Adding Value to GIS - Some Spatial-Analytical Techniques and Their Applications

Clive R. Orton

University College London Institute of Archaeology

Introduction

Geographical Information Systems are excellent vehicles for storing, manipulating and displaying spatial data, but frequently lack analytical capability, or are designed to answer the sorts of questions that archaeologists do not commonly ask. To make the best use of them, therefore, archaeologists may have to use additional software that is capable of answering more relevant questions, through a range of techniques grouped together under the heading of *spatial analysis* (Hodder and Orton 1976). The aim of this paper is to present some techniques of spatial analysis, made possible by recent developments in computing power, which may be of value to archaeologists, and to demonstrate their use in case studies.

Types of spatial data

Spatial data can be divided into four types (Bailey and Gatrell 1995, 11-18):

- *point-pattern data*, in which the point locations of objects are of prime interest, although there may well be additional data about attributes of the objects (e.g. their type),
- *spatially continuous* or *geostatistical data*, for which point locations are chosen, and the values of certain variables are observed or measured at those locations,
- *area data*, data which are only available for areas,
- *spatial interaction data*, data on flows which link a set of locations (areas or points).

O'Sullivan and Unwin (2003, 5-7) call the first two types the *object view* and the *field view* respectively, linking them to vector and raster representations of space. This paper is concerned only with these two types of data.

Spatial patterns derive from the operation of spatial processes, and can be seen as the result of two sorts of variation in the process - global or large-scale trends (*first-order effects*) and local or small-scale (*second-order*) effects (Bailey and Gatrell 1995, 32). The latter result from spatial dependency in the process, i.e. from a tendency for values of the process at nearby locations to be correlated with each other. In archaeological terms, this effect might show itself in the form of clusters of sites or artefacts. Many spatial patterns are the result of a mixture of these two effects.

Second-order processes can be divided into *homogeneous* (or *stationary*) processes and heterogeneous (or *non-stationary*) processes. A spatial process is called homogeneous if its statistical properties are independent of absolute location, i.e. if its mean and variance do not vary according to location. Full homogeneity further implies that the covariance between values at two locations depends only on the distance between them (Bailey and Gatrell 1995, 33). Techniques of spatial analysis are usually devised to explore first-order effects and second-order variations in the mean of a process, under the assumption of homogeneity of the variance and covariance.

A further important point is that the nature of spatial patterning can depend on the scale at which it is examined. Since a spatial pattern can demonstrate completely different characteristics at different scales, any characterisation of a pattern must make it clear at which scale it has been observed. Modern techniques of analysis exploit this property by seeking the scales at which certain characteristics are most pronounced.

Point-pattern analysis

A brief history

Early applications of point-pattern analysis were made in the field of ecology, and were based on either a *quadrat* approach (Greig-Smith 1964) or a *nearest neighbour* approach (Clark and Evans 1954). Pioneering attempts to apply such techniques in archaeology (for example, Dacey 1973; Whallon 1973; 1974) were reviewed in a general study of spatial analysis in archaeology (Hodder and Orton 1976). This also included ideas and techniques 'borrowed' from geography and applied at regional level, in contrast to the ecological models which tended to be applied at intra-site level. There followed a series of innovations (e.g. Johnson 1977; Berry *et al* 1984; Whallon 1984; Barceló 1988; Blankholm 1991). A slightly more detailed account can be found in Orton (2005).

Problems of conventional approaches

A main problem is that of 'edge effects'; some analytical techniques are based on the assumption of a theoretically infinite study area, while all surveys or excavations have boundaries or 'edges'. The need to modify techniques to allow for such effects has been a major theme in spatial analysis.

The quality and internal consistency of any particular dataset will not only affect the choice of technique, but will also determine the suitability of the dataset for spatial analysis at all. In practice, this means that spatial analytical techniques are best suited to small discrete datasets, preferable collected by a single individual or organisation over a relatively short period of time. This argument favours the use of intra-site spatial analysis over inter-site or regional spatial analysis, although archaeologists study patterns over a wide range of scales.

Developments in statistical theory

While these developments were taking place in archaeology, there were parallel developments in statistical theory, associated in Britain with Ripley (1976; 1977; 1981) and Diggle (1983; 2003). They were based on the idea of modelling the stochastic processes that produce spatial patterns, and introduced the *K function* as a tool for characterising spatial patterns. The *K function* is defined by

$\lambda\ K(h) = E(\#(\text{events within distance } h \text{ of an arbitrary event}))$,

where # means 'the number of', $E(\ )$ denotes an expectation, and λ is the *intensity* or mean number of events per unit area (assumed constant) (Bailey and Gatrell 1995, 92). A related function, the *L function*, was found to be a useful indicator of clustering at particular scales (Besag 1977). It was defined by

$\hat{L}(h) = \sqrt{\hat{K}(h)/\pi} - h$ (Bailey and Gatrell 1995, 94).

