenomena. The purpose of these essay has been to introduce some elements for a *theory of spatial relationships* needed to study *social spaces*. To this end we have proposed an operational *multidimensional approach* to discover social activity areas, which can be easily integrated into a GIS framework. This proposal involves the use of some already existing analytical tools in a complementary fashion (geostatistics, intrasite spatial tests, digital image processing, artificial intelligence, etc) in such a way that allows the model building of social interaction between different social units and therefore a better approximation to historical explanation.

JUAN A. BARCELÓ
MARIA PALLARÉS
Universitat Autònoma de Barcelona
Área de Prehistòria. Facultat de Lletres
Bellaterra - Barcelona

Acknowledgements

In this paper we discuss some theoretical assumptions and some methods we are using in our research on hunter-gatherer archaeology (M. Pallarés) and early iron age research (J.A. Barceló). We acknowledge fundings from the Spanish Ministry of Science and Education (Project PB93-0397 and Project PB93-0867) and a FI grant from the Catalan Comission for Research and Technology.

BIBLIOGRAPHY

ALDENDERFER M. 1987 (ed.), *Quantitative Research in Archaeology. Progress and Prospects*, Newbury Park, California, Sage Publications.

AMMERMAN 1992, *Taking stock of quantitative archaeology*, «Annual Review of Anthropology», 21, 231-55.

ANSELIN L., GETIS A. 1992, *Spatial statistical analysis and geographic information systems*, «Annals of Regional Science», 26, 19-33.

BARCELÓ J.A. 1995 (in press), *Podemos determinar arqueológicament la interacción social?*, in *II Jornadas de Arqueología a la Carta*, Carratraca, Diputación de Málaga.

BARCELÓ J.A. 1996 (in press), *Arqueología Automática. La utilización de las técnicas de Inteligencia Artificial en Arqueología*, Sabadell, Ed. AUSA.

BARCELÓ J.A., ESTÉVEZ J., PALLARÉS M.,WÜNSCH G. 1994, *The easier, the better. The proper role of Statistics in Archaeology*, in I. JOHNSON (ed.), *Methods in the Mountains. Proceedings of the UISPP-Commission IV (Data Management, Mathematical Methods and Computing in Archaeology). Inter-Congress Meeting '93*, Sydney, Sydney University Archaeological Methods Series, 2.

BATTY M. 1988, *Informative planning: the intelligent use of information systems in the policy-making process*, Technical Reports in Geo-Information Systems, Computing and Cartography, 10.

BIRKIN M., CLARKE G.P., CLARKE M., WILSON A.G. 1987, *Geographical Information Systems and Model-based Locational Analysis; Ships in the Night or the Beginnings of a Relationship?* (WP-498), School of Geography, University of Leeds, Leeds.

BLANKHOLM H.P. 1991, *Intrasite Spatial Analysis in Theory and Practice*, Aarhus, Aarhus University Press.

BURROUGH P.A. 1990, *Methods of spatial analysis*, «International Journal of Geographic Information Systems», 4 (3), 221-230.

CARR C. 1991, *Left in the Dust. Contextual Information in Model-focused Archaeology*, in E.M. KROLL, T.D. PRICE (edd.), *The Interpretation of Archaeological Spatial Patterning*, New York, Plenum Press.

CASTLEFORD J. 1992, *Archaeology, GIS and the time dimension: an overview*, in G. LOCK, J. MOFFET (edd.), *Computer Applications and Quantitative Methods in Archaeology 1991*, BAR Int. Series S577, Oxford.

CAUDILL M., BUTLER M. 1992, *Understanding Neural Networks. Computer Explorations*, 1, Cambridge (MA), The MIT Press.

CLARKE M. 1990, *Geographical Information Systems and Model Based Analysis: Towards effective decision support systems*, in H.J. SCHOLTEN, J.C.M. STILLWELL (edd.), *Geographical Information Systems for Urban and Regional Planning*, Dordrecht, Kluwer Academic Publishers, 165-175.

CRESSIE N. 1991, *Statistics for Spatial Data*, New York, John Wiley.

FISCHER M., NIJKAMP P. 1992, *GIS and Spatial Analysis*, «Annals of Regional Science», 26, 3-17.

FLETCHER M., SPICER D. 1992, *The display and analysis of ridge-and-furrow from topographically surveyed data*, in P. REILLY, S. RAHTZ (edd.), *Archaeology and The Information Age. A Global Perspective*, London, Routledge.

FOTHERINGHAM A.S., ROGERSON P.A. 1993, *GIS and spatial analytical problems*, «International Journal of Geographical Information Systems», 7, n. 1, 3-19.

GAFFNEY V., VAN LEUSEN M. 1995, *Postscript-GIS, environmental determinism and* *archaeology: a parallel text*, in G. LOCK, Z. STANČIČ (edd.), *Archaeology and Geographical Information Systems*, London, Taylor & Francis, 367-382.

GATRELL A.C. 1983, *Distance and Space: A Geographical Perspective*, Oxford, Oxford University Press.

GATRELL A.C. 1991, *Concepts of space and geographical data*, in D.J. MAGUIRE, M.F. GOODCHILD, M.F., D.W. RHIND (edd.), *Geographical Information Systems: principles and applications*, 1, London, Longman, 119-34.

GETIS A., ORD J.K. 1992, *The analysis of spatial association by use of distance statistics*, «Geographical Analysis», 24, n. 3, 189-206.

GOODCHILD M.F. 1987, *A Spatial analytical perspective on Geografic Information Systems: principles and applications*, «International Journal of Geographical Information Systems», 1, n. 4, 327-334.

GOODCHILD M., HAINING R., WISE S. *et al.* 1992, *Integrating GIS and spatial analysis: problems and possibilities*, «International Journal of Geographical Information Systems», 6, n. 5, 407-423.

GORENFLO L.J., GALE N. 1990, *Mapping regional settlement in information space*, «Journal of Anthropology», 9, 240-274.

GREGG S.A., KINTIGH K.W, WHALLON R. 1991, *Linking ethnoarchaeological interpretation and archaeological data. The sensitivity of spatial analytical methods to postdepositional disturbance*, in E.M. KROLL, T.D. PRICE (edd.), *The Interpretation of Archaeological Spatial Patterning*, New York, Plenum Press.

HARRIS T., LOCK G. 1995, *Toward an evaluation of GIS in European archaeology: the past, present and future of theory and applications*, in G. LOCK, Z. STANČIČ (edd.), *Archaeology and Geographical Information Systems*, London, Taylor & Francis.

HERNANDEZ D. 1994, *Qualitative Representation of Spatial Knowledge*, Berlin, Springer-Verlag (Lecture Notes in Artificial Intelligence, 804).

KINTIGH K.W. 1991, *Intrasite Spatial Analysis: A commentary on Major Methods*, in A. VOORRIPS (ed.), *Mathematics and Information Science in Archaeology: A Flexible Framework*, Studies in Modern Archaeology, 3, 165-199.

KINTIGH K.W., AMMERMAN A.J. 1982, *Heuristic approaches to spatial analysis*, «American Antiquity», 47, (1), 31-63.

KOHONEN T. 1988, *Self Organization and Associative Memory*, 3rd. Ed. Berlin, Springer-Verlag.

LEFEBVRE 1974, *La production de l'Espace*, Paris.

LOCK G. 1992, *Visualizing spatial data: the importance of Geographic Information Systems*, in P. REILLY, S. RAHTZ (edd.), *Archaeology and the Information Age: A Global Perspective*, London, Routledge, 81-96.

MILLS B. 1994, *MapAlgebra: an emerging standard for analysis in GIS*, «GISEurope», 3, n. 9, 18-20.

OLIVER M.A., WEBSTER R. 1990, *Kriging: a method of interpolation for geographical information systems*, «International Journal of Geographical Information Systems», 4, n. 3, 313-332.

OPENSHAW S. 1990, *Spatial analysis and geographical information systems: a review of progress and possibilities*, in H.J. SCHOLTEN, J.C.H. STILLWELL (edd.), *Geographical Information Systems for Urban and Regional Planning*, Dordrecht, Kluwer Academic Publishers, 153-163.

PALLARÉS M. 1993 (in press), *Cuestiones teórico-metodológicas sobre el estudio de la organización espacial de los asententamientos de comunidades cazadoras recolectoras. Sota Palou, un caso práctico*, in *II Congreso Internacional de Informática y Arqueología*, Bilbao 1993.

PEUQUET D.J. 1988, *Representations of geographic space: toward a conceptual synthesis*, «Annals of the Association of American Geographers», 78, 375-94.

REILLY P. 1992, *Three-dimensional modeling and primary archaeological data*, in P. REILLY, S. RAHTZ (edd.), *Archaeology and the Information Age. A Global Perspective*, London, Routledge.

SCHIEPATTI F.J. 1985, *Intrasite structural analysis using semivariograms*, «American Archaeology», 5 (1), 47-58

SOJA E. 1980, *The socio-spatial dialectic*, «Annals of the Association of American Geographers», 70, 207-25.

SOJA E. 1985, *The spatiality of social life: towards a transformative retheorisation*, in D. GREGORY, J. URRY (edd.), *Social Relations and Spatial Structures. Critical Human Geography*, London, 90-127.

SONKA M., HLAVAC V., BOYLE R. 1993, *Image Processing, Analysis and Machine Vision*, London, Chapman & Hall.

TOMLIN C.D. 1990, *Geographic Information Systems and Cartographic Modelling*, Englewood Cliffs (NJ), Prentice-Hall.

TRICOT C. 1987, *Statistique Spatiale*, Brouillons Dupont, n. 15, Avignon, Groupe Dupont.

TRICOT C., RAFFESTIN C. 1979, *Concentration et connexité ou pour un approche du concept de territorialité*, in *Actes du colloque Géopoint76: Théorie et géographie*, Genève, Groupe Dupont, Universités de Genève et de Lausanne, Avignon, 49-62.

USERY E.L. 1993, *Category theory and the structure of features in Geographic Information Systems*, «Cartography and Geographic Information Systems», 20, n. 1, 5-12.

VERHAEGEN P., MCGLADE J., GILI S., RISCH R. 1995, *Some criteria for modelling socio-economic activities in the Bronze-Age of south east Spain*, in G. LOCK, Z. STANČIČ (edd.), *Archaeolgy and Geographic Information Systems: An European Perspective*, London, Taylor & Francis.

VOIRON CANICIO C. 1993, *Espace, Structures et dynamiques regionales: L'Arc mediterranéen*, «Revue d'Analyse Spatiale Quantitative et appliquée», n. 33, 1-296.

WHALLON R. 1984, *Unconstrained clustering for the analysis of spatial distributions in archaeology*, in H. HIETALA (ed.), *Intrasite Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press.

ABSTRACT

The purpose of this paper is to critically explore the role of Geographical Information Systems in the archaeological research. Currently some archaeologists seem largely captivated by new computing technologies believing that the sophistication of powerful software outputs will lend respectability by itself. In our opinion GIS is merely a set of techniques to visualise and manage large amounts of georeferenced data. Thus there must be other tools to move from *visualisation* to *explanation*, which fall within the domain of *Spatial Analysis*.

The ultimate aim of this paper is to show how we can integrate these already existing tools (geostatistics, intra-site statistical tests, digital image processing, artificial intelligence, etc.) in a GIS framework, in order to move from beautiful images to hard analysis. Finally we criticise the lack of theoretical background in archaeological uses of GIS technology arguing that GIS is the only software and may benefit our research only if we use well defined *archaeological* problems on a well-based theory.

Archeologia e Calcolatori
7, 1996, 327-335

THINKING ABOUT THE SPATIAL ARTICULATION OF THE ARCHAEOLOGICAL RECORD: THE ROLE OF STATISTICAL TECHNIQUES

1. Introduction

In archaeological research, as well as in the rest of the disciplines which study social phenomena and its interaction with natural phenomena, the application of quantitative methods cannot be obviated. Nowadays, aims and research programs of most of the professionals of this discipline imply, necessarily, a quantitative and computerized processing of the empirical record.

The ultimate goal when using quantitative methods, especially statistical techniques, is to improve the objectivation of the record manipulation process, that is, to be more rigorous in the archaeological praxis (Fig. 1). This so-called improvement can be observed, above all, in the standardization of the data description it strengthens the reasoning and the inferences. Once the quantification's importance is accepted as the objectivation vehicle of the observations of the empirical record data, the next step must be to carry out an adequate selection of the quantitative tools to ensure the validity and reliability of the data. A good research design must include rigorous proceedings of observation, recording and systematization of the empirical data.

Undoubtedly, the irruption of quantitative methods and computers in archaeological research has had a considerable impact. Overall, its appearance is clearly positive, even though many imprecisions and inevitable errors have been made. Apparently, a very wide integration has been attained regardless of absurd refusals of minority sectors. However, there are several aspects still to be solved:

1) the professionals of the archaeological discipline must improve their training in their school years in order to integrate and normalize, once and for all, the application of quantitative methods and the use of computers. This improvement would minimize errors and the incorrect usage product of the inexperience and carelessness, or what is more, of the disregard for the instrumental proceedings.

2) the usual tendency to unnecessary sophistication, which has caused the understandable refusal by the less receptive sectors, must be overcome (Barceló *et al.* 1994; Wünsch *et al.* 1995). Along this line, we consider adequate to defend an approach based on the principle "the easier, the better". More simple technical solutions must be given as much priority as possible because, usually, the posed problems are simple too.

3) the reflection about the quantitative methods must be integrated within the frame of the theoretical-methodological general discussion. It is

important to set an adequate connection between the evaluated theoretical aspects and the design of the instrumental methodology. The best solution would be to establish a gradation from theoretical to empirical aspects, without leaving any gap.

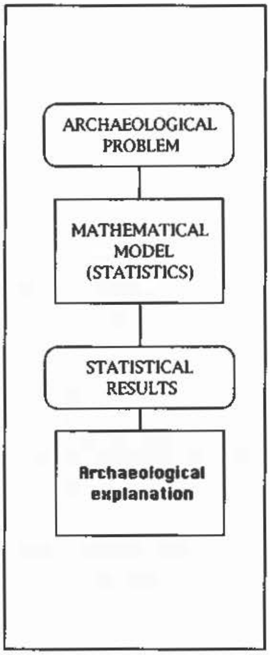
This process of reflection is closely linked to the archaeological praxis, since the nature of the archaeological record is defined as the representative material entity of the social activity. Several properties of the archaeological material record which are prone to be quantified are structured by means of the overlapping, association and recurrence principles (LUMBRERAS 1981). We are going to focus on only those related to the spatial articulation because they enable us to obtain information about the organizational strategies of human communities.

2. FROM THE SPATIAL ARTICULATION OF ARCHAEOLOGICAL RECORD TO THE MANAGEMENT OF THE SOCIAL SPACE

In the archaeology of hunter-gatherer communities a line of research about the management of social space has been clearly defined (WÜNSCH 1991-92, 1992). Synthesizing, this line presupposes that the spatial articulation of the archaeological record reflects aspects which are implied in the management of the social space (Fig. 2). This management includes not only the logistical aspects related to the conditioning and cleaning of the space, but also the articulation among social units, the location of the working processes, the distribution of food and goods, etc. It also affects the development of the general productive process, and therefore, it includes the results of the interaction between the production, distribution and consumption processes.

It is operational to propose an "analytical gradation" of the social space. Such gradation ranges from the most wide category, the territory, as the physical frame in which the dialectics between the human communities vs. environment takes place, to the smallest category, the settlement or occupation place. Generally, and leaving aside terminological nuances, there is a plural approach to the social space resulting from the transformation of the environment through work and the productive activity aimed at the obtention of material goods.

This wide and restricted "spatiality" gradation means to face the notion of record representativeness. As in other research fields, it is essential to evaluate the "archaeological universe" that is being studied in order to decide its validity and degree of reliability. Therefore, we meet the problem of the representativeness of the sample because, undoubtedly, the archaeology professionals study samples of a total "population" of dimensions which are seldom known. Nevertheless, any approach to this wide spatial notion of interaction among complementary settlements can profit from the applica-

1 

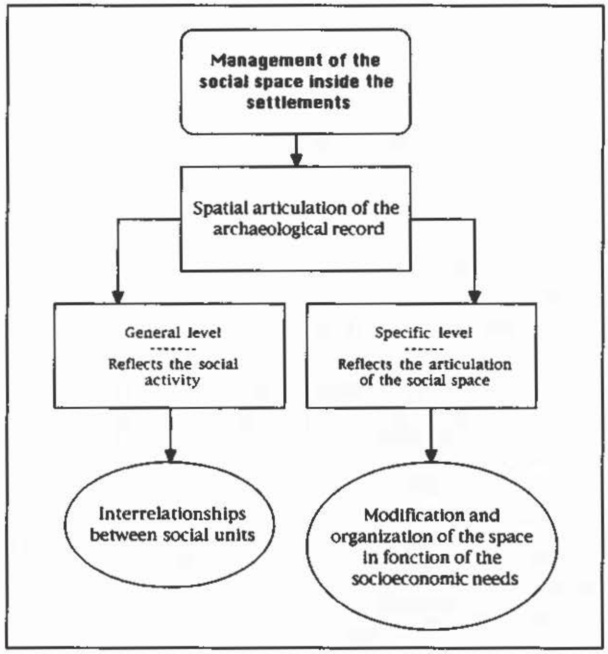
 2

Fig. 1 – The major steps in the application of the quantitative methods, basically statistics, in the archaeological research.

Fig. 2 – The diagram include the most important features included in the concept of the management of the social space.

tion of the probabilistic sampling techniques.

With these techniques the archaeological record is connected to the notions of probability and randomness. This notion of probability which lays upon the probability theory becomes the credibility or confirmation degree in the light of the available empirical evidence. Moreover, it is very useful when establishing the validity degree of an inference or a reasoning. The notion of randomness must be understood as the absence causal determinism. It is obvious, thus, that randomness is essential in relation to the informative potential of the spatial articulation of the empirical record.

The contextual condition of the archaeological record attaches the importance of the association and recurrence principles which bestow relevance and social signification to the analyzed material remains. It is for this reason that a spatial notion restricted to the inner space of the settlements implies a previous reflection about the objectivation possibilities by means of quantification. The problem does not lie so much on the sample's representativeness, if we accept the settlement or occupation place as a significant social context, but on the establishment of a processing design of the record according to the needs and to the informative expectations.

On the whole we see that, from the beginning, it is almost impossible to think about all the different questions implied in the application of quantitative methods in archaeology. Above all, we ought to outline the importance of the randomness concept versus non-randomness. This is a basic concept for emphasize the significant data and eliminate the irrelevant information.

