

17. Quantitative Methods in the 1990s

Clive Orton

Institute of Archaeology, University College London, 31-34 Gordon Square, London WC1H 0PY, U.K.

17.1 Introduction

Archaeologists of all people should know that to see where we are going, we need to know from where we have come. So before looking at the 1990s we need to go back and look at the role of quantitative methods in archaeology throughout this century.

17.2 A brief history

For most of the century, there was really only one quantitative method in archaeology — seriation, invented at the end of the 19th century (Petrie 1899) and enthusiastically adopted in the USA early in the 20th century (Kroeber 1916). Right through to the 1960s it was a true archaeological tool, devised to meet archaeological needs that did not exist in other disciplines, consisting of various *ad hoc* procedures based on a simple model with little theoretical input. But above all, it was designed by archaeologists for archaeologists and it worked.

Things began to change in the 1960s. Several reasons could be put forward:

- (i) the explicitly scientific approach of the New Archaeology,
- (ii) the arrival of growing amounts of quantitative data from scientific techniques (e.g. C14 dating, XRF and other analytical techniques), possibly traceable to the founding of the Oxford Laboratory in 1957 and other similar institutions,
- (iii) the growing availability of computer power and the software needed to carry out sophisticated analyses (e.g. multivariate statistics),
- (iv) cross-fertilisation from other disciplines, notably biology and geography.

The first two may have supplied the motivation, the third the opportunity, but the fourth provided the weapon, i.e. the methodology. The 1960s saw the introduction of taxonomic ideas and structures from biology, with cluster analysis as the leading method. Indeed, when I started my involvement in quantitative archaeology in 1966 our 'bible' was *Principles of Numerical Taxonomy* (Sokal & Sneath 1963). Geography underwent its quantitative revolution in the 1960s, and its reverberations were felt in archaeology in the 1970s with the growth of spatial analysis (Hodder & Orton 1976). Central Place Theory and locational analysis came in as we studied site distributions and settlement patterns under the aegis of Christaller and Lössch. Overlapping, but perhaps a little later, came intrasite spatial analysis, based on models taken from plant ecology, with nearest-neighbour and quadrat analyses, with thanks to Clark and Evans, Grieg-Smith and Pielou.

Inevitably, reaction came in the 1980s. It was pointed out (quite rightly)

- (i) that artefacts do not (except possibly in dark

- cupboards) breed and evolve like living species,
- (ii) that the main problem with archaeological sites is not their neighbours but whether we know where they are in the first place and even (according to some iconoclasts) whether they actually exist, and
- (iii) that flint scatters do not grow like daisies on the lawn.

At the same time, the inappropriateness and statistical naivete of many early applications of general quantitative techniques (e.g. factor analysis) were exposed (Thomas 1978; Aldenderfer 1987), and we questioned whether the emperor had any clothes. Archaeologists demanded techniques that had been devised by archaeologists with archaeological models and data in mind — the call for 'congruency' (Whallon 1984:242). An early example was the k-means version of cluster analysis (Doran & Hodson 1975:180–5) which had already emerged out of dissatisfaction with other techniques. Perhaps the area where this approach went furthest was in intrasite spatial analysis, with techniques such as local density analysis (Johnson 1977) and unconstrained clustering (Whallon 1984) being developed by archaeologists, who at the same time managed to ignore the very general theoretical developments being made by statisticians such as Besag, Diggle and Ripley (e.g. Ripley 1977). This we might call the Audrey Syndrome ('An ill-favoured thing, sir, but mine own', *As You Like It*, Act V Scene iv). The gut feeling was right but it led to a blinkered approach.

So where do we go from here? Anyone expecting me to list new techniques which will appear in the 1990s will be disappointed, for two reasons:

- (i) I do not possess the gift of prophecy,
- (ii) If I did, I would be busy writing grant applications instead of giving this paper.

All I can do is to present my personal agenda for the 1990s and hope to carry some of you with me.

17.3 An agenda for the 1990s

There are four areas where I hope progress can be made:

- (i) methodological progress in specific well-defined areas, e.g. intrasite spatial analysis and multiple correspondence analysis;
- (ii) greater attention to the quality of data, involving greater use of sampling in order to release the necessary resources;
- (iii) the integration of different types of data, e.g. C14 dates and stratigraphy, pottery and 'small finds';
- (iv) education — the techniques are there, but archaeologists must appreciate their value and learn how to use them.