Initially, the *K* function was defined for the distribution of a single type, but later the bivariate function, the *cross K function*, was defined by

$\lambda_j\ K_{ij}(h) = E(\#(\text{type } j \text{ events} \leq h \text{ of an arbitrary type } i \text{ event}))$,

with the analogous cross *L function*

$\hat{L}_{ij}(h) = \sqrt{\hat{K}_{ij}(h)/\pi} - h$ (Bailey and Gatrell 1995, 120-1).

These functions are surprisingly similar to *local density analysis* (Johnson 1977), but their exploratory use was based on a plot of *K* or *L* against *h*, looking for peaks (indicating clustering at scale *h*) or troughs (indicating regularity at scale *h*) in the *L* function. Edge effects were recognised and accommodated into the theory, and it became possible to calculate confidence zones for *K* and *L*, so that the significance of any observed clustering or regularity could be assessed (Besag and Diggle 1977).

A natural development is to relax the conditions under which techniques such as *K* and *L* functions can be used. Their use is based on the assumption of a homogeneous and isotropic point process; there may well be practical reasons why this assumption does not hold. Pélissier and Goreaud (2001) suggested a three-stage approach to such problems:

- Detection of possible heterogeneity through the observation of a peak in the *L* function at large scales,
- Division of the study area into homogeneous sub-regions,

- Separate analysis of each sub-region.

They also demonstrated a useful tool (proposed by Getis and Franklin 1987) of mapping the values of $L_i(h)$ across a study area to show local variation at a range of scales. The case studies of their applications are entirely ecological, concerned with forestry.

ADE-4 (ADS) - what it is and what it does

The ADE-4 (*Ecological Data Analysis*) package is a set of tools for exploratory data analysis, available from the University of Lyon, France (pbil.univ-lyon1.fr/ADE-4/ ADE-4.html) which can run on either Windows or MacOS. Most modules are bilingual (French and English). Its ADS (*Spatial Data Analysis*) module contains three programs: *Ripley* (for univariate analysis), *Intertype* (for bivariate analysis) and *ADSutil* (for data manipulation). Various plotting routines within the package, such as *Curves* and *Plot*, are used to display the output. *Ripley* calculates K and L functions and the data needed to map the L function across the study area. *Intertype* calculates cross K and cross L functions, and enables the cross L function to be mapped. Edge effects are dealt with according to Goreaud and Pélissier (1999), and the general approach is as described by Pélissier and Goreaud (2001). The programs are well documented, with worked examples.

Case study

The case study is Barmose I, an early Maglemosian (mesolithic) site dated *c.* 7500-6000 b.c. and located in Barmosen (Barmose Bog) in South Zealand, Denmark (Johansson 1971; 1990). The excavated area was almost 100 sq. m., in an irregular shape. There is evidence for the presence of a hut floor with a single internal hearth, but its outline can only be approximated (Blankholm 1991, 185). All artefacts, tools and ecofacts were recorded in three dimensions to the nearest centimetre, from a 'culture layer' up to 5 cm thick. The dataset consists of the location and class of each of 470 flint artefacts that had been plotted exactly; the numbers in each class are shown below. The data were stored as a tab-separated text file with three columns: *x*-coordinate, *y*-coordinate, class type. The class codes and counts of artefacts are as follows:

Class code	*class*	*abbreviation*	*count*
1	scraper	SCR	38
2	burin	BUR	25
3	lanceolate microlith	LAN	36
4	microburin	MIC	16
5	flake axe	FLA	28

6	core axe	CAX	4
7	square knife	SQK	192
8	blade/flake knife	KNI	16
9	denticulate/notched	DEN	26
10	core	COR	80
11	core platform	CPL	9

The dataset has been analysed using the techniques of K-means Cluster Analysis, Unconstrained Clustering, Correspondence Analysis and Presab (Blankholm 1991, 183-205).

The new analyses (for a detailed account of which see Orton 2005) comprised:

- a *K* function and an *L* function for each class
- a map of the *L* function for each class
- a cross *K* function and a cross *L* function for each class with each other class
- a map of the cross *L* function for each class with each other class.

The functions were plotted at 0.1m intervals from 0.1m to 3.0m, and maps of the *L* functions and cross *L* functions were produced at the same interval. This gave a wealth of graphical output to be interpreted. Examples of a *K* function, an *L* function and a cross *L* function map are shown as Figs 1-2.

The horizontal interval is 0.1m; curve 1 represents the data and curves 2 and 3 delimit a confidence zone for a uniform distribution. Evidence of aggregation can be seen between 0.6 and 0.9m.

The first step in interpretation was to look at the *K* and *L* functions. The values of the functions for $h = 0.1$m were ignored, as they could be unduly influenced by rounding in the recording of locations. The functions revealed strong aggregation at large scales for all classes, a clear indication of spatial inhomogeneity in the data, and an indication that the space should be divided for finer-grained analysis.