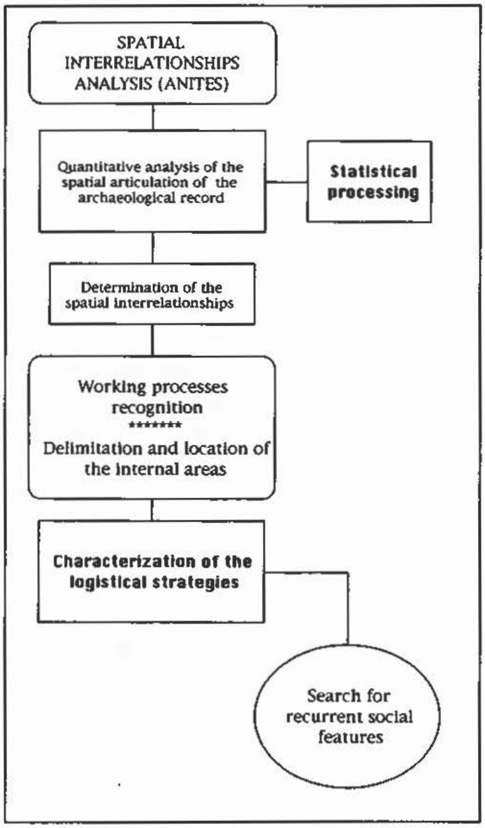
3. The spatial properties of the archaeological record

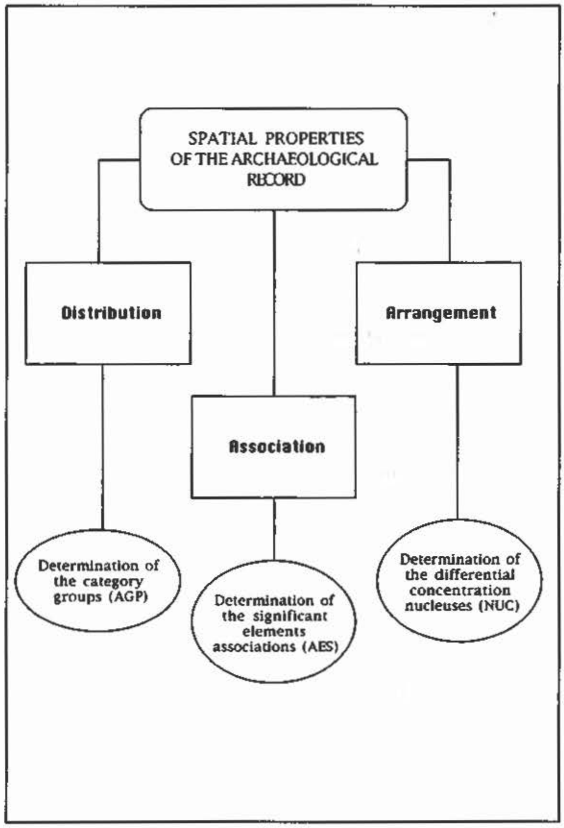
To focus our attention on the study of the spatial articulation of the archaeological record means to define the connection between the archaeological context and the mathematical models, in this case, statistical models which are more suitable for its analysis. We should be able to clearly define the archaeological problem that is to be solved, and subsequently select the statistical models and techniques which are relevant for the obtention of the significant and informative results. For this reason to make the spatial articulation into a study object implies to profit from the spatial properties of the archaeological record.

Our proposal of approach to the study of the spatial articulation is built upon the design of the spatial interrelationships analysis (ANITES). In this previous elaboration process of the methodology, special attention has been given to the application of quantitative methods (Fig. 3). The main idea is to profit from the quantifying possibilities of the spatial properties of the archaeological record. In order to do so, several statistical techniques have been selected. With these techniques we can evaluate aspects related to causality, probability, signification and randomness (Wünsch 1989a, 1989b, 1991, 1991-92, 1992, 1994, 1995).

We have given special attention to the elaboration of the instrumental methodology, which is understood as the whole set of instrumental proceedings implied in the processing of the empirical record. In the process of reflection about the operational design, several relevant features and its eventual relationships with mathematical models are determined. Namely, we have selected three of the most relevant and informative properties of the archaeological record (Fig. 4):

1) the <u>distribution</u> of the material remains. The aim is to measure the degree of difference with respect to the randomness of a distribution of remains considered as points within a limited three-dimensional space. This measure based on mutual distances between points allows us to discriminate the non-random patterns and keeps it from being causally interpreted. The approach to the non-random patterns is carried out by means of the nearest-neighbour analysis remodeled in 3D (Wünsch 1994, 1995). This test determines the existence of random and non-random patterns, by discriminating in the latter case its tendency to clustering or scattering.

3




4

Fig. 3 – The grafic shows the main aspects of the operational design of the ANITES methodology.

Fig. 4 – The diagram shows the three spatial properties of the archaeological record studied by the ANITES methodology.

The determination of the distribution patterns is the first, but not the most important, step in the characterization of the spatial properties of the archaeological record. In fact, it only provides us with an example of non-homogeneity, nevertheless, the inferences based on these patters must be alternative and accurate. However, this determination is basic because it facilitates the study of the non-random patterns and, consequently, the search of causal explanations.

We ought to take into account that it is not advisable to interpret mechanically the patterns in terms of its implications respecting the social activity that they reflect. For instance, the scattered distribution needs not be interpreted as a reflect of an anthropical or postdepositional alteration without taking into account other alternative hypotheses. Moreover, a clustered distribution not always shows the exact location of the working processes, considering the tools, remains and refuse accumulations, with more or less disturbing postdepositional impact.

We must insist on the fact that the determination of the distribution pattern only presents a dim view about the spatial articulation. Its validity is directly conditioned to the analytical determination of the significant category groups (AGP), if they exist. In our design of the ANITES, we have used as a measure of the grouping a critical distance of rupture (WÜNSCH 1995) from which the different groups of points are formed. These groups are a first general characterization of the basic tendencies about the distribution of the material remains within the analyzed space (the excavated surface).

2) the arrangement of the material remains. The aim is to measure the degree of spatial autocorrelation, that is, the intensity of the existent relationship between the adjacent remains or values. Parallely, it can inform us about the signification of eventual tendencies to an specific location, such is the case of the semicircular arrangements or the aligments. The measure of the arrangement must be understood as a complementary view which is necessary in respect to the distribution and, moreover, cannot be regarded as an alternative. In general, it provides new information and enables a greater comprehension of the spatial articulation of the archaeological record. In its minimal informative level, it enables us to prove once more the non-homogeneity of the record (the non-randomness) and also to determine the tendency to clustering or to scattering.

In the case of the processing of three-dimensional remains there are not simple measures and the quantification is difficult to carry out. For the processing of non three-dimensional remains we have selected the application of spatial autocorrelation tests, namely the Moran's I coefficient (WÜNSCH 1995).

As in the above mentioned case, the determination of the arrangement pattern of the remains has relative interest by itself. Its informative potential and validity are conditioned by an analytical determination of the differential concentration nucleuses (NUC), and by means of a comparisons grid, square by square, using the Chi-squared test (WÜNSCH 1995). And hence, it is possible to obtain a concise description of the basic arrangement tendencies of the material remains as far as location and/or concentration degree is concerned.

When the applicability conditions of the statistical techniques are proper, especially concerning the minimal number of effectives, the results of the distribution patterns and those of the arrangement patterns must be complementary and, above all, they should not be contradictory. In case there is a significant pattern of clustered distribution, we would be able to state parallely a clustered arrangement pattern, that is, a clear tendency to concentration.

3) the spatial association of the material remains. The most immediate aim is to measure the degree of spatial association of the selected analytical categories. We try to determine the signification of the contextual relationship based on the differential location within the analyzed space. While the distribution and arrangement measures give information on the categories of the material remains that have been obtained at an individual level, the measures of spatial association are focused to state the interrelationship among categories.

This analytical approach to the spatial association of the archaeological record enables us to obtain a global and strong view of the spatial articulation. It is also the base for the isolation of the significant elements associations (AES). In order to determine these associations we have selected several techniques that furnish more details about complementary aspects. On the one hand, we have selected the association or similitude tests, namely the Jaccard's I coefficient and the Phi-coefficient (Wünsch 1995). These tests are adequate for the measuring of the relation degree between those categories which are studied two by two, at a presence/absence level. This is an important aspect of the operational design of the ANITES because qualitative categories can be processed.

Furthermore, we have selected an approach which offers a wider view of the relationship among categories and the basic excavation grid that is used to delimit the locations. In most of the cases, the percentages table of the Lien offers a fairly close view of the most significant presences or absences. However, very often the application of the correspondence analysis (AFC) helps to strengthen the already observed tendencies and/or to extract complementary data because it relates all the variables each one with another (Wünsch 1995).

The final determination of the significant elements associations is based on the integrated evaluation of the different analyzed aspects. It takes into account not only the differential locations or concentrations within the analyzed space but also the significant associations among categories. We end up having a final synthesis of the spatial articulation of the archaeological record.

4. Discussion

The joint evaluation of the three spatial properties of the archaeological record is the analytical base of the spatial interrelationships concept which defines the methodology of ANITES. In any case, the problem is to establish the nexus between the statistical signification and the archaeological signification, as the base for the explanation of the social phenomenon. Since this explanation intends to acquire a generalizing character, it is important to overcome the irrelevant particularism and focus the attention to recurrence.

Another important aspect that is to be taken into account is that those patterns determined using statistical processing are not necessarily preexistent in the data. On the contrary, they respond to an analytical approach which is only valid within the designed theoretical-methodological frame. If we accept a 'codification' of the archaeological record, which gives way to the existence of a spatial articulation which is analytically discriminable, we can extract analytical patterns that can be used as a comparative model for the search of recurrent social features.

This is the most outstanding utility of the spatial interrelationship patterns (PIE) which are configured as the synthesis of the delimitation and location of the internal areas of the studied settlements (WÜNSCH 1992). These patterns must serve as synthesizing hypotheses of the main tendencies of the management of the social space. They are obtained through an analytical characterization of the spatial articulation of the archaeological record. Its real relevance will only stem from the later prove of significant recurrences.

The first exploratory applications and the controlled experimental processing of the ethnoarchaeological records (WÜNSCH 1993) seem to evidence the operativity of the statistical data processing and hence the interest of the ANITES methodological proposal.

GERMÀ WÜNSCH
Consell Superior d'Investigacions Científiques
Inst. Milà i Fontanals – Lab. d'Arqueologia
Barcelona

BIBLIOGRAPHY

BARCELÓ J.A. *et al.* 1994, *The easier, the better. The role of statistics in archaeological research*, in I. JOHNSON (ed.), *Methods in the Mountains, Proceedings of UISPP, Comm. IV meeting, (Sidney, Australia, 1993)*, Sidney Univ. Archaeological Method Series, 2, 43-48.

LUMBRERAS L.G. 1981, *La arqueología como ciencia social*, Lima, Ediciones PEISA.

WÜNSCH G. 1989a, *Cuantificación en arqueología: el análisis de las interrelaciones espaciales de los elementos arqueológicos*, «Revista de Arqueología», 95, 5-9.

WÜNSCH G. 1989b, *La organización interna de los asentamientos de comunidades cazadoras-recolectoras: el análisis de las interrelaciones espaciales de los elementos arqueológicos*, «Trabajos de Prehistoria», 46, 13-33.

WÜNSCH G. 1991, *Del estudio de la organización espacial al análisis de las interrelaciones espaciales de los elementos arqueológicos*, in A. VILA (ed.), *Arqueología*, Col. Nuevas Tendencias 19, Madrid, CSIC, 195-208.

WÜNSCH G. 1991-92, *La organización del espacio interno de los asentamientos de comunidades cazadoras-recolectoras: Revisión crítica y alternativas*, «Xama», 4-5, 161-218.

WÜNSCH G. 1992, *Aproximación a las estrategias organizativas de las comunidades cazadoras-recolectoras*, «Archidiskodon», 2, 1-36.

WÜNSCH G. 1993, *Aplicación del análisis de las interrelaciones espaciales (ANITES) sobre datos etnoarqueológicos: El sitio Túnel VII (Tierra del Fuego, Argentina)*, Paper presented to the *Segundas Jornadas de Arqueología de la Patagonia, Puerto Madryn* (Chubut, Argentina), 1-15.

WÜNSCH G. 1994, *Modifying the nearest-neighbour analysis for 3D*, «Archaeological Computing Newsletter», 39, 9-15.

WÜNSCH G. 1995, *The spatial interrelationships analysis and its simple statistical tools*, in J. HUGGET, N. RYAN (eds.), *Computer Applications and Quantitative Methods in Archaeology, CAA'94*, Glasgow, BAR International Series, 600, 75-80.

WÜNSCH G. *et al.* 1995, *Dissecting the palimpsest: an easy computer-graphic approach to the stratigraphic sequence of Túnel VII site (Tierra del Fuego, Argentina)*, Paper presented to the *Computer Applications and Quantitative Methods in Archaeology, CAA'95*, 1-10.

ABSTRACT

The aim of this paper is to present a short reflection about the connection between the spatial articulation of the archaeological record and the role of the statistical techniques. The basic idea is to process three spatial properties of the archaeological record: the distribution, the arrangement and the association. We include this idea inside the theoretical-methodological framework of the ANITES proposal. Briefly, we present the informative potential of these properties and the statistical techniques selected.

Archeologia e Calcolatori
7, 1996, 337-355

STUDIO DEI PROCESSI FORMATIVI DEL DEPOSITO E RICOGNIZIONE DI CONFIGURAZIONI SPAZIALI IN INSEDIAMENTI ALL'APERTO: ANALISI STATISTICHE DELLE EVIDENZE ARCHEOLOGICHE DI TERRAGNE (MANDURIA, TARANTO)

1. Un commento introduttivo

L'analisi dei *site formation processes* è certamente uno dei campi di indagini più sviluppati dell'archeologia degli ultimi trenta anni, con una produzione bibliografica estremamente ampia, di cui non è certamente possibile effettuare una trattazione storiografica in questa sede. L'impulso fornito dalla *New Archaeology*, e il forte sviluppo dell'etnoarcheologia rappresentano probabilmente gli elementi fondamentali alla base dell'enorme produzione scientifica, prevalentemente anglosassone e statunitense (e.g. Binford 1981; Schiffer 1987).

In Italia, eccetto alcuni lavori gravitanti perlopiù nella sfera etnoarcheologica, lo sviluppo e l'interesse per l'analisi dei processi formativi è stato essenzialmente fornito da ricercatori "padovani" (e.g. Leonardi 1992), in uno scenario prevalentemente di reazione (De Guio 1988) al concetto dominante di unità stratigrafica in senso "harrisiano" (Harris 1979), sebbene approcci geoarcheologici all'analisi dei depositi preistorici, costituiscano in realtà significative esperienze *ante litteram* sul *site formation processes debate* (e.g. Cremaschi 1990). Ad una analisi della letteratura teorica e analitica sui processi formativi in senso post-harrisiano, sembra emergere comunque, al di là delle differenze teoriche e negli approcci operativi, una spiccata caratterizzazione settentrionale. Questo interesse sperequo crea di fatto, già all'inizio degli anni '80, una robusta forbice in tale segmento di studi tra Nord e Centro-Sud Italia: tale forbice si concretizza, tra l'altro, nella edizione della rivista «Archeologia Stratigrafica dell'Italia Settentrionale» (di breve vita: un solo numero nel 1988). La formalizzazione del predominio teorico del Nord Italiano, e nello specifico padovano, si evince infine dalla produzione, in tempi più recenti, degli atti del convegno "Processi formativi della stratificazione archeologica", editi da Giovanni Leonardi (1992). In Italia meridionale tale dominio di studi è pressoché assente: una possibile eccezione sembra fornita dalle ricerche svolte nel sud-est tarantino, dove particolare enfasi alle problematiche della *site formation* è stato posta nell'ambito delle indagini di scavo nel sito che discutiamo (Gorgoglione *et al.* 1991: in particolare pp. 65-75).

Quello che mi sembra opportuno sottolineare, aldilà delle varie posizioni che rendono multiforme la problematica della *site formation*, è come, nell'ambito della vasta produzione scientifica, siano in realtà molto rari i contributi analitici (i *tools*, per intenderci) esportabili in altri contesti di ricerca, e come, al contrario, ricorrano principalmente gli approcci descrittivi.

2. Il sito: le domande e i problemi

Il progetto di ricerca sul sito di Terragne (Manduria, Taranto) rientra nell'ambito delle attività di tutela e valorizzazione che la Soprintendenza Archeologica della Puglia, sede di Taranto, effettua da molti anni nel sud-est tarantino, dirette dalla Dott.ssa Gorgoglione (Fig. 1). Le indagini sono state affrontate con analisi multidisciplinari, cercando di potenziare al massimo le possibili informazioni del sito.

In un certo senso, tra le motivazioni fondamentali di questa ricerca vi era da un lato quella di raccogliere la sfida di Binford (1981), tentare di utilizzare quindi i *distorted stuff* e non cercare "rare Pompei"; dall'altro, quella di provare a ridurre, almeno in parte, la forbice Nord-Sud relativa proprio a questi problemi.

Questo è stato possibile innanzitutto per le strategie di collaborazione all'interno del gruppo di ricerca, in una reale ottica interdisciplinare. Un rapporto ottimizzante tra costi e benefici è stato possibile sullo scavo grazie a massicce e sistematiche forme di campionamento, coordinate da G. Fiorentino. Il basso costo di *software* dedicati all'analisi dei dati archeologici ha reso accessibili tali strumenti (anche a personale scientifico non strutturato), consentendo quindi una elaborazione quantitativa multivariata e multidimensionale di una grande quantità di dati, almeno per quanto concerne un sito antico olocenico dell'Italia meridionale (Fig. 2). Il lavoro che qui presento è una trattazione specifica di un più ampio lavoro effettuato in collaborazione

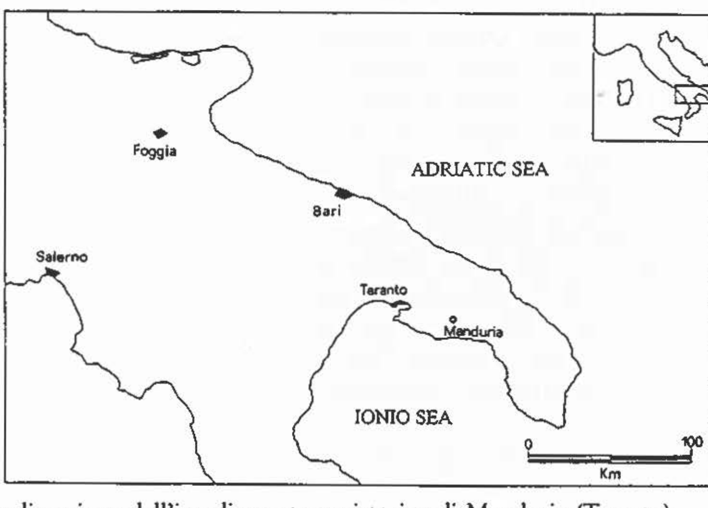


Fig. 1 – Localizzazione dell'insediamento preistorico di Manduria (Taranto).

con Fiorentino e Gorgoglione, i cui dati e risultati sono già editi altrove (GORGOGLIONE *et al.* 1995; DI LERNIA, FIORENTINO c.s.).

Lo scopo è quindi di tentare di approfondire alcuni aspetti archeologici, già interpretati in senso quantitativo, valutandone analogie e differenze in termini qualitativi.

Il sito di Terragne è caratterizzato da un deposito archeologico all'aperto poco spesso, in una zona periferica della città di Manduria. Questo insediamento si caratterizza per la presenza di due livelli archeologici sovrapposti riferibili alla fine del Mesolitico (US 5), datato al radiocarbonio 7260±70 bp, e a fasi avanzate del Neolitico Antico (US 3), datato 6830±70 bp (Fig. 3). Il deposito è chiuso da due unità stratigrafiche (US 1 e 2), con caratteri di arativo recente e subrecente. L'estensione di scavo è di circa 30 mq per il livello più antico "Mesolitico-US 5", e di 40 mq per il livello "Neolitico-US 3". Va precisato che, in relazione ai pesanti interventi antropici, l'area di scavo rappresenta sostanzialmente il deposito sopravvissuto alle arature.

Il deposito si è formato su un suolo di tipo rendzina ed è affetto da profondi fenomeni turbativi postdeposizionali, sia attivi che passivi: le analisi chimico-fisiche dei sedimenti e lo studio micromorfologico di sezioni sottili indisturbati di suolo hanno messo in rilievo la notevole omogeneità delle unità stratigrafiche, e un modesto rimaneggiamento del terreno, già a partire dalla US 2 (CREMASCHI, DI LERNIA 1995). Le evidenze della distribuzione stratigrafica e della configurazione spaziale risentono fortemente di tali processi.

Le caratteristiche dello scavo, effettuato inizialmente come intervento di emergenza, hanno orientato la nostra raccolta dati organizzata per quadrati di 1 metro di lato; per alcuni indicatori archeologici, con valori dimensionali consistenti (ceramica, faune, pietre di dimensioni >5 cm), è stata effettuata la localizzazione per coordinate spaziali; l'industria litica, proprio a causa della sua caratterizzazione ipermicrolitica, risulta invece sempre raccolta per quadrati, essendo stata recuperata pressoché totalmente al setaccio in acqua (2 mm).