17.3a Methodological improvements

17.3a (i) Intra-site spatial analysis

Here a discrepancy between models and techniques seems to have grown in the 1980s. The aim seems to be to detect 'activity areas': distinct zones characterised by the spatial association of two or more types of artefact and/or ecofact, which can be interpreted as relating to the same activity. By definition areas should have edges, which means that the associated types should have discontinuous spatial distributions. But the techniques, e.g. unconstrained clustering (UC), rely on smoothing and contouring these distributions, which only make sense if they are continuous. The theoretical discrepancy shows itself in the practical problems that arise when UC is used, such as the creation of spurious clusters which have to be explained away by the archaeologist (Orton 1987). Other odd features include spurious characterisation — a cluster can be defined as having 'n% of type x' when in fact there are no examples of that type within its area (Blankholm 1991: 106).

Clearly, we need methods that actually look for edges. There are at least three such:

- (i) the 'Poisson forest' approach (Ripley & Rasson 1977). This has been used with apparent success to detect the edges of grave-cuts which had been obliterated by soil-leaching (Alvey 1983). Later thinking suggests that all the burials were in one large grave cut (Glover, pers comm.), but that does not invalidate the method;
- (ii) change-point methods (Buck *et al.* 1988), which specifically look for changes in levels of a variable along spatial transects, i.e. for edges. Although designed for 'surface' problems (i.e. readings taken at specific points of a grid), I would like to see them extended to point-patterns. There would at least be no problem of extending them to counts of points in cells of a grid;
- (iii) more sophisticated techniques of image analysis, such as segmentation (Buck & Litton 1989). I doubt if the full power of such techniques would often be needed, as we are usually looking for simple patterns. However, the problem of whether the edges of the distributions of two or more types coincide or not, could keep statisticians busy for some time.

17.3a (ii) Multiple correspondence analysis

References to correspondence analysis (CA) in the archaeological literature have burgeoned since the early 1980s, but the technique seems to be little known outside a magic circle of 'correspondents'. I was surprised at a recent conference to be asked to explain the technique, which was apparently unknown to most of the audience. But even as currently understood, conventional CA is inadequate for our needs, as it can cope with only two-way tables. There is a strong need to go beyond that to at least three-way ones. A partial answer seems to be multiple CA, known since the late 1970s (Lebart *et al.* 1984) and well-established in the social sciences (e.g. Schiltz 1990). Implementation in

archaeology would be a step forward, but not the whole answer, as the technique looks only at the relationships between pairs of variables.

17.3b Sampling and the quality of data

Computer databases now form part of the bedrock of archaeology, as a study of the papers given at CAA shows (35% of all papers in the 1980s). We have had endless discussions of database design and construction, and of the relative merits of different hardware and software, but surprisingly little has been said about the data themselves. How accurate are they? At the end of the day, the usefulness of our databases depends on the attention span of some poor clerk (perhaps called an archaeologist) punching data in on a wet Friday afternoon and thinking of the week-end. My limited experience as a user suggests that our databases may have a significant proportion of errors, but do we care? and what can we do about them?

This brings me to sampling: another of the topics with a chequered history in archaeology. It was introduced on a large scale in the 1970s (Mueller 1975; Cherry *et al.* 1978) to meet two very different needs:

- (i) to cope with the realisation that every archaeological collection is a sample from some (usually ill-defined) population, and that it is the population that matters, not the sample. For example, to say '10% of this pottery recovered from this pit is Alice Holt ware' invites the answer 'so what?', but to say '10% of the pottery in use at this site in the 4th century came from Alice Holt' is potentially more interesting;
- (ii) to supply the rationale for sample field surveys, e.g. in environmental impact assessments.

The first raised very difficult problems of modelling site formation processes, with which we are still struggling today. It is not as simple as we thought, and has led us to question what we can reasonably say. To continue my example, it may only be possible to make comparative statements like 'The proportion of Alice Holt ware in the 4th century was twice as high at site A as at site B', without giving the proportions themselves (Orton *et al.* forthcoming).

The second led to some of the most bizarre applications of statistics in the history of archaeology (Hole 1980). Archaeologists sought desperately for the Holy Grail of the percentage size which would make their samples respectable, and failed because it does not exist.

Sampling comes into our present discussion because greater care over our data needs resources which in the present climate can only be released by putting in fewer data. What scope is there? There is a spectrum, running from the necessity of 100% coverage (e.g. SMRs) to an equally strong necessity for very small samples (e.g. scientific analyses). For the former, sample spot-checking allied to quality control techniques seem essential. Where sampling is a possibility, do we design samples to minimise the workload? As an example I quote my work on the condition of museum collections. Here relatively small samples, collected in a short time, provided useful information. A designed sample of

1500 objects in the Social History Collection of the Museum of London could give as much information about the condition of the collection as an undesigned sample of 7500 (Keene & Orton this volume). While the surveying time would have been much the same, because of the greater overheads associated with the designed sample, the costs of entering and maintaining the data collected into a database would be less, and the scope for error reduced.