To proceed further, separate analyses of the core and peripheral areas are needed. The boundary between the core area (*densezone*) and the periphery (*outerzone*) was determined by visual inspection of the maps of the *L* functions, together with plots of artefact classes. More sophisticated approached to this division (Ripley and Rasson 1977) were felt to be unnecessarily complicated here. The Ripley analyses of the densezone (carried out at up to $h = 1.5$m) showed aggregation for burins, flake axes, square knives, cores and possible scrapers. In each case maximum aggregation seems to occur around $h = 0.6$ to 0.8m. There are no instances of segregation (uniformity), except possibly for scrapers at $h > 1.0$m. The maps of the *L* functions show that the classes which show aggregation do so in different parts of the densezone; scrapers in the north, south-west and south-east, burins in the south-centre, flake axes in the west, square knives in the north-centre and cores in the east.

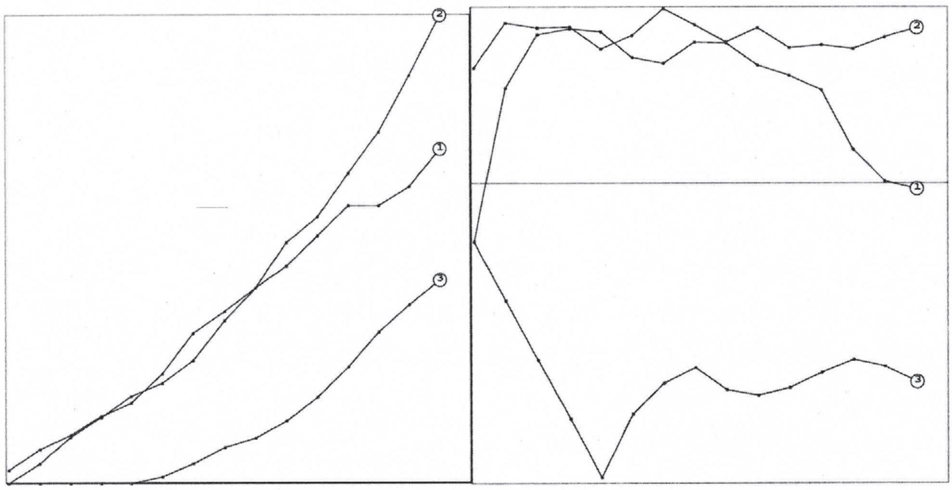

Fig. 1 (left) Example of a *K* function: burins in the densezone. (right) The corresponding L function.

The Ripley and Intertype analyses of the densezone are summarised in Table 1. Aggregation between two classes is rare, occurring only between burins and microliths, and possibly between microliths and denticulates (in both cases at scales above 1.3m). Flake knives are the most consistently segregated class (showing segregation from burins, square knives, denticulates and cores at various scales), followed by scrapers which show segregation from microliths, square knives and cores. In contrast, blade knives show no significant relationships with other classes (which may just reflect how few they are).

The corresponding analyses for the outerzone are not presented here, but can be found in Orton (2005).

The outcomes were compared to Blankholm's (1991, 203) results, showing both similarities and differences (Orton 2005). A more general insight is that, although some clusters of different types do overlap, the level of aggregation between different types across the whole site is low. This suggests that these overlaps may be due to repeated use of the same space for different functions, rather than the association of the types together in the same function.

Compared to the other techniques available for point pattern analysis, ADS has both advantages and disadvantages. It is good at examining variation across a range of scales, and produces rich graphical output for interpretation. It copes well with edge effects, and does not rely on data-smoothing, which tends to create spurious patterns.

The quantity and variety of the graphical output of ADE-4 makes it very suitable for an 'interactive' approach, in which specialist questions are posed and answered, the output giving rise to fresh questions. It may be less suited to providing a single definitive 'result', e.g. in terms of definitive zoning of the site.