I dati così raccolti sono stati informatizzati in appositi *databases*, ed elaborati con *packages* statistici dedicati. Tra i vari *software*, sono stati utilizzati ARCOSPACE, sviluppato da H.P. Blankolhm (University of Aahrus), e il BASP, sviluppato da I. Scollar (Bonn Archaeological Statistic Package), prelevabile all'indirizzo http://www.uni-tuebingen.de/uni/afj/basp.html.

Il tentativo di analisi è stato quindi quello di superare il piano esclusivamente descrittivo, cercando di identificare sistemi di *pattern recognition* affidabili, e isolare conseguentemente elementi archeologici diagnostici. La possibilità di ottenere a basso costo *software* dedicati, e lo sviluppo di tecniche di analisi che gestiscono informazioni su griglie di dati, e non solamente su coordinate spaziali, permettono infatti di recuperare informazioni raccolte in scavi di recupero e salvataggio, dove la migliore documentazione possi-

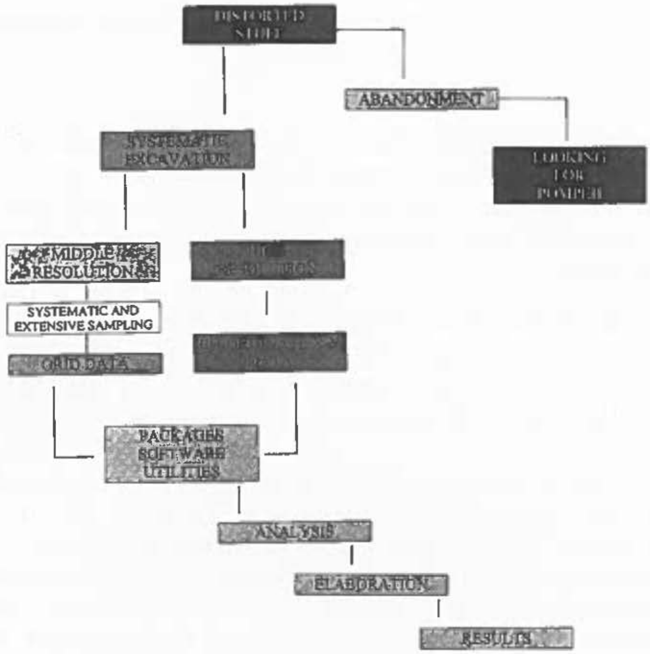


Fig. 2 – Diagramma di flusso del percorso di ricerca effettuato: l'effettuazione di un campionamento sistematico ed estensivo può notevolmente migliorare la qualità delle informazioni raccolte in un intervento di emergenza. L'utilizzo inoltre di dati su griglia, con l'interpolazione di appositi *software*, permette un ulteriore incremento del livello di risoluzione.

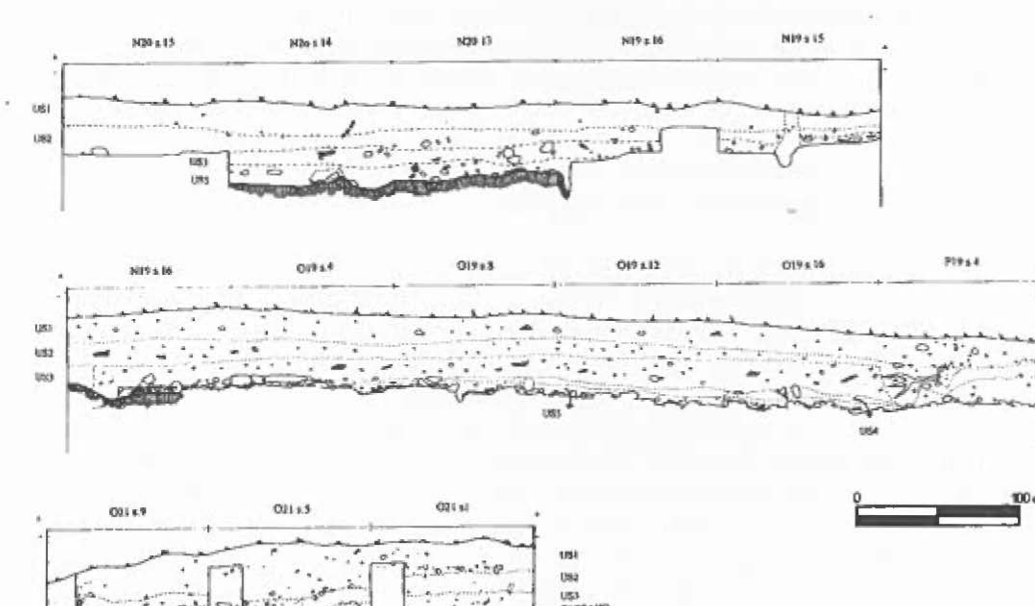


Fig. 3 – Sezioni stratigrafiche "generali" dello scavo estensivo 1988-1991 a Terragne (da Di Lernia, Fiorentino 1995).

bile è spesso solamente quella, appunto, per griglie di 1 metro quadrato. L'assenza di una documentazione di scavo per coordinate spaziali, assai diffusa in interventi con caratteri di emergenza, o con scarsi fondi a disposizione, può e deve essere superata grazie al contributo che tali tecniche di analisi offrono. Non a caso si registra, nella letteratura specificatamente dedicata alla *intrasite spatial analysis*, una enfatizzazione sulla validità dell'utilizzo delle unità di griglia (tra gli altri Blanhkolm 1991), come si evince peraltro dallo sviluppo di *utilities* di trasformazione dedicate (GRID, GRIDI...).

In questa sede vorrei discutere alcuni punti di maggiore interesse, sia per le potenzialità di inferenza, sia su un piano di discussione generale, piuttosto che ripercorrere tutte le tappe operative che hanno caratterizzato la ricerca.

3. Analisi verticale

Per quanto concerne gli aspetti di distribuzione verticale, l'esigenza di analizzare i processi formativi e di tarare i fenomeni postdeposizionali è stata affrontata valutando quantitativamente gli indicatori archeologici, e pesandone gli specifici livelli di correlazione. Ceramica e industria litica richiedono un commento specifico, al fine di analizzare il comportamento del livello neolitico US 3, e i suoi rapporti con il pacco di arativo UUSS 1-2, formatisi a sue spese.

Dei vari indicatori archeologici, certamente la ceramica è tra i più utili e potenti, anche per le notevoli capacità di conservazione che essa offre (per es. Leonardi, Pracchia, Vidale 1989; Orton, Tyers 1990). Accanto alle tradizionali categorie di analisi, è stata messa a punto una specifica codifica morfologica, nell'ipotesi che il coccio possa essere paragonato ad una unità sedimentaria, sottoposta pertanto a leggi specifiche che ne regolino il comportamento.

Lo schema di riferimento adottato e modificato è quello generalmente utilizzato per descrivere le parti scheletriche del suolo in pedologia (Sanesi 1977; di Lernia, Fiorentino 1995). Tale strumento, unitamente alle codifiche delle caratteristiche delle fratture, è servito per distinguere nettamente i processi che hanno contraddistinto la "trasformazione" del livello di occupazione US 3, e la "formazione" del livelli di arativo US 1 e 2. Per gli aspetti dimensionali si è preferito utilizzare un criterio di superficie espresso in cm^2: tale sistema si è dimostrato più utile e flessibile.

A Terragne, grado di elaborazione morfologica, caratteri delle fratture e valori dimensionali sono in funzione dei processi formativi (e delle attività di disturbo post-deposizionali passive e attive), che hanno caratterizzato questo sito sin dal suo seppellimento.

Come è osservabile nelle Figg. 4 e 5, si osservano infatti comportamenti differenti tra il pacco arativo e la US 3, e, all'interno del pacco arativo, tra US 1 e US 2.

È interessante notare come sottoponendo ad analisi fattoriale queste variabili dell'indicatore ceramico (Fig. 6), si osservi una "maggiore" similarità di comportamento tra US 1 e US 3, mentre la US 2 presenti raggruppamenti diversi. Nella mia interpretazione, tale comportamento dovrebbe essere dovuto ai differenti processi di riesumazione ai quali è sottoposta la US 1, che è contraddistinta da frammenti ceramici le cui caratteristiche morfologiche subiscono probabilmente un continuo ravvivamento legato proprio all'impatto arativo (O'Brien, Lewarch 1981; Schiffer 1977).

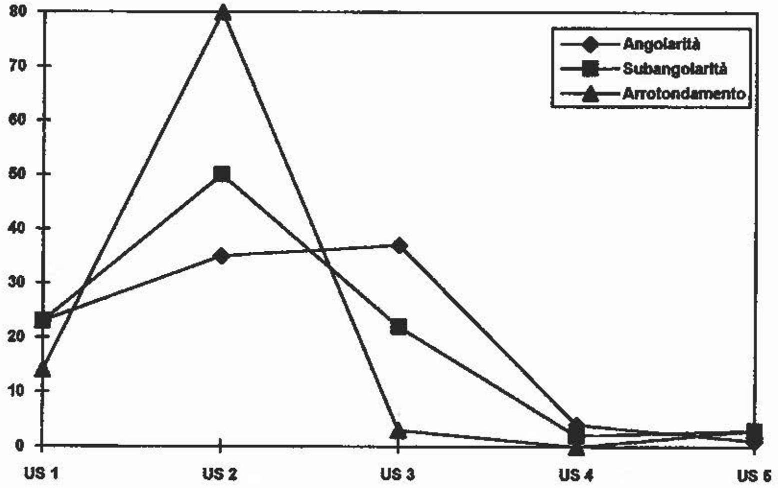
Semplici osservazioni descrittive del comportamento di queste variabili mettono già in evidenza le differenze formative a cui sono state sottoposte. Le caratteristiche dei frammenti ceramici, e loro potenziale ricaduta interpretativa, possono essere quindi così riassunte (Tabella 1):

UNITÀ	EVIDENZA ARCHEOLOGICA	INTERPRETAZIONE
US 1	-media quantità di frammenti -eterogeneità classi dimensionali -morfologia tabulare "smoothed" -pari incidenza di fratture fresche ed elaborate -incidenza giacitura verticale	Riesumazione attiva: impatto arativo recente e subrecente
US 2	-alta quantità di frammenti -standardizzazione classi dimensionali (piccole) -morfologia isodiametrale "smoothed" -netta incidenza fratture elaborate -giacitura "casuale"	Riesumazione e seppellimento ripetuti
US 3	-quantità media di frammenti -eterogeneità classi dimensionali (medio-grandi) -morfologia isodiametrale angolare -netta incidenza fratture fresche -netta incidenza giacitura orizzontale	Discreta stabilità della superficie: lenta, continua cessione di materiali

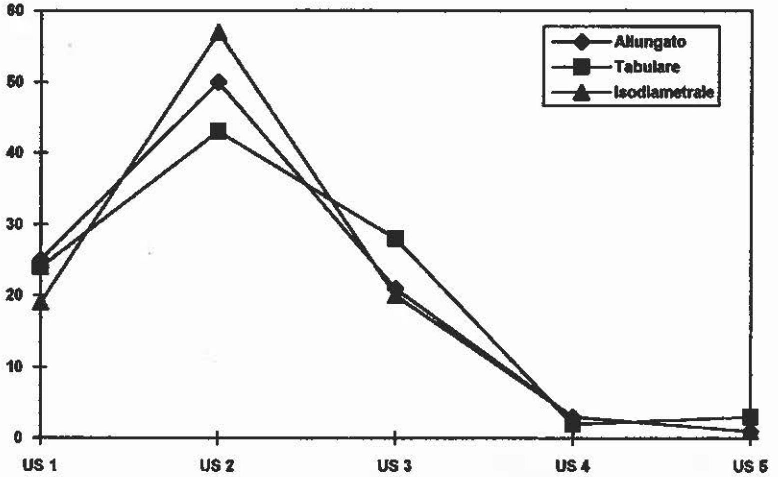
Tabella 1

L'industria litica presenta caratteri forse meno stimolanti, in relazione alla sua resistenza ad una rielaborazione morfologica, e per la manifesta impossibilità (o estrema difficoltà) nell'effettuare *refitting* dei nuclei per controlli sui movimenti verticali dei manufatti (e.g. Villa 1982). È interessante tuttavia notare come due caratteri, spesso utilizzati come elementi "culturali" in letteratura, possano in realtà essere pesantemente sottoposti a processi elaborativi postdeposizionali: mi riferisco all'indice di allungamento e carenaggio.

Nei livelli di arativo non si registrano infatti pezzi allungati, cosi come supporti iperpiatti: la frammentazione a cui comunque è sottoposta l'indu-



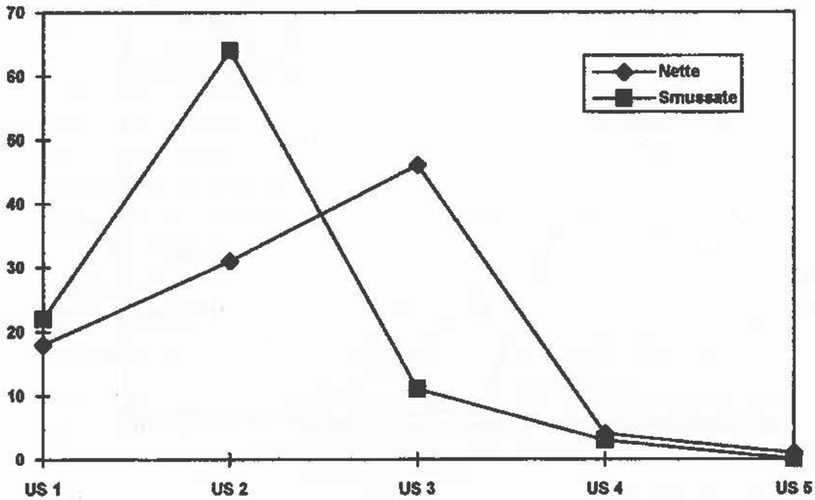
a



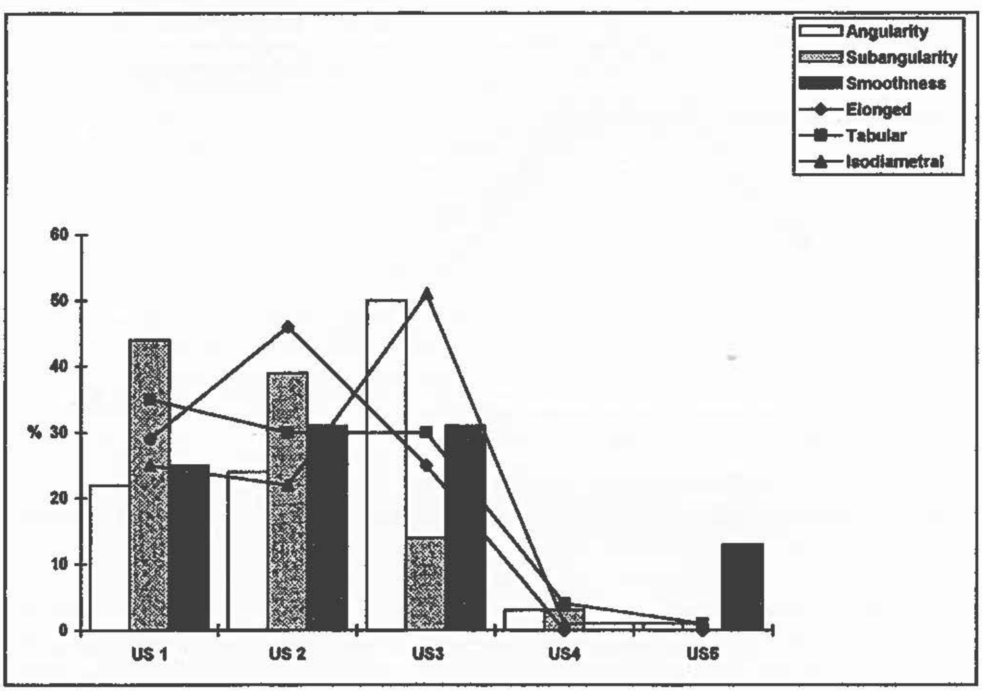
b

Fig. 4 – Frequenze relative della tendenza all'arrotondamento (alto) e all'allungamento (basso) dei frammenti ceramici, sulla base della codifica morfologica "pedologica".

stria indica nei livelli di arativo una vita diversa. Il senso di queste osservazioni è, a mio avviso, isolare degli indicatori precisi di processi postdeposizionali: in caso quindi di ulteriore seppellimento, una eventuale futura valutazione dovrebbe comunque tenere conto di tali aspetti, e caratteri che potrebbero essere interpretati come scelte tecnologiche precise o come fenomeni postdeposizionali *ab antiquo* (e quindi praticamente sin-deposizionali), sono in re-



a



b

Fig. 5 – Frequenze relative delle fratture (alto); combinazione delle frequenze relative dei valori di arrotondamento e allungamento dei frammenti ceramici (basso).

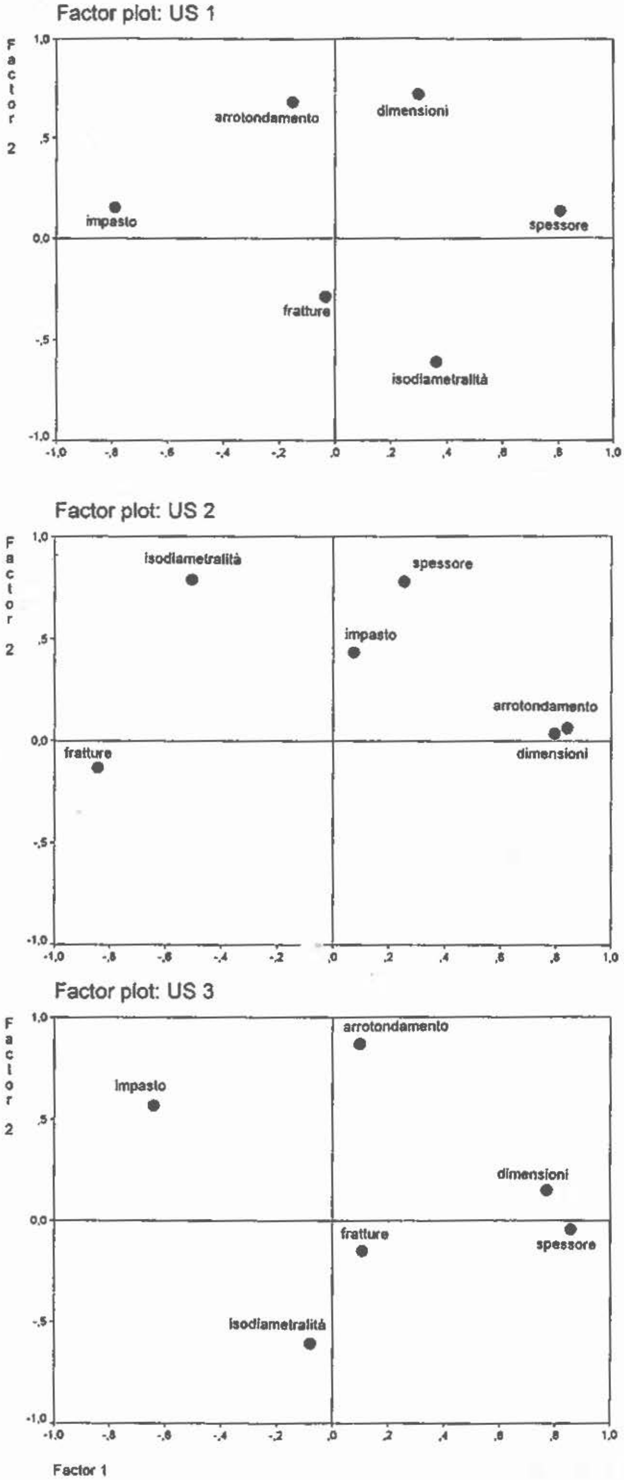
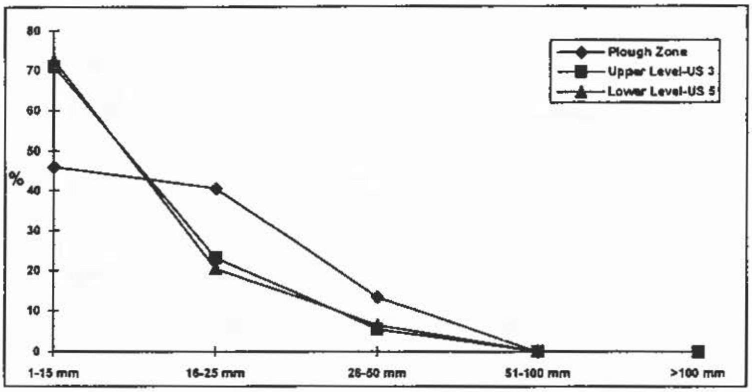


Fig. 6 – Analisi fattoriale delle variabili utilizzate per lo studio della ceramica, operata con il pacchetto SPSS.

a


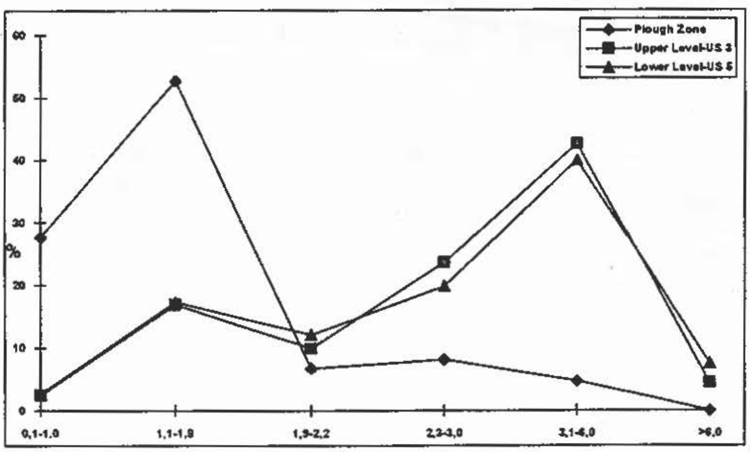
b


Fig. 7 – Frequenze relative delle classi dimensionali (alto) e dell'indice di carenaggio dei manufatti litici (basso).

altà il frutto di altri agenti genetici (nel caso di Terragne l'impatto antropico recente). Infine, come già osservato, l'industria reagisce in misura minore ai processi di riesumazione, per il concetto di *size effect*, evidenziato da O'Brien e Lewarch (1981): le piccole dimensioni infatti risultano in qualche modo meno sensibili a tali disturbi (Fig. 7a).