In general, we need to ask,

- (i) if we are already sampling, could our samples be better designed and more efficient? If people are still asking 'what percentage do we need to sample?' the answer is almost certainly 'yes'.
- (ii) If we are not sampling, is there scope for introducing sampling, not as a cost-cutting measure but as part of a plan to improve the quality of our data?
- (iii) Even if sampling is not an option, can we use sample checking to improve or maintain the quality of our data?

17.3c Integration

For as long as I can remember, archaeologists and especially finds workers, have moaned about the 'bitty' nature of excavation reports: the site report, the pot report, the small finds report, and so on, with the interesting bits either in appendices or (worse) on fiche, and all apparently written in isolation. This has been blamed on (amongst other things) over-rigid division of labour in Units, or the need to put work out to specialists whom no-one understands except other specialists, or just tradition. I suggest that part of the problem lies in our analytical techniques, which force us to study different classes of finds separately, and all of them separately from the site. We need to use the statistical techniques available to bring together data from various sources. The analysis I found myself doing of the medieval small finds from Winchester (Barclay *et al.* 1990) relates classes of find to type of site in what I hope is an interesting way. My own project (*Pie-slice*; Orton & Tyers forthcoming) enables us to use data on pottery (and other 'broken' finds like tile and bones (Rackham, pers comm.)) on an equal footing with the countable classes of find.

A very important point is to get the right level of classification of both finds and site. For sites, a level between the feature and the site seems best (e.g. the phase; the building); for finds, the functional group; for pottery, the functional group, or (for chronology) the ware or fabric type. Too much detail and the random 'noise' obscures any pattern; too little detail and there is no pattern.

17.3d Bayesian statistics

I have for many years advocated the use of the Bayesian approach to statistics as a way of making better use of the statistical information available to archaeologists (see Orton 1980:220), but must confess to have done very little about it. In this approach we see statistics as 'the orderly influencing of opinions by data' and seek to modify our prior belief (which may

itself be based on earlier data) about an archaeological situation by the application of data to arrive at a posterior belief. This approach recognises that we rarely work *in vacuo*, but start from a state of partial understanding about our archaeological problem. The deterrent to using it has been the lack of accessible software and the ferociously heavy computational needs. Both are now being overcome, and useful work is beginning to emerge. For example, the approach has been used to incorporate prior knowledge about stratigraphy into the interpretation of series of C14 dates (Buck *et al.* forthcoming).

17.3e Education

The challenge of the 1990s is not so much to develop new techniques as to achieve a wider and more fruitful usage of what we already have. Much of this will be quite 'low-tech': it is not always the most sophisticated techniques that give the most useful results. In my consulting experience, the most widely-used techniques have been the Jaccard similarity coefficient and single-link cluster analysis, both very simple techniques. They led to results which were startlingly obvious, (e.g. the division of a corpus of Aztec sculptures into two groups), but which the archaeologist had not noticed until it had been pointed out to him/her. 'Of course, why didn't I think of that?' is often the highest accolade that a statistical technique can earn.

There is still a great gulf between the sort of statistics that is presented at CAA and much of what is done (or not done) at the level of everyday archaeology. We have tended to rely on the 'trickle-down' model of the spread of ideas but, like the same model for the giving of aid to the Third World, this approach seems to have failed. Pursuing that analogy, one might suggest the development of a 'basic needs' or 'bottom-up' approach. But how can one stimulate a demand for a basic statistical expertise at the grass roots? The number of qualified statisticians in the archaeological profession will always be small, so perhaps we should be looking for a class of parastatisticians as argued by Roberts (1990) for the business community. As he suggests, they would not only carry out simple statistical tasks, but would be able to spot areas where professional expertise is required. They would be able to permeate quantitative thinking into their organisations from the level at which most of the real work is done. This implies a need for training — for basic numeracy for all and more specialised skills for those with the special role. Partly this is a job for the schools, but we cannot wait that long. We must do all we can, by both encouragement and criticism, to show that innumeracy is as unacceptable as illiteracy and that, as Colin Renfrew once said (an uncomfortably long time ago) 'the days of the innumerate are numbered'. We must take every opportunity we can find to show that numeracy is not a black art for the select few, but part of everyone's cultural heritage and a basic skill necessary for life and work (imagine the uproar if only 20% of archaeologists could read or write), and that quantitative methods are not only useful, but can even be fun!