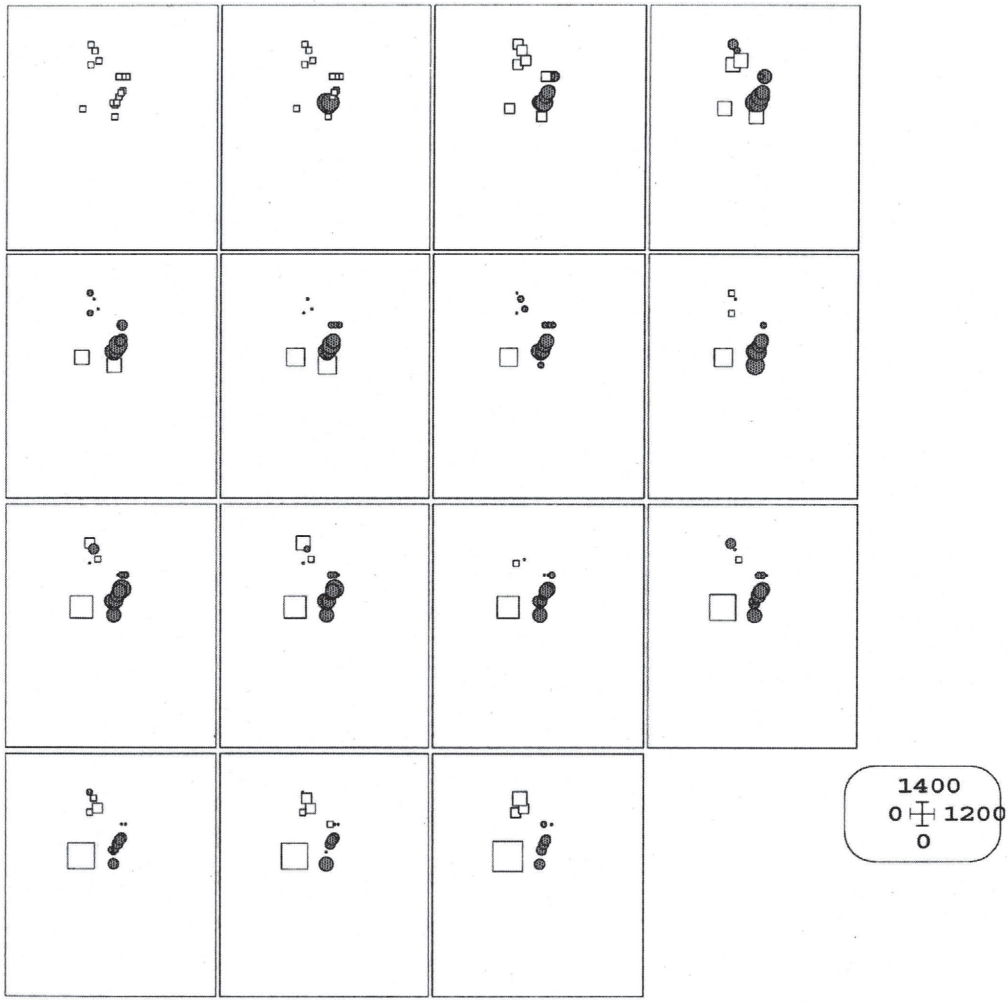

Fig. 2 The *L* function map corresponding to fig. 1. Each symbol represents the location of a burin. Hollow squares indicate 'low' *L* values (i.e. segregation), and shaded circles represent 'high' *L* values (i.e. aggregation). The size of the symbol reflects the strength of the pattern. The 'thumbnail' maps are at 0.1m intervals of h, from $h = 0.1$m (top left) to $h = 1.5$m (bottom right). Strong aggregation can be seen between $h = 0.6$ and $h = 0.9$m, but even here some burins (to the west) are segregated.

Its main drawback is its implicit reliance on a hypothesis-testing paradigm, which is apparent in the confidence zones for the *L* and cross *L* functions, and which forms the basis of the *L* function maps. As always, such an approach is much influenced by sample size, and the significance of a pattern can reflect the number of artefacts in a particular class as much as the nature of the pattern itself.

Spatially continuous (geostatistical) data

History

		Local class							
Base class		SCR	BUR	LAN	FLA	SQK	KNI	DEN	COR
	SCR	A? 0.8m S? >1.0m	ns	S 0.4- 0.8m	ns	S all	ns	ns	S 0.5m
	BUR	ns	A 0.6-0.9m	ns	S 0.5-0.8m	ns	ns	ns	ns
	LAN	ns	A >1.3m	ns	ns	ns	ns	A? >1.3m	ns
	FLA	ns	S 0.5-1.1m	ns	A 0.3-1.0m	S 0.1, 0.4m >1.4m	ns	S 0.5m	A 0.4m
	SQK	S 0.2m >1.3m?	ns	ns	S 0.2m >1.4m	A 0.2-1.3m	ns	ns	ns
	KNI	ns	ns	ns	ns	ns	ns	ns	ns
	DEN	ns	ns	ns	S 0.4, 0.8m	ns	ns	ns	ns
	COR	S 0.4, 0.6m	ns	ns	S 0.3m	ns	ns	ns	A 0.7m

Table 1 Relationships between tool classes in the densezone, as expressed by the *K* and *L* functions. The diagonal elements refer to the Ripley analyses.
A = aggregated, S = segregated (the numbers show the scales at which these occur), ns = no significant relationship.