4. Analisi spaziale orizzontale

L'analisi delle configurazioni spaziali è stato effettuata sui livelli di occupazione del sito (Superiore-US 3 ed Inferiore-US 5). Alcuni aspetti sono stati utilizzati per valutare l'impatto arativo sui livelli superficiali (di Lernia, Fiorentino 1995); in questa sede trattiamo invece i problemi legati alla conservazione di aree specifiche, strutturali o funzionali.

Anche in questo caso, la semplice valutazione descrittiva delle frequenze permette una serie di considerazioni. La buona corrispondenza di alcune categorie di materiali, che qui non affrontiamo in dettaglio, ha permesso di definire alcuni punti salienti:

1) differenziazione di due aree (indicatori ceramica e intonaco) con un "corridoio" centrale, nella US 3 (Fig. 8)
2) all'interno di questo "corridoio", la sovrapposizione delle frequenze di fauna e industria litica alternativa alle aree sopra indicate (Fig. 8)
3) la sovrapposizione delle frequenze di fauna e industria litica nella US 5 (Fig. 9)

Queste particolari configurazioni sono state ulteriormente testate, nel primo caso per verificare il carattere delle aree differenziate, nel secondo per valutare se la sovrapposizione di fauna e litica rifletta una configurazione specifica o se debba invece essere relazionata a giaciture secondarie (discariche, pulizia).

Al primo punto, analisi di correlazione multivariata evidenziano ovviamente un raggruppamento significativo di ceramiche, pietre e intonaco, e specificatamente nella porzione orientale dello scavo. La più interessante sembra l'analisi di *Uncostrained Clustering* (*Ward's method*: WHALLON 1984): come è noto, essa consiste in una serie di procedure che permettono di raggruppare gli oggetti gerarchicamente, minimizzando la varianza tra i *clusters*. Sono stati considerate 5 categorie di oggetti: frammenti ceramici (>16 cmq); industria litica; frammenti di intonaco; resti faunistici; pietre (>10 cm).

Come aspettato, si osserva una brusca caduta del coefficiente di significatività dopo 5 *clusters*: al di là si osserva una omogeneizzazione puntiforme che di fatto identifica associazioni a sé stanti. Osservando i primi raggruppamenti identificati, è stata scelta una visualizzazione su 5 *clusters*, che si sovrappongono per composizione interna e configurazione spaziale sulle mappe "smussate" degli indicatori archeologici (Fig. 10a, b). I *clusters* 2 e 3 sono in massima parte caratterizzati da pietre ed intonaco, mentre il cluster 5 è rappresentato da resti faunistici, industria litica e fauna.

I *clusters* presentano quindi una distribuzione spaziale che ben si accorda con le semplici analisi descrittive: sembrerebbe un esempio in cui i caratteri di configurazione spaziale non sono molto "nascosti".

In realtà, tali analisi sono state effettuate considerando l'US 3, un classico strato "sottile" (spessore massimo 10-12 cm), come un episodio sedimentario uniforme, che potremmo definire in un certo senso "unità stratigrafica minima" (*minimal units of deposition*: SCHIFFER 1987, 266, ma "quale" è l'unità minima?...).

Ad un livello di scavo, e di documentazione, eseguito con dettaglio maggiore, realizzato con mappe di frequenza della US 3 su micro-tagli di 2 cm, si possono osservare comportamenti interessanti valutando i semplici valori delle densità di frequenza: questi comportamenti sembrano possano essere collegati a specifiche dinamiche di crollo, come l'alternativa presenza/assenza

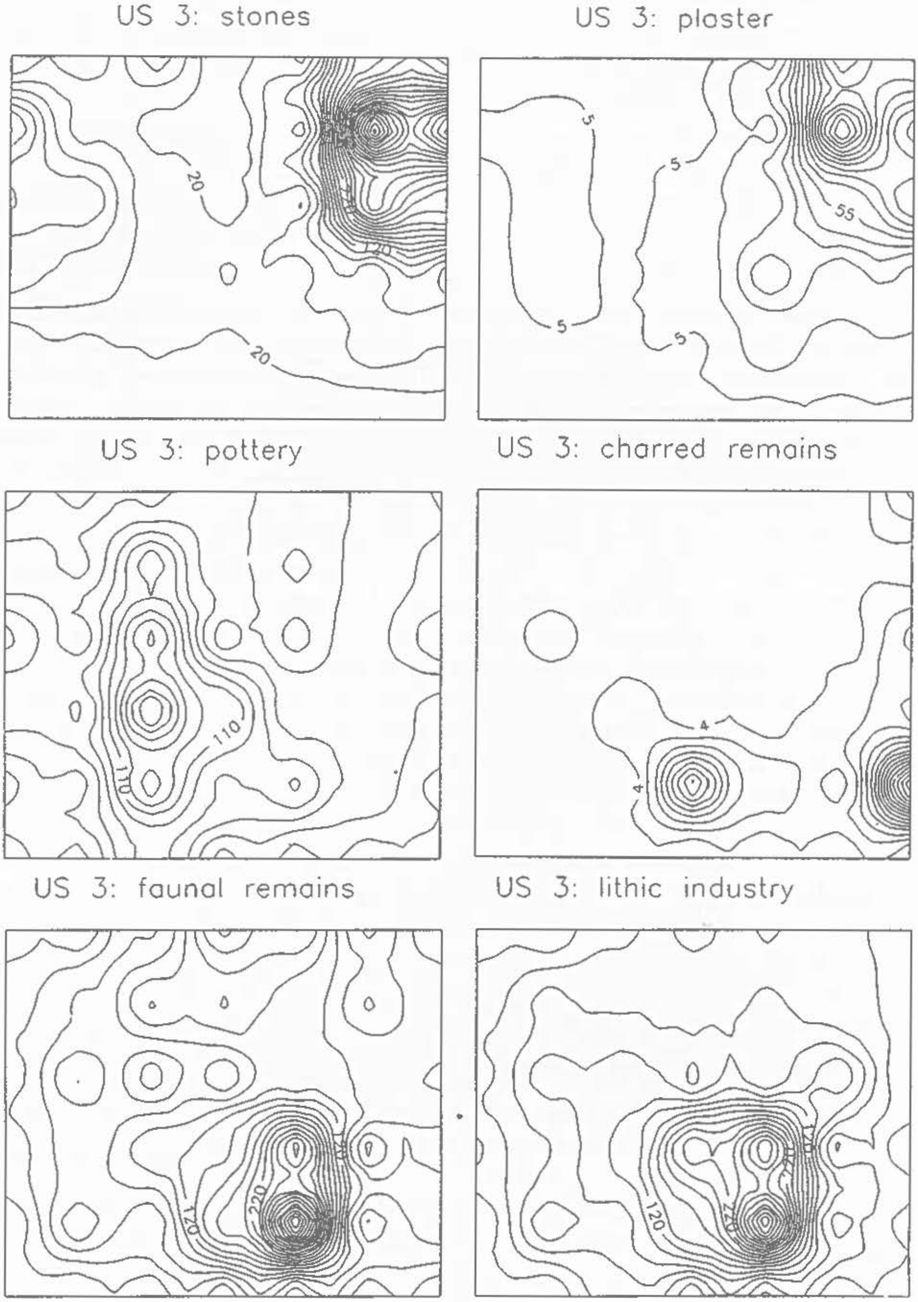


Fig. 8 – Piante di distribuzione delle evidenze archeologiche "smussate" (frequenze assolute) del livello Superiore Neolitico – US 3.

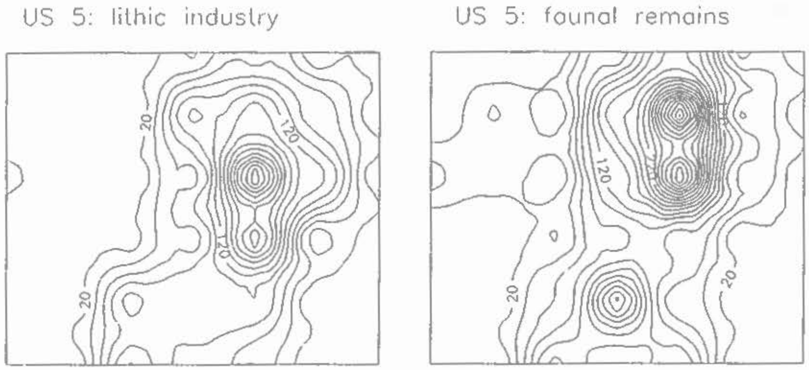


Fig. 9 – Piante di distribuzione delle evidenze archeologiche "smussate" (frequenze assolute) del livello Inferiore Mesolitico – US 5.

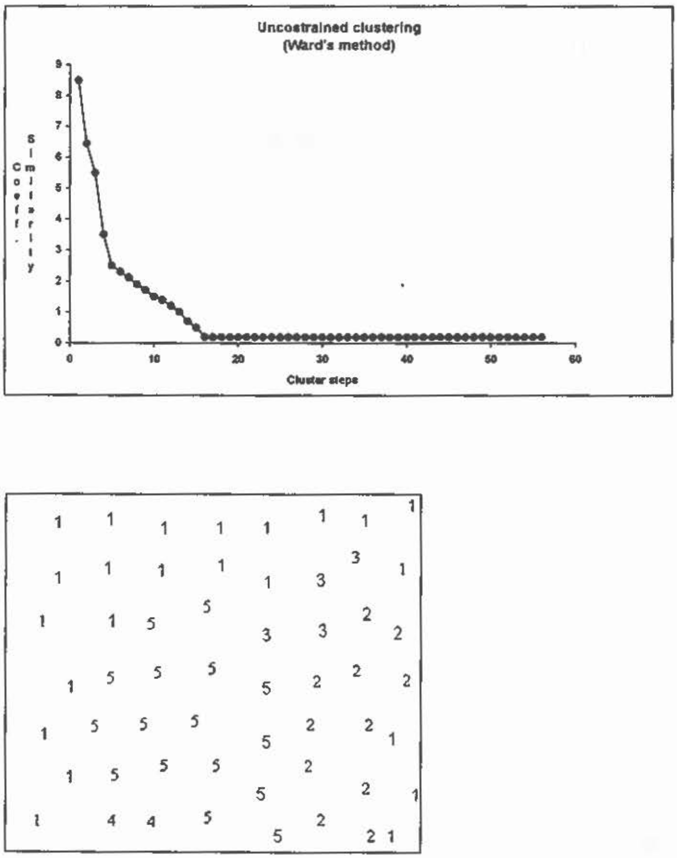


Fig. 10 – *Uncostrained Clustering Analysis* di cinque categorie di evidenze archeologiche (ceramica > 16 cmq; pietre; intonaco; industria litica; resti faunistici). In alto è visibile la curva della formazione dei *clusters* in relazione al coefficiente di similarità. È stata scelta la soglia di cinque *clusters* (vedi nel testo), visualizzata nella figura in basso. È stato utilizzato il pacchetto ARCOSPACE, sviluppato da Blankholm (1991), editando però i grafici con altri programmi.

di ceramica e intonaco in differenti aree dello scavo lasciano pensare (Fig. 11).

Quello che è utile sottolineare è la differenza di interpretazione, sia in termini quantitativi che qualitativi, che uno scavo di dettaglio offre: in realtà, le mappe di frequenze realizzate su unità stratigrafiche grossolane rappresentano generalmente una sommatoria di situazioni, il cui esito è spesso in contrasto con la reale caratterizzazione formativa del deposito e del record archeologico nel suo complesso. Complesse elaborazioni multivariate multidimensionali, se non effettuate con cautela, possono allontanarci fortemente da interpretazioni plausibili delle evidenze del record archeologico.

Per quanto concerne il rapporto tra la distribuzione di fauna e litica, questo rappresenta certamente un problema differente. I caratteri socio-culturali dei gruppi che si sono avvicendati, e le differenti storie che i differenti insediamenti hanno vissuto, dovrebbero riflettersi ovviamente nelle configurazioni spaziali. Gruppi di cacciatori-raccoglitori tardo mesolitici (US 5), che processano delle parti di animali, si differenziano certamente da gruppi di agricoltori-allevatori neolitici (US 3), che probabilmente spazzano accuratamente una capanna. Il punto è valutare come discriminare tali configurazioni.

Osservando la distribuzione di industria litica nel livello US 3 Neolitico, è possibile isolare picchi consistenti (Fig. 8). Questo comportamento potrebbe essere indicatore di configurazioni residuali di aree di scheggiatura, in relazione alle alte frequenze di nuclei e *waste*. È stato quindi effettuato un controllo per valutare se le associazioni dei gruppi di strumenti fossero qualitativamente e funzionalmente coerenti: sono state usate a tale scopo diverse tecniche di analisi multivariata. Di queste, la più interessante sembra essere l'analisi fattoriale dell'Indice di Densità Locale (JOHNSON 1984). I calcoli sono stati effettuati utilizzando come sorgente i valori di correlazione dell'indice di densità locale degli strumenti. L'analisi fattoriale, operata con il pacchetto SPSS, è stata effettuata estraendo e ruotando tre fattori con il metodo VARIMAX.

Osservando la Fig. 12a, si notano medi valori di associazione, che potrebbero essere messi in relazione ad un generale processo di omogeneizzazione, che, anche sulla base di osservazioni condotte su altri materiali, sembra avere interessato l'unità stratigrafica. Questi, inoltre, non sembrano isolare associazioni funzionali tra gruppi di strumenti. Quindi, in accordo con la presenza di un "corridoio" tra due strutture come evidenziato dalle mappe di frequenza dei materiali in generale e dal grado di correlazione tra ceramica, pietre e intonaco in particolare, si potrebbe propendere per azioni ripetute di pulizia di alcuni spazi, con manufatti gettati e accumulati in una zona che a sua volta è caratterizzata anche da alte frequenze di frammenti faunistici, con particolari caratteri tafonomici (DI LERNIA 1995; DI LERNIA, FIORENTINO c.s.).

Nel caso del Livello Inferiore-US 5, Tardo Mesolitico, lo stesso tipo di analisi isola dei *clusters* (Fig. 12b), che in alcuni casi sembra plausibile riferire ad associazioni funzionali di strumenti, come ad esempio lame a dorso e grattatoi; la possibilità di effettuare analisi tracceologiche sembrerebbe però