References

- ALDENDERFER, M.S. 1987. "Assessing the impact of quantitative thinking on archaeological research", in M.S. Aldenderfer (ed.), *Quantitative Research in Archaeology*, Newbury Hill, California, Sage Publications: 9–29.
- ALVEY, B.A.P. 1983. "There is no stratigraphy: there is no proof", unpublished MA dissertation, University College London Institute of Archaeology, London.
- BARCLAY, K., M. BIDDLE & C.R. ORTON 1990. "The chronological and spatial distribution of the objects", in M. Biddle (ed.), *Object and Economy in medieval Winchester*, Oxford, Clarendon Press: 42–73.
- BLANKHOLM, H.P. 1991. *Intrasite Spatial Analysis in Theory and Practice*, Aarhus, Aarhus University Press.
- BUCK, C.E., W.G. CAVANAGH & C.D. LITTON 1988. "The spatial analysis of site phosphate data", S.P.Q. Rahtz (ed.), *Computer and Quantitative Methods in Archaeology 1988*, British Archaeological Reports (International Series) 446, Oxford, British Archaeological Reports: 151–160.
- BUCK, C.E. & C.D. LITTON 1989. "Image segmentation techniques for archaeological geochemical data", in S. Rahtz & J. Richards (eds.), *Computer Applications and Quantitative Methods in Archaeology 1989*, British Archaeological Reports (International Series) 548, Oxford, British Archaeological Reports: 121–131.
- BUCK, C.E. C.D. LITTON & A.F.M. SMITH forthcoming. "Calibration of radiocarbon results pertaining to related archaeological events", *Journal of Archaeological Science*.
- CHERRY, J.F., C. GAMBLE & S. SHENNAN 1978. *Sampling in contemporary British Archaeology*, British Archaeological Reports (British Series) 50, Oxford, British Archaeological Reports.
- DORAN, J.E. & F.R. HODSON 1975. *Mathematics and Computers in Archaeology*, Edinburgh, Edinburgh University Press.
- HODDER, I.R. & C.R. ORTON 1976. *Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press.
- HOLE, B. 1980. "Sampling in archaeology: a critique", *Annual Review of Anthropology*, 9: 217–234.
- JOHNSON, I. 1977. "Local density analysis", in *Computer Applications in Archaeology 1977*, Birmingham, University of Birmingham: 90–98.
- KROEBER, A.L. 1916. "Zui Potsherds", *Anthropological Papers of the American Museum of Natural History*, 18(1), New York: 1–37.
- LEBART, L., A. MORINEAU & K.M. WARWICK (eds.) 1984. *Multivariate Descriptive Statistical Analysis and related Techniques for Large Matrices*, New York, John Wiley.
- MUELLER, J.W. (ed.) 1975. *Sampling in Archaeology*, Tucson, Arizona, University of Arizona Press.
- ORTON, C.R. 1980. *Mathematics in Archaeology*, London, Collins.
- ORTON, C.R. 1987. "Review of Intrasite Spatial Analysis in Archaeology", *Bulletin of the Institute of Archaeology Review Supplement*, 24: 50–51.
- ORTON, C.R. & P.A. TYERS forthcoming. "Counting broken objects: the statistics of ceramic assemblages", British Academy.
- ORTON, C.R., P.A. TYERS & A.G. VINCE forthcoming. *Pottery in Archaeology*, Cambridge, Cambridge University Press.
- PETRIE, W.M.F. 1899. "Sequences in prehistoric remains", *Journal of the Royal Anthropological Institute*, 29: 295–301.
- RIPLEY, B.D. 1977. "Modelling spatial patterns", *Journal of the Royal Statistical Society Series B*, 39(2): 172–212.
- RIPLEY, B.D. & J-P. RASSON 1977. "Finding the edge of a Poisson forest", *Journal of Applied Probability*, 14: 483–491.
- ROBERTS, H.V. 1990. "Applications in business and economic statistics; some personal views", *Statistical Science*, 5(4): 372–390.
- SCHILTZ, M-A. 1990. *A French reanalysis of a British survey. Comparative study of statistical methods applied to social science data*, Paris, Centre d'analyse et de mathématique sociales.
- THOMAS, D.H. 1978. "The awful truth about statistics in archaeology", *American Antiquity*, 43(2): 231–244.
- WHALLON, R. 1984. "Unconstrained clustering for the analysis of spatial distributions in archaeology", in H.J. Hietala (ed.), *Intrasite Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press: 242–277.