As the name suggests, techniques for analysing this type of data have been developed mainly in the geological and earth sciences (Howarth 1983), and have only recently come to archaeology. An exception must be made for geophysical survey techniques in archaeology, where specialised equipment for resistivity and magnetometry has long been supported by dedicated software (e.g. *Geoplot*, www.geoscan-research.co. uk/page9.html). However, this has made no impact outside its own specialised field and ideas about treating archaeological data as explicitly geostatistical have only recently been put forward (Ebert 2002; Wheatley and Gillings 2002), although there have been attempts to use techniques designed for this sort of data, in an implicit fashion (see below).

One area where an explicitly statistical approach has been used in archaeology is in predictive modelling of the locations of sites (Kvamme 1990; Kamermans, this conference). The purpose of predictive modelling is to predict the probability of the occurrence of an archaeological 'event' (e.g. a site) in a spatial unit (e.g. one square of a grid). It is thus an 'area data' type of technique. The approach used is to construct a relationship between the existence of 'events' and the values of chosen independent (usually environmental) variables, through a statistical technique such as *discriminant analysis* or *logistic regression* (Kvamme 1990, 275-276), based on data from a subset (*training set*) of the spatial units. This function can be used to predict occurrences in other units of the study area.

Problems

Problems arose in early experiments (see Hodder and Orton 1976, 164-174), because either the techniques were too strongly model-based, or they were trying to perform unhelpful functions on the data. An example of the former is the attempt to fit trend surfaces to the distribution of a class of Roman pottery in southern England, which gave very strange results if extrapolated beyond the confines of the data. Smoothing is an example of the latter: archaeological data (for example, artefact distribution patterns) are frequently discontinuous, and there is a temptation to smooth them to make them more continuous, and so more amenable to geostatistical techniques. This seems fundamentally misplaced - archaeological data are often already smoothed (e.g. by site formation processes) and need rather to be 'sharpened', than smoothed further (Whallon 1984).

Variograms and kriging

The second-order patterns in spatially continuous data are often at least as interesting as any first-order patterns. We used *K functions* to study the second-order properties of point pattern distributions; the corresponding statistics for spatially continuous data are known as *covariance functions* or *covariograms*. The covariance function $C(h)$ is defined as the covariance between the values of a spatial variable at locations that are at a distance h apart. This definition only makes sense if the variable is both *stationary*, i.e. if its mean and variance are constant across the whole study area, and *isotropic*, i.e. if $C(h)$ depends on the value of h and not on its direction. If the variance of the variable is σ^2, then the corresponding correlation, $\rho(h) = C(h)/\sigma^2$ is known as the *correlogram*. Finally, the *variogram* is derived from the correlogram by the formula $\gamma(h) = \sigma^2 - z(h)$. All three functions give much the same information, but the variogram is the one usually employed (Bailey and Gatrell 1995, 161-166). A variogram is characterised by four parameters: its *model* (for example, spherical, exponential, power, or gaussian, or a combination of them, see Bailey and Gatrell 1995, 179), its *range* (the value of h at which the curve 'levels off'), its *sill* (the value of $\gamma(h)$ at which the curve levels off, or towards which it converges), and its *nugget* (the value of $\gamma(h)$ at $h = 0$; non-zero values indicate discontinuity in the data).

The main purpose of *kriging* is to interpolate values of variables at points that were not sampled. The name derives from the South African mining geologist, D.G. Krige, who developed an early version of the approach. The simplest estimate of the value of a spatial process at a chosen point would be its overall mean, or a value based on an observed trend. Kriging uses knowledge of the autocorrelations in a dataset to improve on this, by adding an estimate of the second-order effects, based on the values observed at the data points (Bailey and Gatrell 1995, 183-199). The addition is a weighted linear combination of the residuals at the data points, and carries an estimate of the likely error (the *mean square prediction error*).

For computational reasons, it is often the practice to use only data points within a *search neighbourhood* of the point where a predicted value is required. If a constant mean value is assumed, the technique is known as *ordinary kriging*; if a trend is assumed, the term *universal kriging* is used. If an estimate is required, not at a point location, but the average value over a small area or *block*, the technique of *block kriging* can be used (Bailey and Gatrell 1995, 208, 210-212). Although its use was suggested as early as the 1970s (Zubrow and Harbaugh 1978), kriging has been little used in archaeology. Interest has recently been revived by Ebert (2002).

Case studies

The first case study is a simple illustration of the use of variograms as a descriptive tool. The data are raw phosphate readings (mgP/100g) from site LS165 of the Laconia Survey in Greece (Buck et al 1988). Readings were taken at 10m intervals on a 16 by 16 grid; a few observations are missing. The data file is in the standard Geo-EAS format (see www.ai-geostats.org/software/Geostats_software/geoeas.htm), and the software used is *Variowin* 2.21 (available from www-sst.unil.ch/research/variowin/; see Pannatier 1996). Its modules include:

- *Prevar2D.exe,* which computes a pair comparison (.pcf) file from a Geo-EAS data file,
- *Vario2DP.exe*, which carries out exploratory variography in 2D using a .pcf file,
- *Model.exe*, which models experimental variograms originating from vario2DP.exe, and includes interactive modelling of geometric and zonal isotropies.