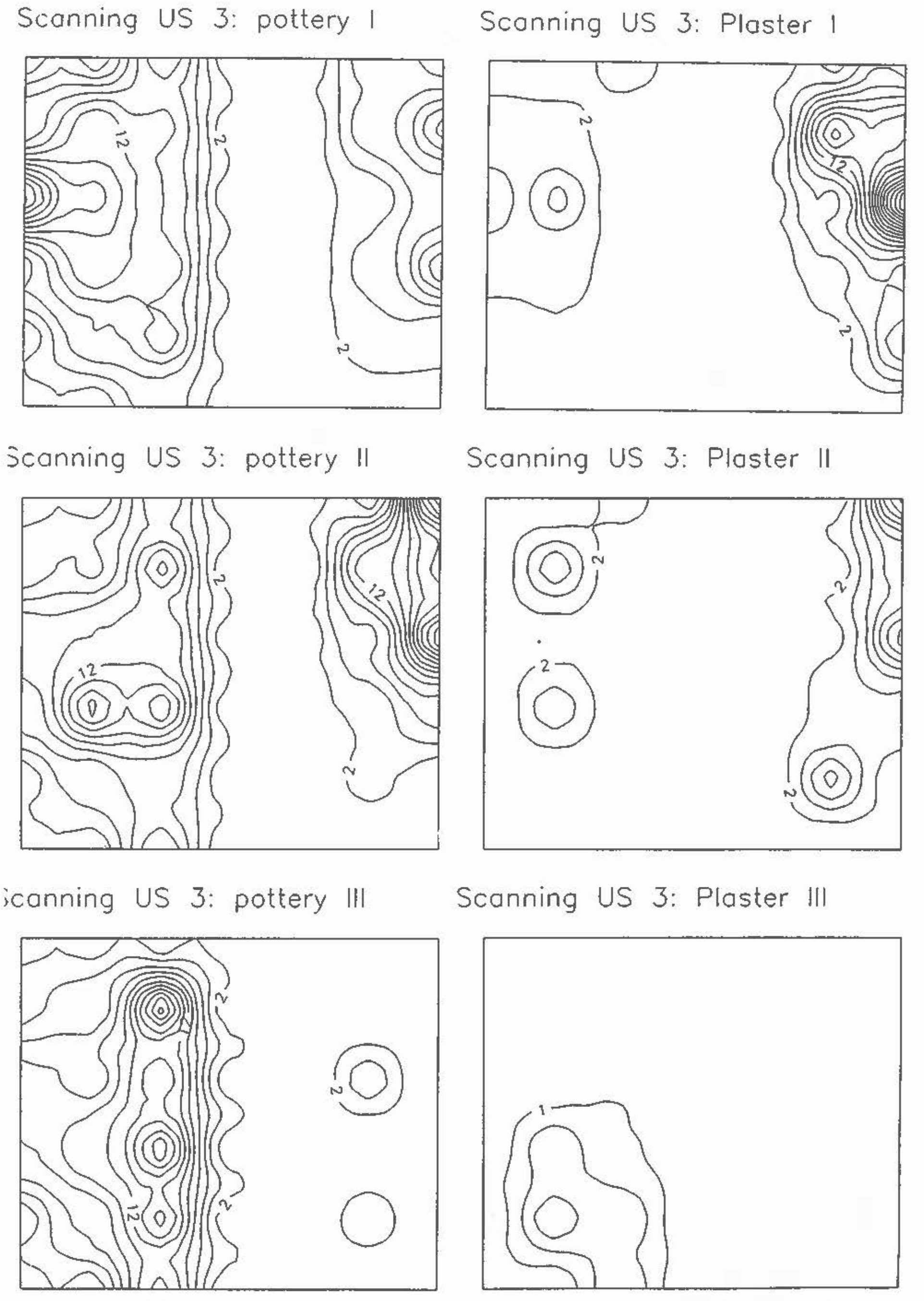
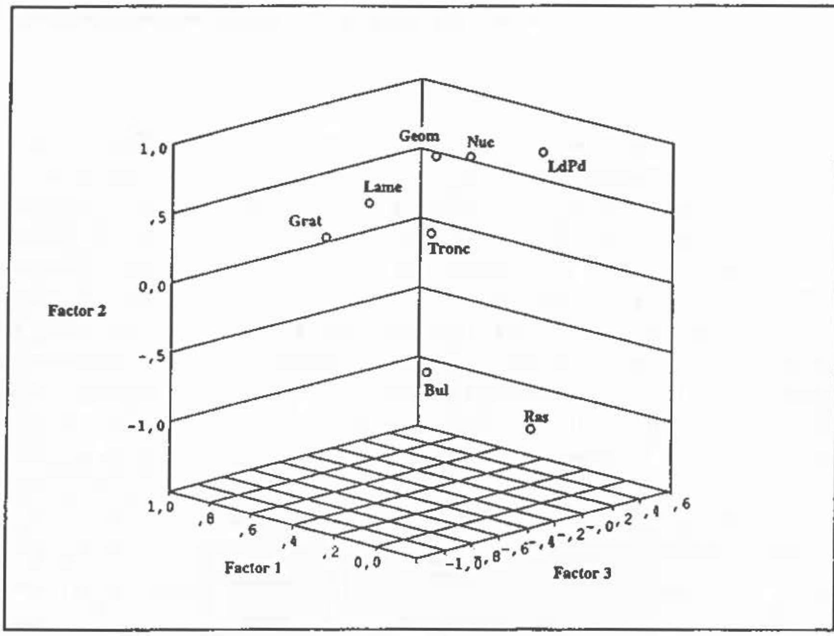
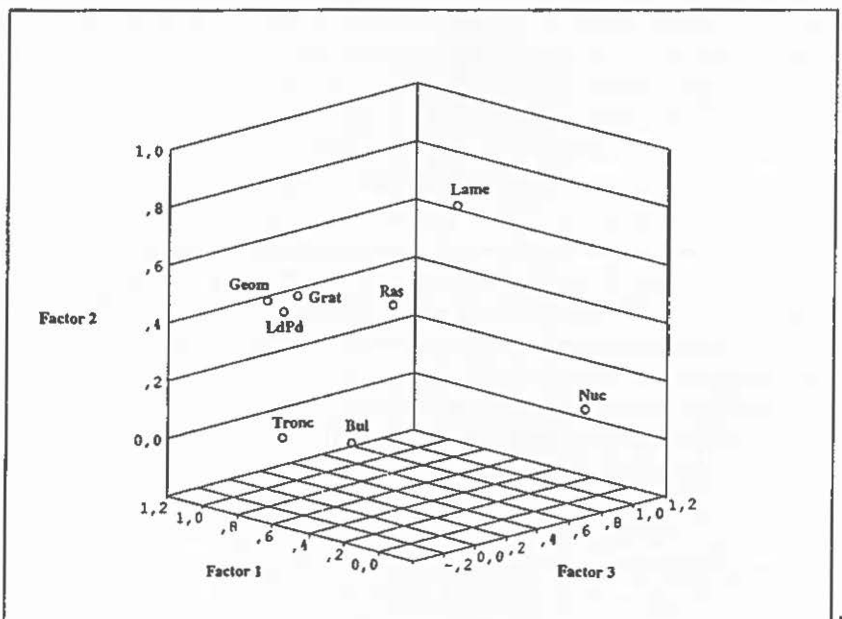


Fig. 11 – Piante di distribuzione di ceramica e intonaco "smussate" (frequenze assolute) del livello Superiore Neolitico – US 3, prendendo in considerazione tagli artificiali di 2 cm.



a

FACTOR PLOT IN ROTATED FACTOR SPACE



b

FACTOR PLOT IN ROTATED FACTOR SPACE

Fig. 12 – Analisi fattoriale degli Indici di Densità Locale dei manufatti litici relativi al livello Superiore Neolitico – US 3 (alto) e al livello Inferiore Mesolitico – US 5 (basso). Gli Indici sono stati utilizzati come sorgente per l'analisi fattoriale, effettuata con il pacchetto SPSS.

allo stato attuale inficiata dall'abrasione che il sedimento ha operato sui manufatti (Lemorini, com. pers.): viene pertanto a mancare un elemento decisivo per confermare ulteriormente o rigettare questa ipotesi.

La ridotta vita dell'occupazione tardo mesolitica, testimoniata dalle datazioni radiometriche e dal ridotto spessore del suolo, dovrebbe essere in relazione ad un rapido seppellimento del sito, come è possibile evincere da un lato per una migliore conservazione delle configurazioni spaziali e dall'altro per le caratteristiche tafonomiche delle faune.

5. Considerazioni conclusive

Riassumendo, le analisi univariate descrittive e multivariate unidimensionali e multidimensionali forniscono una buona lettura della configurazione spaziale delle evidenze archeologiche, evidenziando *pattern* distributivi e caratteri identificativi specifici dei due livelli archeologici, sia su un piano verticale che orizzontale, contribuendo ad una lettura più approfondita del record archeologico.

Credo sia utile segnalare alcuni punti: 1) già le "proprietà semplici dei manufatti" (per es. grandezza, densità, forma, danneggiamenti: *sensu* Schiffer 1987, 267 e ss.) offrono un notevole grado di elaborazione, permettendo un consistente avvicinamento alla comprensione dei processi di formazione. Per esempio, le semplici mappe di densità, "smussate" con gli appositi *software* dedicati, consentono in numerosi casi un buon livello di approssimazione; 2) l'importanza di un attento controllo delle caratteristiche di scavo, ed una coerente esplicitazione dei metodi che ne hanno regolato l'effettuazione: il problema non è la qualità dello scavo (qualsiasi essa sia), ma in primo luogo la qualità delle informazioni che vengono rese disponibili inerenti le specificità dell'intervento. L'applicazione sfrenata di sviluppatissimi *software* a contesti di cui non è resa nota la metodologia di intervento crea forti problemi nel poter concretamente valutare e confrontare le esperienze di ricerche diverse. Il caso di Terragne mi sembra emblematico: laddove è stato possibile effettuare uno scavo (ed un rilievo) di maggiore dettaglio, le potenzialità di elaborazione sono ovviamente aumentate, rivelando però molto di più non tanto sulla comprensione della problematica generale del sito, quanto piuttosto su un piano qualitativo di alcuni specifici problemi; 3) una maggiore cautela nell'uso di *software* commerciali e controllo del dominio teorico dei metodi utilizzati. Come da più parti sottolineato (Bietti 1993; Blankolm 1991) un solo metodo di analisi non è certamente sufficiente, ma più metodi dovrebbero essere provati ai fini di una interpretazione plausibile del record archeologico.

Vorrei concludere mettendo in rilievo la capacità che questi depositi all'aperto hanno comunque di mantenere discretamente le configurazioni orizzontali, e in seconda battuta quelle verticali, sebbene siano poco spessi e assai disturbati. Ritengo che approcci di questo tipo siano quantomeno auspicabili, soprattutto in considerazione dell'elevato tasso di antropizzazione riscontrabile in determinate regioni italiane, in cui le Pompei sono le eccezioni, e i *distorted stuffs* la norma.

SAVINO DI LERNIA
Dipartimento Scienze Storiche
Archeologiche e Antropologiche dell'Antichità
Università degli Studi di Roma "La Sapienza"

BIBLIOGRAFIA

ARCOSPACE ver. 2.0 1987, *A Package for the Spatial Analysis of Archaeological Data*, by T.D. PRICE and H.P. BLANKHOLM.

BASP 5.0 1994, *Bonn Archaeological Statistical Package*, by I. SCOLLAR, Unkelbach Valley Software Works.

BLANKHOLM H.P. 1991, *Intrasite Spatial Analysis in Theory and Practice*, Aarhus, Aarhus University Press.

BIETTI A. 1993, *Nuove prospettive nelle analisi di correlazioni spaziali in preistoria* «Archeologia e Calcolatori», 3, 39-60.

BINFORD L.R. 1981, *Behavioral Archaeology and the "Pompeii Premise"*, «Journal of Anthropological Research», 36, 3, 195-208.

CREMASCHI M. 1990, *Geoarcheologia: metodi e tecniche delle scienze della terra nella ricerca archeologica*, in T. MANNONI, A. MOLINARI (edd.), *Scienze in Archeologia*, Quaderni del Dipartimento di Archeologia e Storia delle Arti, Università di Siena, 20-21, 395-422.

CREMASCHI M., DI LERNIA S. 1995, *Caratteri sedimentologici e micromorfologici della successione stratigrafica*, in GORGOGLIONE *et al.* 1995, 37-44.

DE GUIO A. 1988, *Unità archeostratigrafiche come unità operazionali: verso le archeologie possibili degli anni '90*, «Quaderni di Archeologia Stratigrafica dell'Italia Settentrionale», 1, 9-22.

GORGOGLIONE M.A., DI LERNIA S., FIORENTINO G. 1991, *L'insediamento preistorico di Terragne, Manduria (Taranto). Indagine territoriale e metodologia d'intervento*, «Taras», 11, 1, 57-83.

GORGOGLIONE M.A., DI LERNIA S., FIORENTINO G. (edd.) 1995, *L'insediamento preistorico di Terragne (Manduria, Taranto). Nuovi dati sul processo di neolitizzazione nel sud est italiano*, Manduria, CRSEC.

HARRIS E.C. 1979, *Principles of Archaeological Stratigraphy*, London, Academic Press.

JOHNSON I. 1984, *Cell frequency recording and analysis of artifact distributions*, in J.H. HIETALA (ed.), *Intrasite Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press, 75-96.

LEONARDI G. 1992 (ed.), *Processi formativi della stratificazione archeologica*, «Saltuarie dal Laboratorio del Piovego», 3, 13-47.

LEONARDI G., PRACCHIA S., VIDALE M. 1989, *L'indicatore ceramico nei percorsi archeologici*, in *Atti della XXVII Riunione Scientifica I.I.P.P. (Ferrara 17-20 novembre 1987)*, 85-104.

DI LERNIA S. 1995, *Analisi spaziale delle industrie litiche di Terragne*, in GORGOGLIONE *et al.* 1995, 73-85.

DI LERNIA S., FIORENTINO G. 1995, *Analisi del deposito archeologico: il caso della ceramica*, in GORGOGLIONE *et al.* 1995, 45-72.

DI LERNIA S., FIORENTINO G. c.s., *Decoding an open-air archaeological deposit: the prehistoric settlement of Terragne (Manduria - Taranto, South Eastern Italy). Formation* *processes and spatial analysis*, «Origini», 19 (in corso di stampa).

O'Brien M.J., Lewarch D.E. (edd.) 1981, *Plowzone Archaeology: Contributions to Theory and Technique*, Vanderbilt University Publications in Archaeology, 27.

Orton C.R., Tyers P.A. 1990, *Statistical analysis of ceramic assemblages,* «Archeologia e Calcolatori», 1, 81-110.

Sanesi G. 1977, *Guida alla descrizione del suolo,* Firenze, C.N.R.

Schiffer M.B. 1977, *Toward a unified science of the cultural past*, in S. South (ed.), *Research Strategies in Historical Archaeology*, New York, Academic Press, 13-50.

SPSS 5.0 1992, *Statistical Package for Social Sciences.*

Villa P. 1982, *Conjoinable pieces and site formation processes*, «American Antiquity», 47, 276-290.

Whallon R. 1984, *Uncostrained clustering for the analysis of spatial distributions in archaeology*, in J.H. Hietala (ed.), *Intrasite Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press, 245-277.

ABSTRACT

The analysis of site formation processes seems quite disregarded in the archaeological literature of Southern Italy. In this paper, we discuss the case-study of Terragne (Taranto, South-Eastern Italy), an open-air site characterised by two occupation layers (US 5- Late Mesolithic; US 3- Early and Middle Neolithic). Statistical analyses of different archaeological indicators were performed, in order to value the formation features of the deposit and to identify possible spatial configurations. Particular emphasis has been drawn to the identification of diagnostic tools, able to isolate specific formative phenomena (sin-depositional and post-depositional).

Archeologia e Calcolatori
7, 1996, 357-363

LA SABINA TIBERINA. ANALISI ESPLORATIVA PER LA ZONA ARCHEOLOGICA DELL'*AGER ERETANUS*

Premessa

Come è noto la Sabina Tiberina è un territorio caratterizzato da notevoli tracce di antichi insediamenti antropici che vanno dall'epoca arcaica fino al tardo medioevo.

Col presente lavoro si intende portare un contributo alla conoscenza delle consuetudini di quelle antiche popolazioni mediante l'uso di metodologie proprie della statistica spaziale e più precisamente delle tecniche di *point pattern analysis*. A questo proposito, pur considerando come area di studio l'intera Sabina Tiberina, l'interesse è stato prevalentemente incentrato sull'*Ager Eretanus* poiché al riguardo la letteratura, oltre ad un'ampia documentazione (Quilici Gigli, Santoro 1994), rende immediatamente disponibile una ricca e ben organizzata raccolta di dati archeologici e topografici (Ogilvie 1965).

È qui appena il caso di ricordare che uno studio di questa natura necessita di un importante supporto informatico quale un Sistema Informativo Territoriale (G.I.S.) opportunamente organizzato. Il lavoro è comunque strutturato in tre parti. Nella prima sezione vengono presentate sinteticamente le modalità di organizzazione di un Sistema Informativo Territoriale orientato prevalentemente ad applicazioni per la ricerca archeologica nella zona indicata in premessa. Nella seconda sezione vengono proposti e commentati i risultati di un'applicazione di *point pattern analysis* ai siti di epoca sabina cui si è fatto riferimento sopra (Ogilvie 1965). Nel terzo paragrafo infine, oltre ad uno studio sulla robustezza del modello esplicativo identificato, viene proposto un metodo di rappresentazione ad isolinee di densità di siti archeologici (Getis, Franklin 1986).

1. Il Sistema Informativo Territoriale (G.I.S.)

Il G.I.S. è stato sviluppato inserendo sul data base cartografico innanzitutto la cartografia di base della zona. Più precisamente sette tavolette in scala 1:25.000 dell'I.G.M. sono state digitalizzate in formato raster per riferimento topografico su video, mentre le infrastrutture e la idrografia principale presenti sulle zone di interesse sono state digitalizzate per costituire riferimento topografico nella stampa di mappe tematiche. Successivamente è stata inserita l'orografia numerica in scala 1:25.000 dell'I.G.M. e la copertura aerea del comune di Magliano Sabina sempre alla stessa scala. Inoltre è stata georeferenziata ed inserita nel G.I.S. un'immagine da piattaforma orbitante LANDSAT di tutta l'area della "Sabina Tiberina" nonché la geologia regionale in scala 1:100.000.

Sono stati infine inseriti nel Sistema Informativo dati geometrici e descrittivi relativi alla tavoletta I.G.M. di Passo Corese (*Ager Eretanus*) desunti da letteratura (Ogilvie 1965) ovvero, per quanto riguarda *Eretum*, sono stati studiati e classificati più di 100 siti tra sabini-arcaici e romani. Una prima aggregazione dei dati, è stata ottenuta costruendo delle tabelle di codifica suddivise in classi di appartenenza per epoca e per tipologia di sito come indicato nelle tabelle seguenti:

	A	*B*	*C*	*D*
0	**Protostorico**	**Arcaico - Sabino**	**Romano**	**Medioevale**
1			**Repubblicano**	
2			**Imperiale**	

Tab. 1 – Codifica per epoche storiche.

	A	*B*	*C*	*D*	*E*
	GENERICO	*RELIGIOSO*	*ABITATIVO*	*FUNERARIO*	*PRODUTTIVO*
1	**Non definito**	**Tempio**	**Insediamento**	**Tomba**	**Cisterna**
2		**Deposito votivo**	**Casa**	**Tomba a camera**	**Acquedotto**
3		**Chiesa**	**Capanna**	**Cimitero**	**Forno**
4			**Villa**	**Mausoleo**	**Pozzo**
5			**Torre**	**Sarcofago**	**Macina**
6				**Ossa**	**Pressa**

Tab. 2 – Classificazione e codifica di siti archeologici.

Per Magliano Sabina dati storici e geologici di superficie sono stati acquisiti mediante ricognizione e georeferenziazione sul terreno. Il G.I.S. così organizzato ha consentito tutta una serie di elaborazioni riguardanti sub-aree archeologiche della Sabina Tiberina. In particolare, sono state prodotte mappe tematiche (ricostruzione D.T.M. del terreno, mappe delle pendenze, studio delle esposizioni etc.), anche a tematismi sovrapposti, nonché analisi statistiche preliminari, in vista dello studio modellistico della zona archeologica oggetto della presente indagine.

2. L'analisi del II ordine di una mappa di punti

Dopo una classificazione per epoca e per tipologia di sito, la prima operazione effettuata è stata quella di trasformare la distribuzione spaziale dei siti, selezionati per epoca storica, in *pattern* di punti distribuiti in uno spazio omogeneo (Espa *et al.* 1995). Occorre precisare che in questo lavoro si è fatto riferimento ai siti di epoca sabina poiché la domanda storica relativa ai sabini di epoca storica, in questo periodo è interessata allo studio di quelle strutture che spesso sono alla base dei fenomeni sociali ed economici dei vari aggregati umani. Si è quindi proceduto all'analisi preliminare della mappa di punti così ottenuta attraverso la funzione $\hat{K}(d)$ di Ripley (RIPLEY 1977) e cioè:

$$\hat{K}(d) = A \sum_{i=1}^{N} \sum_{j=1}^{N} K_{ij}(d) / N^2 \qquad [1]$$

o più precisamente di una sua trasformata (la funzione $\hat{L}(d)$ dovuta a Besag (BESAG 1977) ovvero

$$\hat{L}(d) = [\hat{K}(d) / \pi]^{1/2} \qquad [2]$$

Analizzando la Fig. 1 si può osservare che:

1. il grafico $\hat{L}(d)$ mostra una marcata tendenza alla clusterizzazione nell'intervallo 487 m < d < 1625 con un livello di significatività pari al 5%
2. $\hat{L}(d) = 0$ nell'intervallo $0 \leq d \leq 130$ m significa che non si hanno siti in un intorno di 130 m di raggio con centro in uno qualunque dei siti considerati (distanza del vicino più prossimo)
3. d = 487 m è la soglia al di là della quale la clusterizzazione è statisticamente significativa.
4. d = 1625 m è la soglia di massima clusterizzazione

Tutte queste informazioni sono estremamente importanti perché consentono, almeno sul piano teorico, il confronto tra varie configurazioni spaziali di siti di interesse storico. Esse costituiscono altresì la base per la costruzione e la verifica di modelli e di ipotesi socio-culturali.

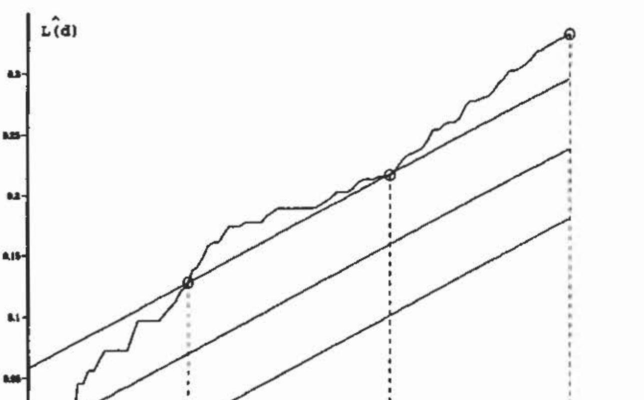


Fig. 1 – Siti di epoca arcaico-sabina. Diagramma della funzione $\hat{L}(d)$ calcolata per N = 25 siti.