Use of Vario2DP shows that there is considerable autocorrelation, but that it is not isotropic, i.e. it varies according to the direction. The variogram for the north-to-east sector,

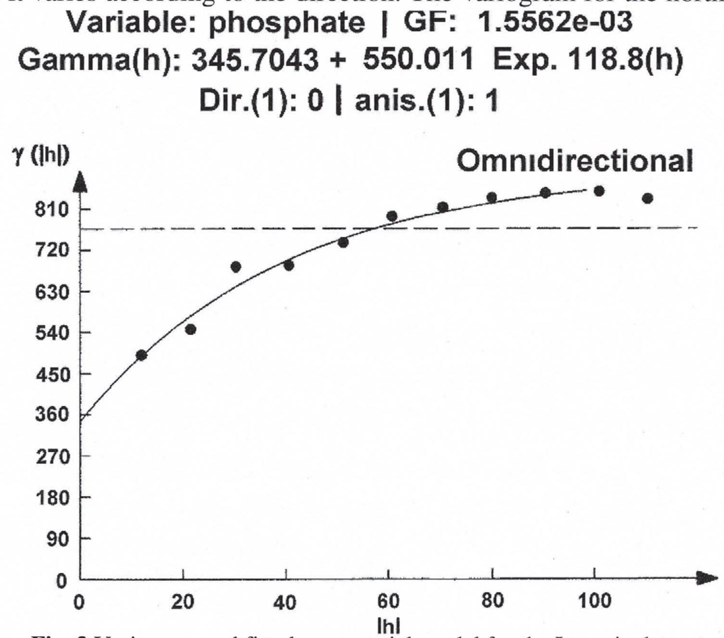

Fig. 3 Variogram and fitted exponential model for the Laconia dataset.

for example, increases steadily throughout, levelling off at about 80m, while that for the east-to-south sector peaks at about 60m and then declines. The Model module is used to empirically fit various models to directional variograms. In the case of an omnidirectional model, a good fit was achieved by an exponential model with a nugget of about 350, a sill of 550 and a range of 120m (Fig. 3). Interpretation of such models is difficult; here the high relative nugget value (about 40%) suggests fairly low short-range autocorrelation (indicating random noise, or possible discontinuities), but the high range suggests low autocorrelation effects continuing over considerable distances.

Perhaps the most useful role of a variogram is as an input into kriging, which can be used predictively. Gjesfjeld (2004) attempted to exploit this feature in a study of alternative routes for a by-pass at Hortonville, Wisconsin, USA. The variogram (produced by Variowin 2.21) was fairly flat, with a high relative nugget effect (about 70%), indicating only weak spatial autocorrelation. The best fit was obtained from an exponential model, and the data appeared to be isotropic. Gjesfjeld found that the data on site locations were too sparse for ordinary kriging, and that block kriging was preferable. It was carried out using the Geostatistical Analyst function of ArcGIS. The outcome was a prediction map (Fig. 4), showing the predicted densities of sites along the two suggested routes for the by-pass. On the basis of this map, he expressed a strong preference for the northern route, as being likely to impact on fewer archaeological sites than the southern route.

Discussion

It must be recognised that most archaeological spatial data are of the point-pattern type, at a wide range of scales, and that genuinely continuous data area relatively rare and often handled by highly specialised software. Both types can be downgraded to area data, with a resulting loss of information, and intrinsically point-pattern data may in fact be collected in this way (e.g. as counts of artefacts in grid squares). Ideally, analytical techniques should be chosen to match the type of data, but this does not appear to always be the case, with a

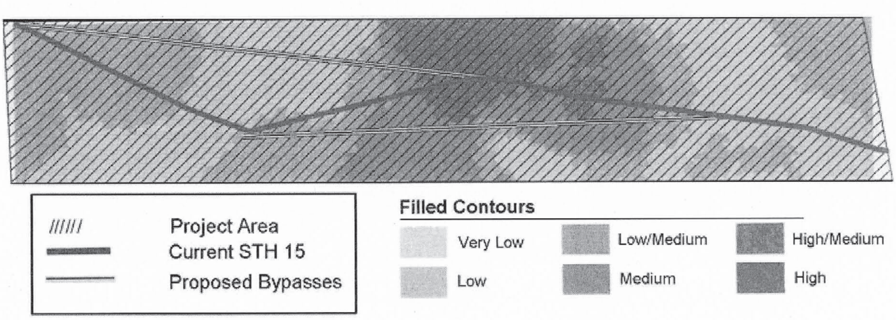

Fig. 4 Prediction map for route of STH15 and proposed by-passes (Gjesfjeld 2004, Fig. 22). The shading shows the probability of the existence of archaeological sites.