3. Verifica della robustezza del modello identificato

I dati trattati nel precedente paragrafo, provenendo da letteratura e quindi da rilevazioni non recenti, possono prestarsi ad alcune obiezioni. Per esempio: il rilevatore ha adottato criteri e metodi di misura diversi da quelli accettati dagli strumenti di elaborazione attualmente disponibili oppure, riguardo alla precisione, la moderna tecnologia avrebbe certamente consentito rilevazioni più accurate.

Per verificare l'influenza sul modello identificato di eventuali errori di localizzazione, si è proceduto ad una serie di simulazioni in cui la posizione originaria dei siti viene modificata mediante l'introduzione di errori casuali.

Il processo di generazione dei dati (D.G.P. - *Data Generating Process*) ha seguito i seguenti due criteri:

1) ogni sito viene spostato casualmente su una circonferenza di raggio variabile centrata sul sito originario (ipotesi forte);

2) ogni sito è stato spostato a caso in un cerchio di raggio variabile centrato sul sito originario (ipotesi debole).

Ogni punto della mappa originaria è stato spostato di una certa quantità casuale lungo le coordinate [x,y]. In altri termini il procedimento seguito è il seguente: per ogni punto sono stati calcolati, in coordinate polari (modulo e anomalia) gli spostamenti casuali della posizione iniziale. Inoltre, per effettuare questi spostamenti si è proceduto fissando intervalli di variazione del raggio (modulo) sempre più ampi (di 10 m in 10 m), con la possibilità di estrarre valori compresi tra zero e l'estremo superiore dell'intervallo, mentre l'angolo (anomalia) viene fatto variare tra 0° e 359° di grado in grado. L'intervallo max simulato è un raggio di 100 metri.

La scelta di intervalli successivi è stata fatta per determinare eventualmente una soglia oltre la quale il modello esplicativo identificato per la mappa esaminata assumesse caratteri significativamente diversi. Un'ulteriore elaborazione con ipotesi di casualità più restrittive (ma con errore di posizionamento maggiore), è stata eseguita con un procedimento analogo al precedente.

L'intervallo di variazione del raggio sempre di 10 m fino ad un max di 100 m, coincide con l'estremo superiore del raggio, mentre l'angolo viene generato casualmente di grado in grado. In questo caso si avrà una distanza dal punto originario comune per tutti i punti della mappa ognuno con un valore dell'angolo diverso e casuale. Entrambi i processi sono stati ripetuti 100 volte e per ogni replicazione è stata disegnata la funzione $\hat{L}(d)$. Dall'analisi dell'insieme dei grafici così ottenuti, si può osservare che solo nei casi limite si ha una modesta variazione dei parametri del modello, mentre resta sostanzialmente identico l'andamento dei grafici.

Ciò significa che, anche ammettendo la presenza di consistenti errori di misura, le conclusioni inferenziali circa la natura del modello stocastico da adattare ai dati non mutano. I risultati mostrano come le conclusioni inferen-

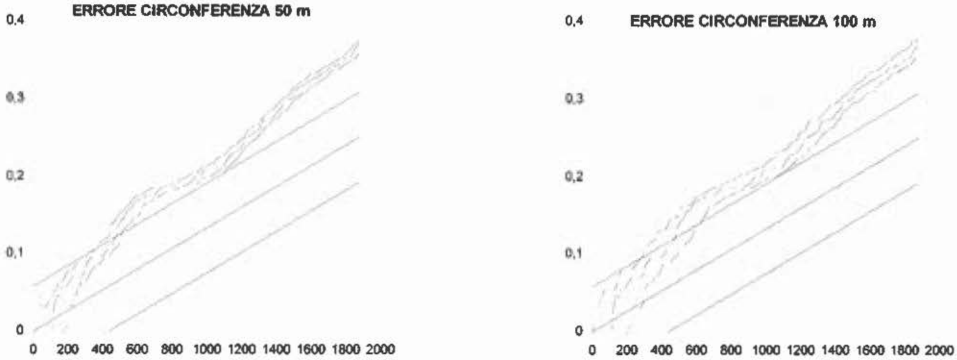


Fig. 2a – Andamento della $\hat{L}(d)$ per 25 siti Sabini con errori casuali di localizzazione su un cerchio di r = 50 m con variazione angolare 0° ± 359°, passo 1° (ipotesi forte).
Fig. 2b – Andamento della $\hat{L}(d)$ per 25 siti Sabini con errori casuali di localizzazione su un cerchio di r = 100 m con variazione angolare 0° ± 359°, passo 1° (ipotesi forte).

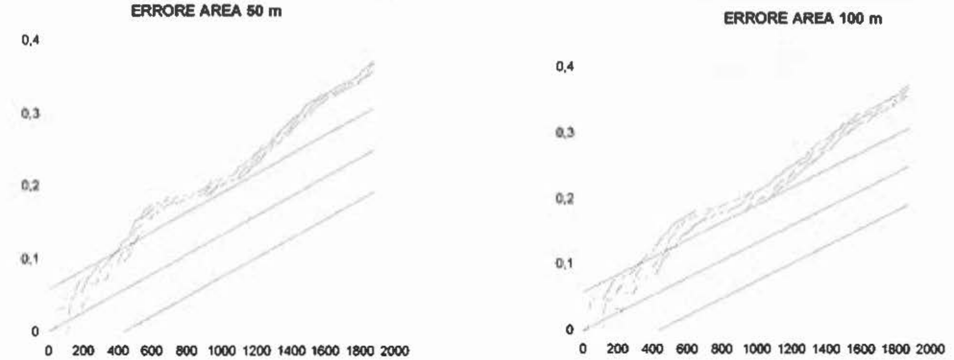


Fig. 3a – Andamento della $\hat{L}(d)$ per 25 siti Sabini con errori casuali di localizzazione su un cerchio di r_{max} = 50 m con variazione angolare 0° ± 359°, passo 1° (ipotesi debole).
Fig. 3b – Andamento della $\hat{L}(d)$ per 25 siti Sabini con errori casuali di localizzazione su un cerchio di r_{max} = 100 m con variazione angolare 0° ± 359°, passo 1° (ipotesi debole).

ziali formulate non subiscono variazioni di rilievo (cfr. Figg. 2 e 3).

3.1 *Isolinee di densità*

Infine, utilizzando ancora una trasformata della funzione $\hat{K}(d)$ di Ripley ovvero $\hat{L}_i(d)$ (Getis, Franklin 1986)

$$\hat{L}_i(d)=\left[A\sum_{j=1}^{n}K_{ij}/\pi(n-1)\right]^{1/2} \qquad [3]$$

si sono ottenute delle rappresentazioni sia ad isolinee, sia in scala cromatica, che evidenziano zone di intensità di sito variabile (Tav. XXa-b). Informazioni di questa natura rivestono un'importanza fondamentale in quanto consentono di quantificare fenomeni di clusterizzazione a differenti livelli di scala.

4. Considerazioni conclusive

Fin qui si sono mostrati i risultati relativi all'utilizzo di una statistica funzionale, cioè la funzione $\hat{L}(d)$, la quale rappresenta in un certo senso "la media" delle relazioni che intercorrono fra le distanze tra i siti archeologici osservati. Per quantificare i fenomeni di attrazione e repulsione che interessano le configurazioni spaziali in esame, si può procedere costruendo la funzione $\hat{L}_i(d)$ per ogni singolo sito archeologico. In tal modo si può stabilire a quale livello di distanza un sito appartenga o meno ad un *cluster*.

Lo strumento statistico adottato è, fra l'altro, di grande utilità qualora il *pattern* totale non si discosti dalla casualità completa per studiare fenomeni locali di inibizione e clusterizzazione che possono comunque essere presenti. In particolare, nel caso in esame **non** è stato affrontato il problema di condurre un'analisi del secondo ordine via $\hat{L}_i(d)$ per alcuni siti particolarmente significativi e rappresentativi. L'attenzione è stata però incentrata su una rappresentazione grafica di immediato riscontro pratico.

Prendendo in considerazione la sola soglia d = 487 m (distanza alla quale la clusterizzazione diviene statisticamente significativa) sono stati calcolati i valori $\hat{L}_i(d = 487)$, i = 1, 2, ..., 25. Una rappresentazione per isolinee $\hat{L}_i(d = 487)$, i = 1, 2, ..., 25 è riportata in Tav. 0,1a; una lettura più semplice della Tav. XX, a è consentita dall'esame di Tav. XX, b che mostra una rappresentazione di $\hat{L}_i(d = 487)$ in scala cromatica. Le porzioni di figura più chiare corrispondono alle zone interessate da più intensi fenomeni di clusterizzazione.

La Tav. XX a riporta i siti oggetto di studio e le isolinee per le quali $\hat{L}_i(d)$ supera il valore atteso nell'ipotesi di casualità completa $\hat{L}_i(d =487)$ ed evidenzia in modo più preciso le zone di maggiore intensità di sito archeologico.

In questo lavoro ci si è limitati a presentare un esempio che costituisce una sorta di "fotografia istantanea scattata" ad una soglia pari a 487 m. È interessante, ed è una delle linee di studio del gruppo di lavoro, valutare quali cambiamenti occorrono quando si fa variare la soglia d (magari scegliendo un limitato numero di livelli d particolarmente significativi). In particolare sarebbe interessante valutare se i *clusters* individuati ad un dato livello di risoluzione rimangono tali variando la scala oppure se vanno a costituire dei *clusters* più ampi.

Giuseppe Espa
Università degli Studi di Trento
Istituto di Statistica e Ricerca Operativa

Salvatore Espa
Roberto Gabrielli
Ugo Ricci
Istituto per le Tecnologie Applicate ai Beni Culturali
C.N.R. - Area della Ricerca di Roma, Montelibretti

BIBLIOGRAFIA

BESAG J. 1977, *Contribution to the discussion of Dr. Ripley's paper*, «Journal of the Royal Statistical Society», B,193-195.

ESPA G., ESPA S., GABRIELLI R., PIFFERI A., RICCI U. 1995, *Impiego integrato di metodi statistico-spaziali e tecniche informatiche nell'analisi di aree di interesse storico-archeologico*, «Archeologia e Calcolatori», 6, 35-60.

GETIS A., FRANKLIN J. 1986, *Second-Order Neighborhood Analysis of Mapped Point Patterns*, «Ecology», 65 (3), 473-477.

OGILVIE R.M. 1965, *Eretum*, «Papers of the British School at Rome», 32, 70-112.

QUILICI GIGLI S., SANTORO P. 1994, *Eretum: ricerca topografica sull'abitato in epoca arcaica*, «Archeologia Laziale», 12, 641-663.

RIPLEY B.D. 1977, *Modelling Spatial Pattern (with Discussion)*, «Journal of the Royal Statistical Society», B, 39, 172-212.

ABSTRACT

The aim of this work reported here is the contribution to the knowledge of the ancient population located in the *Ager Eretanus* through the statistical methodology known as "point pattern analysis". We first present a brief discussion on the management of an application-oriented GIS in the field of archaeological research. Then we show the results of a "second order analysis" on the data at our disposal. Finally we approach two interesting methodological problems. The first concerns the evaluation of the influence exerted on the model by possible errors in the location of the points. The second regards the proposal of a graphical representation which proves very useful in showing the variable intensity of the distribution of archaeological site.

Archeologia e Calcolatori
7, 1996, 365-372

FORCASTING STATISTICAL MODELS OF ARCHAEOLOGICAL SITE LOCATION

1. Introduction

Forecasting statistical models are becoming increasingly important in archaeological research (Kohler, Parker 1986). One of the reasons of this popularity is that archaeological sites tend to present themselves in particular environments so that forecasting models can help in identifying areas where the probability is higher based on previously collected statistical information. In the present paper we will consider a class of statistical models designed to produce maps of the probability for archaeological site location (henceforth ASL) which incorporate both deductive and inductive considerations.

Forecasting models for the probability of ASL can be classified into two classes by distinguishing between models on a continuous space and models on a discrete space. The output produced in the two cases is displayed in Fig. 1. In the first instance the models produce a probability surface for ASL (Fig. 1a). In the second instance the space is discretized by superimposing a grid of contiguous quadrats and the output is an array of probability values (Fig. 1b).

In the present paper we will refer about the class of statistical models on a discrete space. The approach based on a continuous space have been exploited elsewhere (for instance by Benedetti, Espa 1995), but is not considered here.

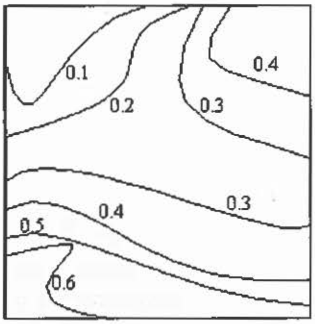
The models described can be of help in practical circumstances exploiting the potentialities of data derived from satellite and aerial photographs and can be easily implemented within the context of a GIS (Arroyo-Bishop 1995).

The paper is organised as follows. Section 2 is devoted to discuss one of the currently most popular approach to forecasting modelling in archaeology that is the "integrated strategy" of Kwamme (1983) and Warren (1990). Starting from the weakness of this approach in Sections 3 and 4 we will discuss possible extensions to correct the procedures using contextual information, we propose a new methodology and we discuss the relative advantages against the current practice. Finally Section 5 is devoted to some conclusions and to outline the research agenda in the field.

2. The integrated approach: logistic models

One of the most popular approach to forecasting modelling in archaeology is the "integrated strategy" developed by Kwamme (1983) and employed by several researchers like Warren (1990). Kwamme strategy exploits the potentiality offered by a GIS to create and process large data sets through the logistic

(a)



(b)

0.1	0.1	0.2	0.3	0.4
0.2	0.2	0.2	0.3	0.4
0.3	0.3	0.3	0.3	0.4
0.4	0.4	0.4	0.5	0.5
0.4	0.4	0.5	0.6	0.6

Fig. 1 – Probability maps of ASL expressed on a continuous space (a), and on a discrete space (b).

regression, a flexible statistical tools which allows to forecast binary variables.

Suppose we are analysing a study area where a number of ASL have been identified by means of a field survey. The study area is discretized into M contiguous quadrat cells. Suppose that in N of such cells ($N<M$) the site survey was able to assess the presence or absence of an archaeological site. On this basis we define a variable Y_i such that:

$$Y_i = \begin{cases} 1 \text{ if cell } i \text{ contains an archaeological site} \\ 0 \text{ otherwise} \end{cases} \qquad i = 1, \ldots, N$$

An example is reported in Fig. 2.

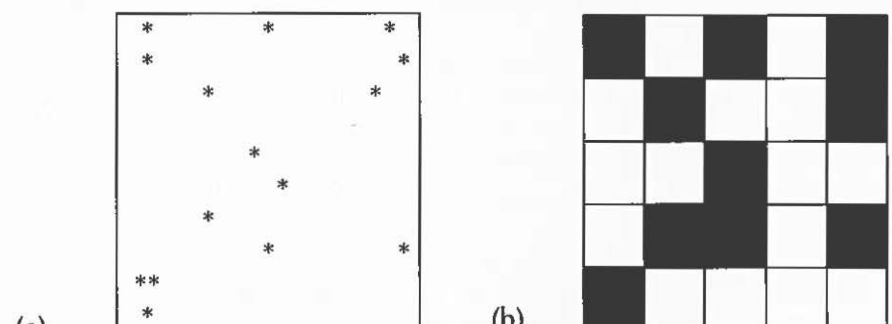


Fig. 2 – Continuous distribution of ASL in a study area (a) and its discretized version on a 5-by-5 quadrat grid (b).

Suppose further that for all M cells it is available a set of auxiliary information about k independent variables $\mathbf{X}_i \equiv (X_{1i}, X_{2i}, \ldots, X_{ki})$ $i = 1, \ldots, M$. The independent variables could be related to the nature of the soil (slope, erosion, exposure, etc.), to the topography (DTM), the hydrology (rivers, channels, lakes, etc.), the topology (nearness to streets, to communication routes, etc.), or to other variables derived from existing cartography, aerial or satellite imagery, survey samples and other sources (CARLA' *et. al.*, 1995).

In such a situation we can define on the basis of the *N* observations a model in which the probability of finding an ASL, say θ_i = Prob{Y_i = 1}, is a function of the vector of predictors X and of a vector γ of parameters, that is:

$$[1] \quad \theta = f(X, \gamma)$$

In particular the linear logistic regression (COX, SNELL 1989) specifies the relationship [1] as

$$[2] \quad \Pr ob\{Y_i = 1\} = \theta_i = \frac{\exp[x'\gamma]}{1 + \exp[x'\gamma]}$$

and similarly

$$[3] \quad \Pr ob\{Y_i = 0\} = \theta_i = \frac{1}{1 + \exp[x'\gamma]}$$

or, in general,

$$[4] \quad \Pr ob\{Y_i = \gamma_i\} = \frac{\gamma_i \exp[x'\gamma]}{1 + \exp[x'\gamma]}$$

Model [2] and [3] can be estimated via a maximum likelihood procedure by choosing a subset *n* (*n*<*N*) of the *N* selected cells (called *training sites*), and cross-validated by contrasting the results with the (*N*-*n*) observed sites. Finally the model can be employed to forecast the probability of ASL on the (*M*-*N*) unobserved cells. The output is a probability map of the kind displayed in Fig. 1b.

However the use of the logistic regression in such a context is statistically incorrect since the model assumes a spatial independence in the Y_i's (COX, SNELL 1989), an hypothesis which is patently violated in all geographical studies where (as stated in Tobler's – first law of geography –, TOBLER 1970) «everything is related to everything else, but near things are more related than distant things». As a role archaeological sites tend to cluster in space and this fact results into higher probability of sites in the neighbourhood of existing sites. As a consequence closeness to other archaeological sites modifies (generally increases) our expectation of other sites. On the other hand it is obvious that neglecting such a contextual information is statistically inefficient and is doomed to produce unreliable estimates and forecasts (see CRESSIE 1992; ARBIA 1993).

3. IMPROVING THE INTEGRATED APPROACH: THE AUTO-LOGISTIC MODEL

The formal way of incorporating the notion of contextual information into binary response variable models has been introduced in the statistical literature by BESAG (1974).

Define for each cell *i* constituting a study area ($i = 1, 2, ...M$), the set N(i) which represents the set of neighbouring cells. For instance one could assume a neighbouring scheme on horizontal and vertical direction (or "rook's neighbourhood" from the chess rook's move) as the one depicted in Fig. 3:

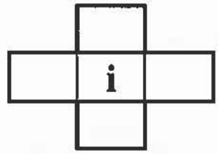


Fig. 3 – Neighbouring cells to the *i*-th cell with a horizontal/vertical criterion.

Now we can model the probability of finding a cell containing an ASL conditional upon the observations of the remaining cells as a function of only the neighbouring cells N(i), and of a set of parameters β

$\theta_i = \text{Prob}\{Y_i = 1 \mid Y_j = 1\ j \neq i\} = f\,(Y_j\,,\ j \in N(i),\ b)$

Besag's "auto-logistic" model is expressed as:

$$[5] \quad \theta_i = P\left(Y_i = y_i \mid y_j, j \in N(i)\right) = \exp\left[y_i\left(\beta_{0i} + \sum_{j \in N(i)} \beta_{ij} y_j\right)\right] \Big/ \left[1 + \exp\left(\beta_{0i} + \sum_{j \in N(i)} \beta_{ij} y_j\right)\right]$$

where $y_i = [0,1]$, N(i) represents the sets of cells close to the i-th pixel of the image and the β_{0i} and the b_{ij}'s are parameters to be estimated. In particular β_{0i} is a location parameter which can incorporates the influence of the independent variables selected from theory, and the remaining β's control for spatial interaction, that is the intensity of the influence of local context on archaeological site expectations (Arbia 1993).