tendency to treat point-pattern data as if it were continuous. It could be argued that an observed point pattern is only one possible outcome of a spatial process, and that we should be interested in studying the process rather than any particular outcome. In other words, we should estimate underlying probability distributions (e.g. through *kernel density analysis*, see Barceló 2002, 244) rather than focus on the detail of any particular dataset. Not all archaeologists would agree with this, so if this approach is followed, the reasoning should be made explicit to permit its discussion. Even if we do follow this view, we have to admit that underlying distributions are likely to contain (in their original state) discontinuities, which may well have become blurred through site formation and other processes. Techniques for sharpening data, such as *change-point analysis* (Buck *et al.* 1988; 1996, 258-276), may therefore be more relevant than those which assume a continuous underlying distribution.

It is therefore clear that archaeologists need to consider very carefully the exact mathematical nature of their data, and of the questions that they are asking of them, before choosing an appropriate technique of spatial analysis. The relative merits of competing techniques (e.g. logistic regression and kriging) need to be assessed under a wide range of circumstances, to provide guidance for professional archaeologists who may not have the time, resources or inclination to tackle every spatial problem from 'square one' as it arises. This is however not a plea for 'rules of thumb' (which have bedevilled contract archaeology over the last decade), but for carefully thought out decision procedures.

Acknowledgements

I am grateful to students in *Spatial Analysis in Archaeology* classes for carrying out the bulk of the point-pattern analyses described here, and especially to Matthew Reynolds for his particularly perceptive comments. The discussion on kriging owes much to work undertaken by Erik Gjesfjeld in his Report for the MA in Field and Analytical Techniques in Archaeology (Gjesfjeld 2004), to whom I am especially grateful for permission to use Fig. 4.

References

Bailey, T.C. and Gatrell, A.C. 1995, *Interactive Spatial Data Analysis*, Harlow, Longmans Scientific & Technical

Barceló, J. A. 2002, Archaeological thinking: between space and time, *Archeologia e Calcolatori* 13, 237–257.

Berry, K.J., Kvamme, K.L. and Mielke, P.W. 1984, Efficient permutation procedures for the analysis of artefact distributions, in H. Hietala (ed.), 1984 *Intrasite Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press, 54-74.

Besag, J. 1977, Contribution to the discussion of Dr Ripley's paper, *Journal of the Royal Statistical Society Series B* 39, 193-195.

Besag, J. and Diggle, P.J. 1977, Simple Monte Carlo tests for spatial patterns, *Applied Statistics* 26, 327-333.

Blankholm, H.P. 1991, *Intrasite spatial analysis in theory and practice*, Aarhus, Aarhus University Press.

Buck, C.E., Cavanagh, W.G., and Litton, C.D. 1988, The spatial analysis of site phosphate data, in S.P.Q. Rahtz (ed.) *Computer and Quantitative Methods in Archaeology 1988*. BAR Int Ser 446, 151-160.

Buck, C.E., Cavanagh, W.G. and Litton, C.D. 1996, *Bayesian Approach to Interpreting Archaeological Data*. Chichester, John Wiley & Sons.

Clark, P.J. and Evans, F.C. 1954, Distance to nearest neighbour as a measure of spatial relationships in populations, *Ecology* 35, 445-453.

Dacey, M.F. 1973, Statistical tests of spatial association in the location of tool types, *American Antiquity* 38, 320-328.

Diggle, P.J. 1983, *Statistical Analysis of Spatial Point Patterns*, London, Academic Press. [2nd edition 2003]

Djindjian, F. 1988, Improvements in intra-site spatial analysis techniques, in S.P.Q. Rahtz (ed.), *Computer and Quantitative Methods in Archaeology 1988*, BAR International Series 446(i), Oxford, 95-106.

Ebert, D. 2002, The potential of geostatistics in the analysis of fieldwalking data, in D. Wheatley, G. Earl and S. Poppy (eds), *Contemporary Themes in Archaeological Computing*, University of Southampton Archaeology Department Monograph 3, Oxford, Oxbow Books, 82-89.

Getis, A. and Franklin J. 1987, Second order neighborhood analysis of mapped point patterns, *Ecology* 68, 473-477.

Gjesfjeld, E.W. 2004, *Spatial Prediction and Contract Archaeology: the Benefits of Geostatistics*, unpublished Master's Report, UCL Institute of Archaeology.

Goreaud, F. and Pélissier, R. 1999, On explicit formulas of edge effect correction for Ripley's K-function, *Journal of Vegetation Science* 10, 433-438.

Grieg-Smith, P. 1964, *Quantitative plant ecology*, London, Methuen.

Hietala, H. (ed.) 1984, *Intrasite Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press.

Hodder, I. and Orton, C. 1976, *Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press.

Howarth, R.J. (ed.) 1983, *Statistics and Data Analysis in Geochemical Prospecting*, Amsterdam, Elsevier Scientific Publications.

Johansson, A.D. 1971, Barmose-gruppen. Præboreale bopladsfund med skiveøkser i Sydsjælland. Foreløbig meddelelse, *Historisk Samfund for Præstø Amt. Arbog 1968* 101-170.

Johansson, A.D. 1990, *Barmose-Gruppen. Præboreale bopladsfund i Sydsjælland*, Aarhus, Aarhus University Press.

Johnson, I. 1977, Local density analysis, in S. Laflin (ed.), *Computer Applications in Archaeology 1977*, Birmingham, University of Birmingham Computer Centre, 90-98

Kvamme, K.L. 1990, The fundamental principles and practice of predictive archaeological modeling, in A. Voorrips (ed.) *Mathematics and Information Science in Archaeology: a flexible framework*, Bonn, Holos, 257-295.

Orton, C.R. 2005, Point pattern analysis revisited, *Archaeologia e Calcolatori* 15, 000-000.

O' Sullivan, D. and Unwin, D.J. 2003, *Geographic Information Analysis*, Hoboken, John Wiley.

Pannatier, Y. 1996, *VARIOWIN: Software for Spatial Data Analysis in 2D* , New York, Springer-Verlag.

Pélissier, R. and Goreaud, F. 2001, A practical approach to studying the spatial structure in simple cases of heterogeneous vegetation stands, *Journal of Vegetation Science* 12, 99-108.

Ripley, B.D. 1976, The second-order analysis of stationary point patterns, *Journal of Applied Probability* 13, 255-266.

Ripley, B.D. 1977, Modelling spatial patterns, *Journal of the Royal Statistical Society, Series B* 39, 172-212.

Ripley, B.D. 1981, *Spatial Statistics*, Chichester, John Wiley.

Ripley, B.D. and Rasson J.-P. 1977, Finding the edge of a Poisson forest, *Journal of Applied Probability* 14, 482-491.

Whallon, R. 1973, Spatial analysis of occupation floors I: The application of dimensional analysis of variance, *American Antiquity* 38, 266-78.

Whallon, R. 1974, Spatial analysis of occupation floors II: The application of nearest neighbour analysis, *American Antiquity* 39, 16-34.

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, 242-277.

Wheatley, D. and Gillings, M. 2002, *Spatial Technology and Archaeology: the Archaeological Applications of GIS*, London, Taylor and Francis.

Zubrow, E.B.W. and Harbaugh, J. W. 1978, Kriging and simulation, in I. Hodder (ed.), *Simulation Studies in Archaeology*, Cambridge, Cambridge University Press, 109-122.

GISに付加価値を－空間分析の手法と適用－

クリーブ・R・オルトン

ロンドン大学 ユニバーシティ・カレッジ

この論文では、考古学者にとって有益な、最近開発されたGISのアドオンで利用する空間分析手法について検討する。その手法について、まず異なったタイプの空間データと、それを生成する可能性のある空間処理過程のタイプと特徴を記述することから始める。議論の焦点は、“ポイントパターン”データと“空間的に継続する面的パターン”の空間統計データの２タイプのデータについてである。

ポイントパターン分析は、生態学などの研究領域から考古学に導入された。このとき、同時にこの技術の持つ３種類の主要な問題事項、すなわち①空間スケールの問題、②エッジエフェクトの問題、③考古学的データの品質とデータ間の相互の整合性の問題、も移入された。最近の統計学的な手法の発展と開発により、①と②の問題は克服されたが、最後の③の問題は、常にそしてこれからも、我々が抱える問題である。ここでは事例研究として、どのようにして最新の技術が遺跡における異なった型式の遺物分布のパターンに対して、診断的な統計と視覚的表示を通じて明らかにすることができるのかを示す。統計パッケージソフトウェアADE-4のADSモジュールを利用する。

面的データの分析は、地質学や地球科学、環境科学などで進められてきた。例えば地球物理学調査のような特定の専門的な分野を除いては、考古学に対しては比較的小さな影響を与えるに止まった。ここでは現象の記述と予測のための空間統計学的手法について２つの事例研究を紹介する。最初は土壌のリン酸塩分のパターンをVariowinパッケージのバリオグラムを用いて読み取る特徴付けの方法について、次に、小規模な地理的範囲を対象として、遺跡分布を特徴づけるために、このVariowinを利用する。後者は、新しいバイパス道路の最適ルートの敷設に関して、遺跡存在予測地図を作製することを目的とした事例で、ブロッククリギング法（単純クリギング法・通常クリギング法とも）を適用する。

最後に、こうした関連の手法や適用による統計学と考古学の問題について概観し、これらの考古学研究への導入の可能性について提案する。