Model [5] incorporates the notion of Markov dependency in spatial processes and is also known in statistical physics as *Ising's law* (see Ising 1925 or Arbia 1993 for a review). The spatial interaction parameters β_{ij}'s can often be decomposed as

$$[6] \quad \beta_{ij} = \beta\, w_{ij}$$

with w_{ij} such that:

$$[7] \quad w_{ij} = \begin{cases} 1 & \text{if } j \in N(i) \\ 0 & \text{otherwise} \end{cases}$$

The parameter β embodies the degree of contextual information accounted for in the model. Consider the following example. Suppose that in Formula [5] we have $\beta_0 = 1$ each i, and $\beta_{ij} = \beta\, w_{ij}$, and consider the following distribution of presence/absence of ASL reported in Fig. 4.

Fig. 4 – Distribution of ASL Presence/Absence in the neighbourhood of site *i*.

In this case Expression [5] becomes:

[8] $$\theta_i = \Pr ob(Y_i = 1) = \frac{\exp(1+3\beta)}{1+\exp(1+3\beta)}$$

If there is no contextual information captured by the model we have that $\beta = 0$ and Formula [8] becomes:

$$\theta_i = \Pr ob\ (Y_i = 1) = \frac{\exp(1)}{1+\exp(1)} = 0.73$$

In contrast, if there is a positive expectation of finding ASL close to one another, we have that $\beta > 0$. For instance, if $\beta = 1/3$ we have:

$$\theta_i = \Pr ob(Y_i = 1) = \frac{\exp(2)}{1+\exp(2)} = 0.88$$

and the probability is higher than in the case of no contextual information. If we want to account for a higher degree of contextual information, for instance $\beta = 1$, we have, instead:

$$\theta_i = \Pr ob(Y_i = 1) = \frac{\exp(4)}{1+\exp(4)} = 0.98$$

which further increases the probability.

The model described above can be estimated via a pseudo-maximum likelihood procedure or, alternatively through the so-called *coding technique* (Besag 1974) by choosing a subset n ($n<N$) of the N *training sites*, cross-validated by contrasting the results with the (N-n) observed sites, and employed to forecast the probability of ASL on the (M-N) unobserved cells.

4. A joint logistic/autologistic approach

The approach presented in the previous section suggests that a gain in the comprehension of ASL can be obtained by considering information collected in the context where each observation is embedded. However, contrasting the logistic model presented in Section 2 which is statistically incorrect, we now have a model which is statistically correct, but does not take into account in a sufficient way of any *a priori* knowledge on ASL derived from auxiliary variables. It seems therefore reasonable to join the two approaches and to introduce a third model that seeks to incorporate both contextual and auxiliary information.

The *log/autolog* model (as we will henceforth refer to this third model) states that the probability of finding an ASL conditional on existing data and on contextual information can be represented as:

$$[9] \qquad \theta_i = \text{Prob}\,\{Y_i = 1 \mid y_j\, j \in N(i);\, X\} = \frac{\exp(\beta_0 + \sum_j \beta_{ij} y_j + X'\gamma)}{\left[1 + \exp(\beta_0 + \sum_j \beta_{ij} y_j + X'\gamma)\right]}$$

where β_0 is a constant, the β_{ij}'s are the spatial interaction parameters with an analogous meaning to those introduced in the autologistic model (see Formula [8]), the vector **X** is a vector of independent variables like those employed in model [2] and [3] related to auxiliary information, and γ the associated parametrization.

Model [9] is conceived so as to remove the problems connected with the non independence of the Y's, and simultaneously, to augment the autologistic model with information coming from auxiliary variable related to the nature of the soil. As such it appears ideal under many respects in that it manages to produce more reliable probability maps of ASL than the logistic model used following the "integrated approach", and to incorporate archaeologists' prior knowledge on the study area. However, as we did in the previous sections while presenting the various models, it is fair to underline the drawbacks of this alternative approach.

The main one seems to be that model [9] is formally more difficult to treat than the other two models, and that some of the statistical problems connected with its estimation are still unsolved. The statistical estimation problem can be described as follows. The usual situation we have to deal with is the case in which we have the study area discretized into *M* cells, but observations are available for only *N* of such cells (see Fig. 5). Since observa-

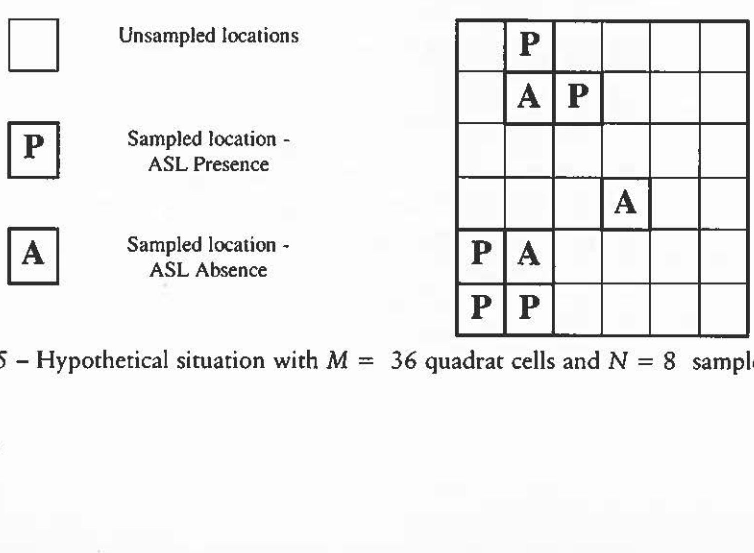


Fig. 5 – Hypothetical situation with *M* = 36 quadrat cells and *N* = 8 sampled locations.

tions are not available for all the random variables which constitute the spatial process, we are in the statistical situation termed "incomplete-data problem". Some solution to this problem has been proposed in the literature (see DEMPSTER, LAIRD, RUBIN 1977), and modifications to account for spatial Markov dependency have been attempted (QIAN, TITTERINGTON 1989; RATHBUN, CRESSIE 1994). However the effectiveness in the case in hand needs to be studied and tested in practical cases.

5. CONCLUSIONS AND RESEARCH AGENDA

In the present paper we have criticised the use of the logistic regression for the production of ASL probability maps proposed by KWAMME (1983), an approach known in the archaeological literature as the "integrated strategy". The application of the method is statistically incorrect since in archaeological studies it is violated the hypothesis of independence between cells which is at the basis of the logistic regression model.

To overcome such limitations we have proposed two alternative models. The first one is an autopredictive model in which the probability of ASL is modelled as a function of the observations coming from field surveys in neighbouring zones. This approach accounts for the problem of non-independency of observations, but neglects *a priori* auxiliary information on the archaeological area. The second approach is a more comprehensive one which overcomes the problems of logistic regressions while preserving the role of *a priori* information.

The research agenda in the field presents three major issues. First of all it is necessary to test the autologistic model (Section 3) and the log/autolog model (Section 4) on real data sets. Secondly the results obtained on a set of training sites need to be compared with those obtained by using the standard "integrated approach" based on logistic regression. Finally a purely statistical problem needs to be faced while devising good estimators for the log/autolog model where (as described in Section 4) we face an incomplete-data problem.

GIUSEPPE ARBIA
Department of Quantitative Methods and Economic Theory
University "G. D'Annunzio" - Pescara

GIUSEPPE ESPA
Institute of Statistics and Operational Research
University of Trento

BIBLIOGRAPHY

ARBIA G. 1993, *Recenti sviluppi nella modellistica spaziale*, in S. ZANI (ed.), *Metodi Statistici per le Analisi Territoriali*, Milano, Franco Angeli, 193-217.

ARROYO-BISHOP D. 1995, MEDAIS: *The potential contribution of GIS to the identification,* *study and conservation of the cultural heritage of the Mediterranean basin,* in *Proceedings of the First International Congress on Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin*, Catania-Siracusa, 103.

BENEDETTI R., ESPA G. 1995, *The use of remotely sensed images as auxiliary information for the interpolation of archaeological data,* in *Proceedings of the First International Congress on Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin*, Catania-Siracusa, 115.

BESAG J.P. 1974, *Spatial interaction and the statistical analysis of lattice systems* (with discussion), «Journal of the Royal Statistical Society», 36, B, 192-235.

CARLA' R., CARRARA A., JACOLI M., ALESSANDRO V., BARONTI S. 1995, *Analysis of multispectral imagery for archaeological investigation,* in *Proceedings of the First International Congress on Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin*, Catania-Siracusa, 129.

COX D.R., SNELL E.J. 1989, *Analysis of Binary Data*, 2nd ed., London, Chapman & Hall.

CRESSIE N. 1992, *Statistics for Spatial Data*, New York, John Wiley and Sons.

DEMPSTER A.P., LAIRD N.M., RUBIN D.B. 1977, *Maximum likelihood for incomplete data via the EM algorithm,* «Journal of the Royal Statistical Society», B, 39, 1-22.

ISING E. 1925, *Beitray sur theorie des ferromagnetismus,* «Zeitschrift Physic», V31, 253-258.

KOHLER T.A., PARKER S.C. 1986, *Predictive models for archaeological resource location*, in M.B. SCHIFFER (ed.), *Advances in Archaeological Method and Theory*, New York, Academic Press, 9, 397-452.

KWAMME K.L. 1983, *Computer processing techniques for regional modelling of archaeological locations*, «Advances in Computer Archaeology», 1, 26-52.

QIAN W., TITTERINGTON D.M. 1989, *On the use of Gibbs Markov chain models in the analysis of images based on second-order pairwise interaction distributions,* «Journal of Applied Statistics», 16, 267-281.

RATHBUN S.L., CRESSIE N. 1994, *A space-time survival process for a longleaf pine forest in Southern Georgia,* «Journal of the American Statistical Association», 89, n. 428, 1164-1174.

TOBLER W.R. 1970, *A computer movie simulating urban growth in the Detroit Region*, «Economic Geography», suppl. 46, 234-240.

WARREN R.E. 1990, *Predictive modelling of archaeological site location: a case of study in the Midwest*, in K.M.S. ALLEN, S.W. GREEN, E.B.W. ZUBROW (eds.), *Interpreting Space: GIS and Archaeology*, New York, Taylor & Francis, 90-111.

ABSTRACT

Forecasting statistical models are becoming increasingly important in archaeological research. One of the reasons of this popularity is that archaeological sites tend to present themselves in particular environments so that forecasting models can help in identifying areas where the probability is higher based on previously collected statistical information.

In the present paper we consider a class of statistical models designed to produce maps of the probability for archaeological site location (ASL) which incorporate both deductive and inductive considerations. In the discussion we criticise the use of the logistic regression for the production of ASL probability maps, a popular approach known in the archaeological literature as the "integrated strategy". The application of the method is statistically incorrect since in archaeological studies the hypothesis of independence between sites, which is at the basis of the logistic regression model, is violated.

To overcome such limitations we propose two alternative models. The first one is an autopredictive model in which the probability of ASL is modelled as a function of the observations coming from field surveys in neighbouring zones. This approach accounts for the problem of non-independency of observations, but neglets *a priori* auxiliary information on the archaeological area. The second approach is a more comprehensive one which overcomes the problems of logistic regressions while preserving the role of *a priori* information.

Archeologia e Calcolatori
7, 1996, 373-384

OBSERVATIONS SUR LA DISTRIBUTION DES TOMBES DANS UNE NÉCROPOLE GRECQUE D'ÉPOQUE ARCHAÏQUE: LE CAS DE LA NÉCROPOLE OUEST DE MÉGARA HYBLAEA

1. Introduction

La nécropole grecque d'époque archaïque présente une dispersion dans l'espace que les archéologues qualifient souvent *d'aléatoire ou au hasard.* La nécropole Ouest de Mégara Hyblaea, fouillée par Paolo Orsi en 1889 (Orsi, Cavallari 1892), présente une telle caractéristique. L'étude réalisée a eu pour but d'effectuer des observations quant à la dispersion des tombes dans l'espace de la nécropole en relation avec les données archéologiques.

Sur un plan plus général la recherche vise une forme d'économie générale de l'espace de la nécropole qui pourrait se profiler derrière une mise en activité des données archéologiques.

A l'horizon d'une telle recherche on trouvera un fonds commun de questionnements:

Jusqu'où le caractère archéologique des données employées est en mesure de contribuer à une telle connaissance?

Dans quelle mesure, faire travailler les données archéologiques entre elles, n'aboutit pas à une notion de syntaxe qui dépasse l'observation archéologique isolée?

Y a-t-il dérivation ou clivage de l'espace funéraire historique à l'espace étudié à caractère archéologique?

2. Présentation de l'étude

L'étude a été menée selon trois phases qui pour l'essentiel recouvrent trois outils à base d'informatique:

1. Reconstitution de l'espace de la nécropole à base de Conception Assistée par Ordinateur (CAO).
2. Élaboration des variables par une analyse du vocabulaire et de la syntaxe à l'aide d'une Base de données et d'Intelligence Artificielle (IA).
3. Calcul et étude des variables dans l'espace à base de statistiques descriptives et leur manipulation dans l'espace reconstitué.

2.1 *1ère phase*

Pour des raisons de cohérence de l'exposé, nous n'allons pas présenter les travaux relatifs à la première phase. L'espace en archéologie, est la plupart du temps une donnée qui n'est pas à reconstituer. Or, dans la publication

Fig. 1 – La distribution des tombes de la nécropole Ouest de Mégara Hyblaea (reconstitution réalisée par CAO).

citée il n'y a pas de relevé de l'ensemble des tombes de la nécropole Ouest, mais plusieurs croquis des différents secteurs fouillés et qu'il a fallu assembler. Les problèmes méthodologiques que l'étude a posés ont un intérêt évident pour des reprises d'anciennes fouilles. Nous nous limiterons à l'emploi du relevé général des tombes obtenu lors de cette première manipulation (Fig. 1).

2.2 $2^{ème}$ *phase*

La deuxième phase constitue un domaine particulier qui vise à une mise en activité des données archéologiques recueillies. Elle a pour but l'élaboration des variables à étudier partant des données archéologiques. Nous avons présenté ailleurs (IACOVELLA, AUDA 1994) les principes qui nous conduisent à différencier les variables archéologiques strictes, des variables soumises au calcul dans l'étude d'un contexte archéologique. Ici nous allons réaffirmer l'importance d'une telle distinction en se donnant des outils qui permettent d'observer de près ce mouvement de dérivation qui met en jeu, outre les classifications archéologiques relatives au matériel, l'observation du discours qui les cimente. Nous avons observé les différences de comportement des variables dans le calcul, lorsqu'elles reposent sur l'unique critère du classement archéologique et lorsqu'elles sont réajustées par l'étude du contexte où elles ont été annoncées. Grâce à des contours bien délimités, ce type d'étude offre des champs d'expérimentations adaptés à l'emploi de l'IA. L'observation au plus près des variations, dans la définition des variables permet de mieux mesurer leur impact au moment du calcul.

Dans notre cas, nous avons constaté comment le glissement dans la définition d'une variable soumise à un même calcul, se trouve affecter différemment la distribution des tombes dans l'espace où les données ont été recueillies.

2.3 *3ème phase*

Nous avons regroupé dans la troisième phase deux traitements distincts qui sont le calcul à proprement parler et la représentation de ses résultats dans l'espace archéologique à étudier.

Au niveau de la représentation des résultats dans l'espace, il ne s'agit pas du seul changement de forme: du quantitatif au graphique. L'activité liée à la représentation des résultats dans l'espace constitue une forme de mise en opération à proprement parler, dont le but est de revenir à l'observation de toutes les variables archéologiques initiales afin de dégager leur impact sur la distribution des tombes. De cette façon se dégagent des contours conformément à la pertinence des associations livrées par le calcul. Dans un second temps, les variables retenues comme pertinentes sont replacées dans le relevé général des tombes pour observer si oui ou non elles contribuent à une clarification des distributions *aléatoires ou au hasard*.

3. LE CAS DES ENFANTS DE MEGARA HYBLAEA

Nous allons illustrer quelques résultats de l'étude en nous limitant au cas des enfants. Les raisons de ce choix tiennent aux enseignements obtenus dans la deuxième phase de l'étude, que nous avons appelée élaboration des variables. En effet l'étude montre que les classifications archéologiques relatives au matériel se déclinent de manière particulièrement fine dans le cas de tombes que Paolo Orsi rapporte à des enfants.

L'analyse du vocabulaire employé démontre la richesse des variations obtenues dans la désignation de l'enfant de Mégara Hyblaea et traduit dans le même temps une implication particulière du fouilleur (Fig. 2).

3.1 *Les enfants Orsi*

Il s'ensuit une double constatation: d'abord l'existence d'un enfant explicitement désigné comme tel par le fouilleur, ensuite un groupe d'enfants implicite qui possède les caractéristiques du précédent mais que le fouilleur n'a pas attribué à des enfants en tant que tels. L'ensemble résultant de ces deux groupes se définit par le seul critère de la classe d'âge dont la paternité revient totalement au fouilleur; nous les appelons les enfants Orsi.

Toujours en opérant dans les limites de la syntaxe Orsi, ce groupe permet déjà quelques observations quant à la répartition de ces enfants dans l'espace de la nécropole.

Dans la Fig. 3 la distribution des tombes à inhumations contenant des enfants *(124 cas)*; dans la Fig. 4 celle des tombes à crémation *(21 cas)* (nous ne discutons pas du comment Orsi peut-il affirmer, dans le cas d'une crémation qu'il s'agit d'un enfant); dans la Fig. 5 on trouvera la distribution des tombes d'enfants qui ont été simulées d'après les caractéristiques livrées par le fouilleur *(33 cas)* (ex: une inhumation dans un vase de 50 cm, ou une déposition dans une petite tombe de 70 cm de long...) et la Fig. 6 résume tous les enfants Orsi de la nécropole occidentale *(178 cas)*.

Les remarques faites par le fouilleur, dans son rapport, au sujet de l'existence de secteurs d'enfants, ne se laissent pas vérifier de manière décisive dans les résultats obtenus par notre étude. Dans l'ensemble (Fig. 6), la distribution des enfants dans l'espace n'épargne aucune zone particulière (tenant compte que les parties est et nord de la nécropole ont subi des pillages de tombes). L'observation montre de manière plus précise que l'enfant affecte plus ou moins l'espace en question sans jamais y être franchement exclu et que des densités plus importantes se laissent grossièrement délimiter, en particulier dans la distribution des inhumations (Fig. 3) on constate tant à l'extrémité nord-est que celle de sud-ouest ainsi que sur la diagonale nord-ouest, sud-est des séries de tombes contigües à inhumations attribuées à des enfants.

3.2 *Les enfants de l'étude*

Après cette première étude basée sur les enfants Orsi, nous avons rapproché les deux séries d'observations: vocabulaire et classifications archéologiques en deux tableaux concernant d'une part ce qui est relatif à la tombe elle-même y compris les dépositions, et d'autre part le seul mobilier.

Nous nous sommes limités à injecter ces observations dans la

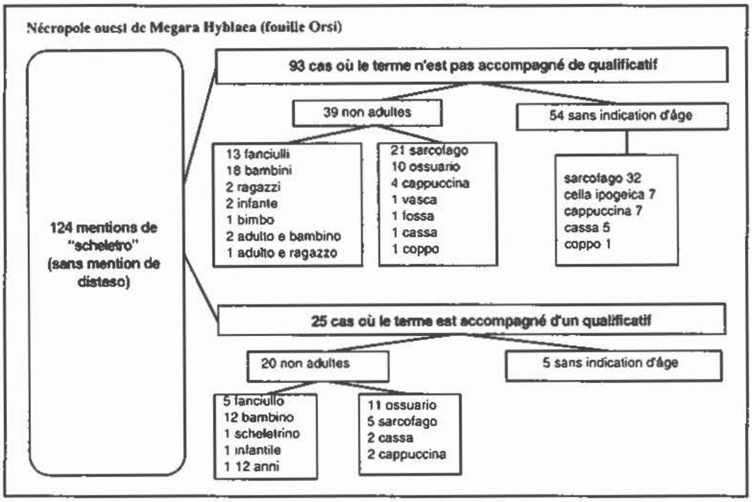


Fig. 2 – Exemple de distribution du terme "scheletro" dans le rapport Orsi.

Fig. 3 – Distribution des enfants Orsi dans les tombes à inhumation.

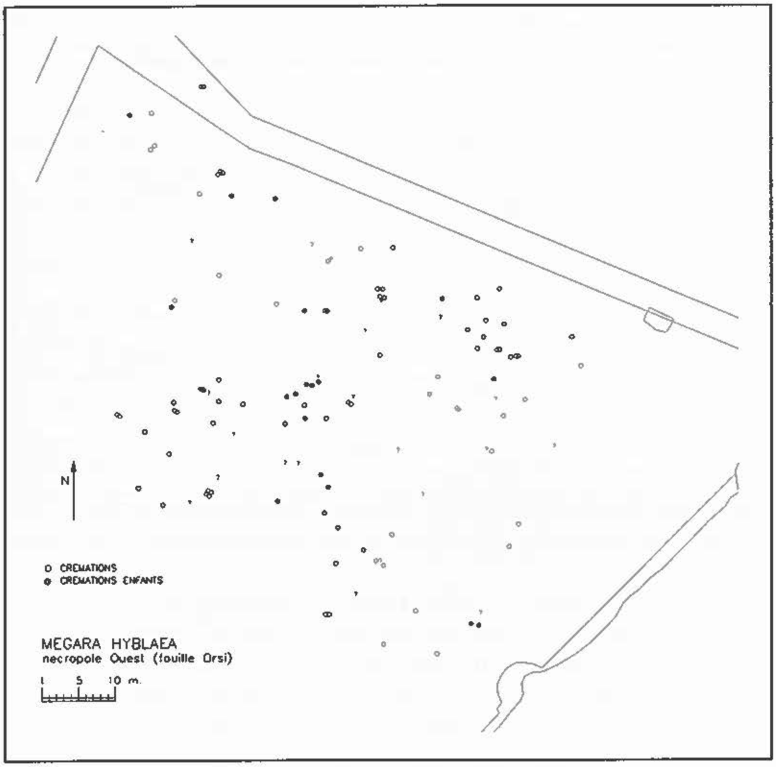


Fig. 4 – Distribution des enfants Orsi dans les tombes à crémation.

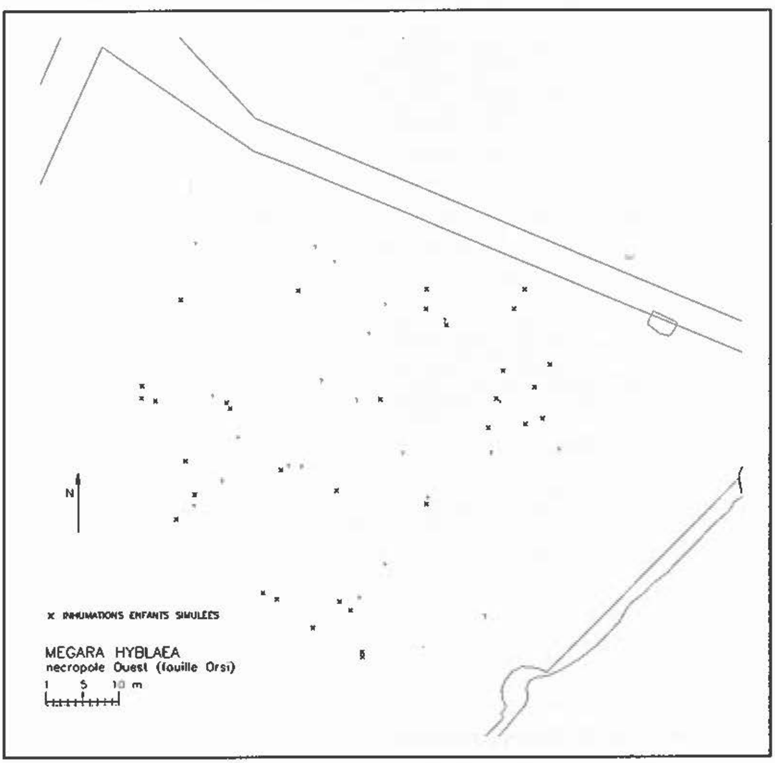


Fig. 5 – Distribution des enfants simulés selon les règles du rapport Orsi.

Fig. 6 – Distribution des 178 enfants Orsi dans la nécropole ouest de Mégara Hyblaea.

caractérisation des variables purement archéologiques afin d'obtenir de nouvelles variables indicées par l'orientation en question. C'est ce dernier lot qui a été soumis à des analyses statistiques[1].

3.3 *L'AFC sur les variables du mobilier*

La table de contingences du mobilier funéraire (quantité de mobilier par type trouvé dans chaque tombe) se prête à une Analyse des Correspondances (AFC de 248 lignes correspondant aux tombes contenant

[1] Pour le calcul nous avons utilisé la *Programmathèque ADE. Analyses multivariées et expression graphique des données environnementales, Version 3.6*, DOLEDEC S., CHESSEL D., *Écologie des Eaux Douces et des Grands Fleuves*, URA CNRS 1451, Université Claude Bernard Lyon 1 – 69622 Villeurbanne Cédex – F, ou encore http://biom1.univ-lyon1.fr:8080/

du mobilier et 12 variables caractérisant le mobilier funéraire). A l'issue du calcul (Fig. 7), les 4 groupes proposés par l'analyse ont été à leur tour injectés en tant que tels dans le tableau suivant pour caractériser le mobilier funéraire de chaque tombe, avec ajout d'un cinquième groupe correspondant au cas des tombes vides de mobilier. On remarquera en particulier le groupe 2, comprenant 44 cas, dont le mobilier est fortement marqué par la présence de vases qualifiés de petits, très petits ou encore de jouets et qui correspondent à des tombes attribuées par le fouilleur à des enfants.

3.4 *L'ACM sur les tombes avec le mobilier*

Les variables correspondant à la description des tombes sont de nature qualitative, d'où l'adoption d'une Analyse des Correspondances Multiples (ACM de 359 lignes correspondant à toutes les tombes fouillées et 6 variables les décrivant avec un total de 34 modalités dont font partie celles issues de l'analyse précédente). A l'issue du calcul (Fig. 8), 4 groupes de tombes se laissent définir, comprenant 95 tombes. On notera en particulier le groupe 3 où nous retrouvons le groupe 2 de l'analyse précédente mais élargi à des tombes uniquement à vases comprenant une unique déposition ainsi qu'un groupe de tombes indéterminé, c'est à dire sans trace de sépulture mais dont le profil se rapproche de ces tombes attribuées à des enfants.

3.5 *Résumé des analyses*

Nous nous sommes limités à l'observation de la distribution des deux groupes obtenus, pour lesquels nous avons tracé les cartes avec les positions respectives des tombes.

Ainsi dans la Fig. 9, le groupe 2 de l'AFC sur le seul mobilier, compte 44 tombes, dont une déclinaison particulière du mobilier caractérise un groupe d'enfants. Les tombes sont réparties dans l'ensemble de la nécropole avec une très nette tendance à être groupées en noyaux entre elles.

Dans la Fig. 10, le groupe 3 de l'ACM compte 95 tombes dont 22 appartiennent au groupe précédent, c'est à dire marqué par un même profil de mobilier. Parmi les 73 cas restants nous trouvons 59 tombes à caractère indéterminé, c'est à dire avec présence de modalités manquantes et qui expriment la difficulté à préciser une catégorie particulière d'enfants. La distribution de ces tombes dans l'espace reste calquée sur la précédente, c'est à dire sous forme de noyaux à l'exception du nord-ouest et du sud-est.

4. Conclusions de l'étude des enfants

A l'intérieur du texte Orsi l'étude a délimité dans un premier temps les contours d'un continent qui s'est avéré particulièrement soigné par le fouilleur lui-même: la désignation des enfants.

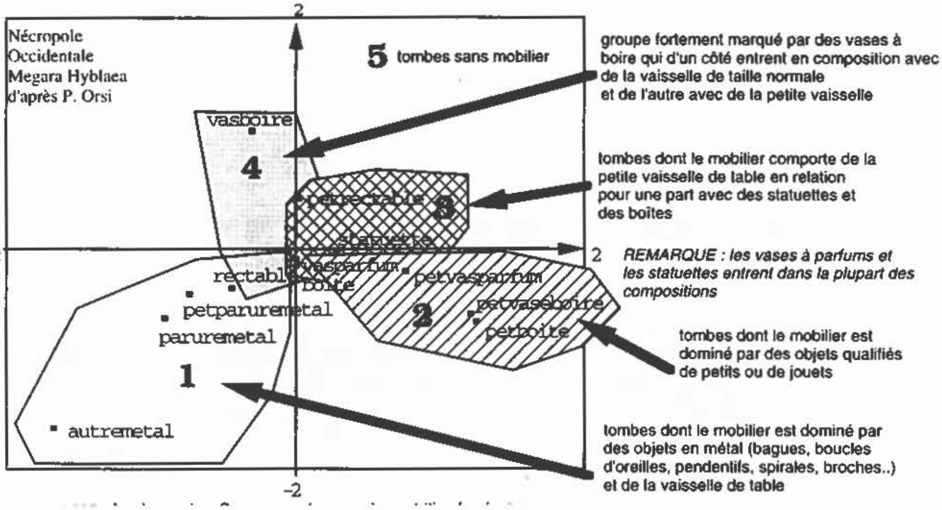


Fig. 7 – Analyse des Correspondences du mobilier funéraire. 248 lignes, 12 variables – Définition des groupes pour étude topographique.

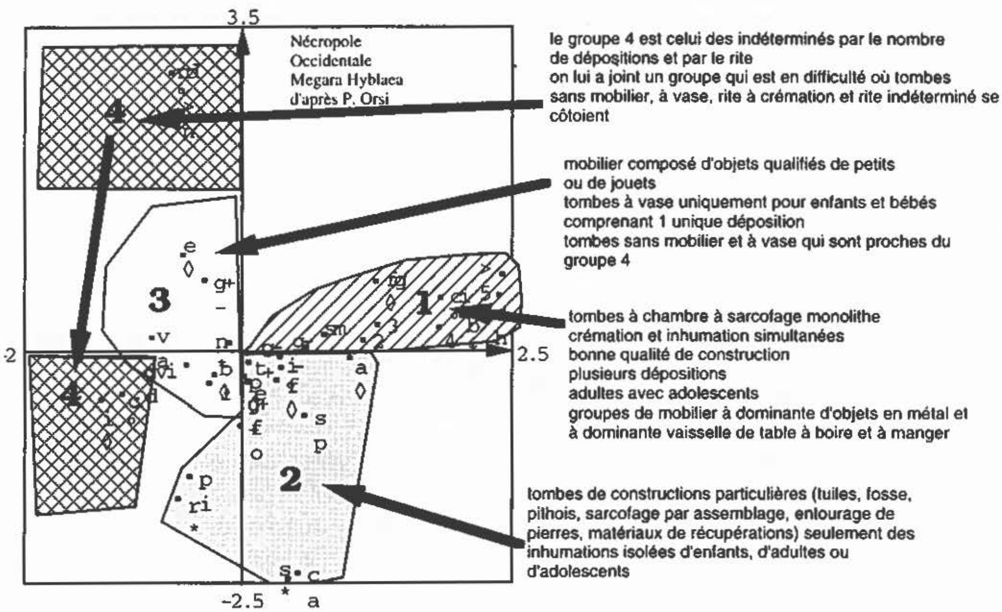


Fig. 8 – Analyse des Correspondences Multiples des tombes + mobilier codé par la Fig. 7. 359 lignes, 6 variables, 34 modalités. Définition des groupes pour étude topographique.

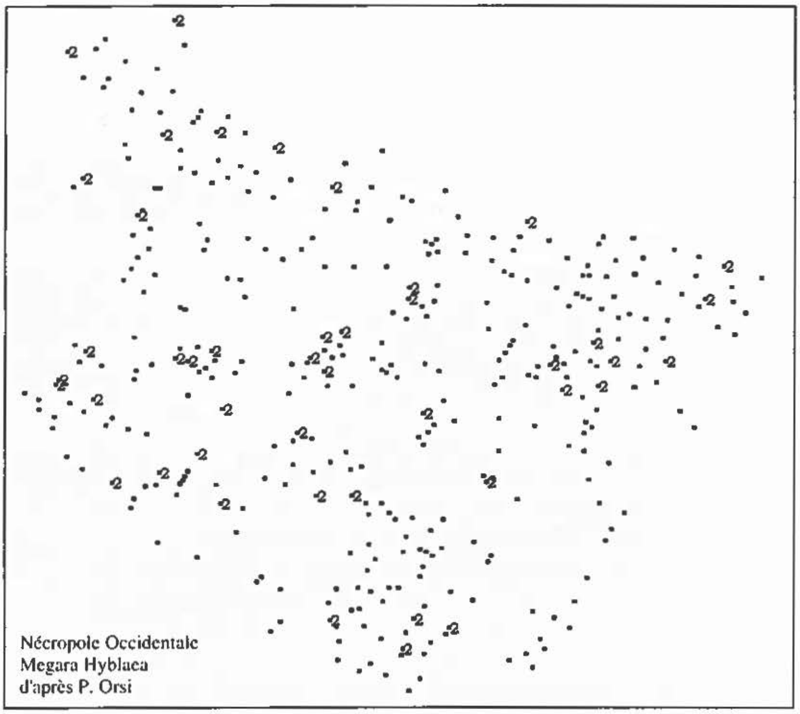


Fig. 9 – Analyse des Correspondances du mobilier funéraire. Situation topographique des tombes dont le mobilier appartient au Groupe 2 (Fig. 7) f1/f2 (+-).

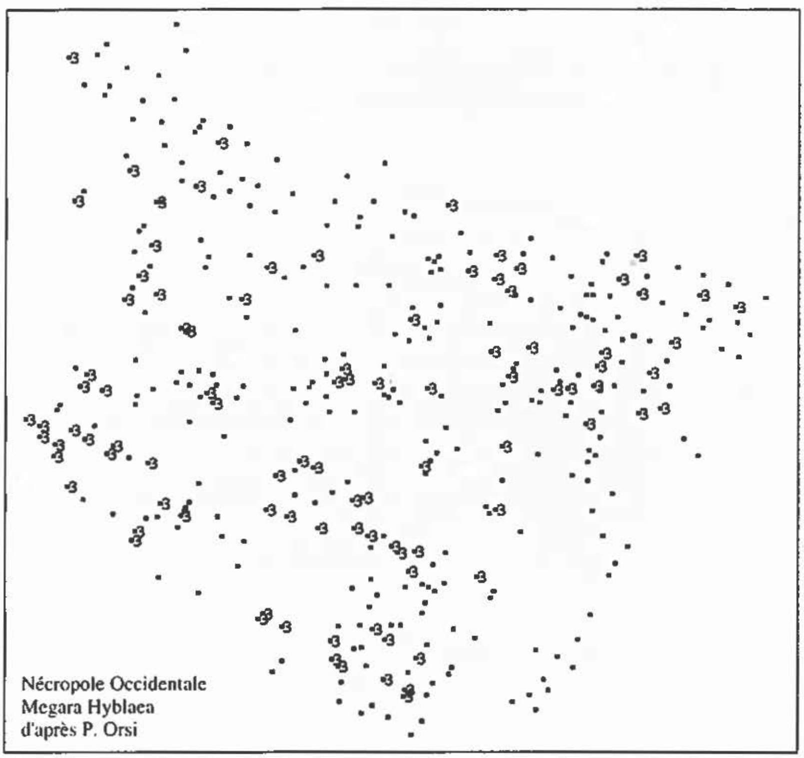


Fig. 10 – Analyse des Correspondances Multiples des tombes + mobilier. Situation topographique des tombes appartenant au Groupe 3 (Fig. 8) (f1 de -0,5 à 2,5, f2 de 0 à 1,75).

Dans un second temps, l'étude des 178 enfants Orsi désignés par le seul critère de la classe d'âge, aboutit à une connaissance renouvelée des données. En particulier par l'apport des caractéristiques du matériel archéologique et de particularités mises en évidence par l'observation du vocabulaire employé par le fouilleur.

Enfin dans un troisième temps, ce nouvel espace livre des contours intérieurs qui délimitent des groupes inédits. Dans le total des 178 cas, un groupe comprenant 44 cas se différencie des autres par un même profil dans la composition du mobilier funéraire, un deuxième groupe réunissant 73 cas est marqué par une coutume à déposer des defunts dans des vases, le dernier groupe qui compte 71 cas est plus complexe, il réunit les plus jeunes des enfants ainsi que des enfants Orsi qui entrent dans d'autres compositions; nous y trouvons les enfants déposés avec des adultes ou encore des enfants dans des tombes d'adultes (quasiment que des sarcophages). Ce dernier groupe se caractérise majoritairement par des tombes à caractère monumental et à mobilier comprenant des objets en métal que l'analyse montrent corrélés, pour une moindre part on trouve des enfants associés à un groupe de tombes à construction particulière et marquées par une quasi absence de mobilier.

Tout au long de l'étude la situation topographique des tombes a été mise à contribution pour délimiter de manière décisive le tracé des contours intérieurs de chacun des groupes. C'est ici un moment privilégié de l'étude où les chemins de l'analyse cèdent à une représentation chaque fois inédite de l'espace funéraire. C'est dans cette perspective que les contours délimités participent d'une économie générale de l'espace qui tout en affirmant sa différence avec l'espace urbain tout proche, ne cesse de s'en réclamer.

5. Perspectives

Au-delà des chiffres le but était de montrer que par rapport à la seule catégorie Orsi, l'étude fait apparaître un enfant plus nuancé en rapport aux données archéologiques. Et l'étendre ce même type d'étude aux autres composantes de la nécropole. De la sorte, l'ensemble des contours délimités à l'intérieur de la nécropole figurent une forme particulière d'activité de l'espace. Dans le cas présenté nous avons remarqué que les groupes d'enfants quels qu'ils soient ont toujours tendance à se présenter sous la forme de petits groupes, voire de noyaux qui essaiment sans épargner pour autant une portion significative de l'espace. Contiguïté donc entre les tombes, voisinage entre les groupes et continuité dans le temps, montrent que l'espace funéraire de l'enfant n'est pas clivé à Mégara Hyblaea du milieu du VII au début du V siècle Av. J.C. Mais la distribution aléatoire ne résulte pas de l'impact direct d'une ou plusieurs variables isolées, mais d'une activité de certaines d'entre elles.

La formation des groupes observés dans la distribution topographique des tombes incite à poursuivre l'exploration des autres tombes et des autres composantes archéologiques car c'est par l'approfondissement de ses contours intérieurs que l'espace funéraire d'époque grecque archaïque marqué par cette distribution AU HASARD des tombes, se prêtera à de nouvelles configurations.

ANDREA IACOVELLA
Service Informatique
École Française d'Athènes

BIBLIOGRAPHIE

ORSI P., CAVALLARI F.S. 1892, *Megara Hyblaea Storia – Topografia – Necropoli e Anathemata*, Monumenti Antichi Pubblicati dall'Accademia dei Lincei, I, col. 685-950.

IACOVELLA A., AUDA Y. 1994, *Nécropoles de Sicile: Étude de l'utilisation des espaces funéraires dans le temps (du IXème au Ier S.AV.JC)*, «Archeologia e Calcolatori», 5, en particulier 69-75.

ABSTRACT

The aim of the study is to observe the random distribution graves, characteristic of Greek cemeteries in archaïc period. From the three computing tools used, only the data analysis is explained. The results were displayed in the topographical space, by the coordinate x,y of each grave.

The definition of archaeological variables destinated to the data analysis has grown by analysis of the vocabulary used by the archaeologist. Correspondence Analysis was used for the study of grave-goods, and Multiple Correspondence Analyses for the study of burials characterised by the grave-goods structure of the previous analysis.

The report concerns one part of results: the whole 178 child burials described by the archaeologist. The study shows four groups characterized by different archaeological patterns. Each group was inserted in the topography of cemetery. The internal outlines delimited by the groups are the components of the distribution model. This last must be extended at the others structures of the archaeology of death.