

Computer assisted pattern perception within post hole distributions

Mike Fletcher
and
Gary Lock

North Staffordshire Polytechnic,
Stafford.

1. INTRODUCTION

Archaeologists are increasingly using numerical techniques as tools to aid them in their interpretation of the evidence of past societies (Doran and Hodson, 1975, Hodson, Kendall and Tautu, 1971). This paper presents another such tool which may prove useful to archaeologists. The results presented here are not claimed to be archaeologically significant in any way other than the methodology used: archaeological analysis and conclusions based upon these results are subsequent steps which must be carried out by the archaeologist concerned.

The need for this tool was realised when trying to interpret the post hole distribution that is the result of the first 10 years of excavation at Danebury hillfort in Hampshire (Cunliffe, 1971 and 1976). We feel that it would be impossible for anyone viewing this distribution to be able to isolate and remember patterns within it. Human pattern perception is just not that good and is certainly not regular between more than one person, hence our programs were developed to search the post hole distribution for a pattern, in this case a circle, with pre-determined attributes. Danebury was eventually rejected as a test site following statistical analysis, to be discussed later, and for more readily apparent archaeological reasons. These are that the few circular houses known at Danebury from the stratified deposits, within the quarry hollows just inside the ramparts, are built in such a way that only the two door posts leave post holes of any considerable size, the walls being of wattle and daub construction. The stake holes left by these walls would not have been preserved over the non-stratified interior of the hillfort which only leaves pairs of door post holes to be found (Cunliffe, 1976 Fig. 9.). Of course there is no reason why other methods of house construction should not have been employed alongside the stake-wall method.

While searching the literature for examples of Iron Age round houses on which to base the figures to be inserted as the parameters of our programs it became obvious that it is impossible to generalise. What did seem apparent, however, is that known examples of round houses seem to divide into groups according to size. The first group includes larger houses such as at Crickley Hill (Dixon, 1976.) with a radius of 5.5 metres and 26 posts in the circle and the Pimperne house with 27 posts arranged in a circle of 5 metres radius (Harding and Blake, 1963). The second group includes smaller structures with less posts, many examples of which occur at Moel-y-Gaer (Guilbert, 1976) a hillfort in North Wales, with an average of 7 posts in a 3 metre radius circle. Although these are all called round houses there is no evidence to suggest that they were all lived in. Features other than the minimum common trait of a circle of post holes, such as a porch, are not included in the search parameters and must be looked for by eye in a subsequent stage of analysis. An important aspect of our programs, however, is that they will search for circular structures of any desired size, it is a trivial job to change the parameters.

The most important parameter is what we call the tolerance. In archaeological terms this represents the accuracy with which Iron Age house builders could, or at least thought it necessary to, position the post holes exactly in a circle. In our programs the tolerance is the width of the annulus

that searches for post holes, therefore the wider it is the more it will find. A tolerance of + 5 cms was used on our test runs which means that the centre of the post could be anywhere within a circle of 5 cms radius from the centre of the post hole. This allows for the posts not being in the centre of the post hole as it seems reasonable to position a post against one side of the hole for support and then pack the other side. It also allows for small errors on our part while transferring the co-ordinates from the site plan onto the machine. It is realised that this tolerance of 10 cms may be unrealistically small but it has statistical implications which will be discussed later. The second parameter is the range of radii which is to be searched. This along with the distance apart of the first two post holes in the circle from where the search starts, was decided by looking at known examples of round houses. We decided, for our test runs, to search from 3 to 8 metres radius with the first two post holes spaced from 1.75 to 2.25 metres apart.

The site used to test the search procedure on is Muntham Court, near Findon, Sussex, excavated about 25 years ago but not yet published. We would like to thank Professor Cunliffe of Oxford University for supplying the site plan. It was a settlement site in pre-Roman times and round houses are implied by circular surface features that were surveyed and planned before excavation. We have not yet seen this surface plan so it could provide a convenient test for our results. This site has 810 post holes distributed over an area of 2224m² giving a density of 0.364 postholes per m². Figure 1 shows a computer drawing of the post hole distribution over the excavated area.

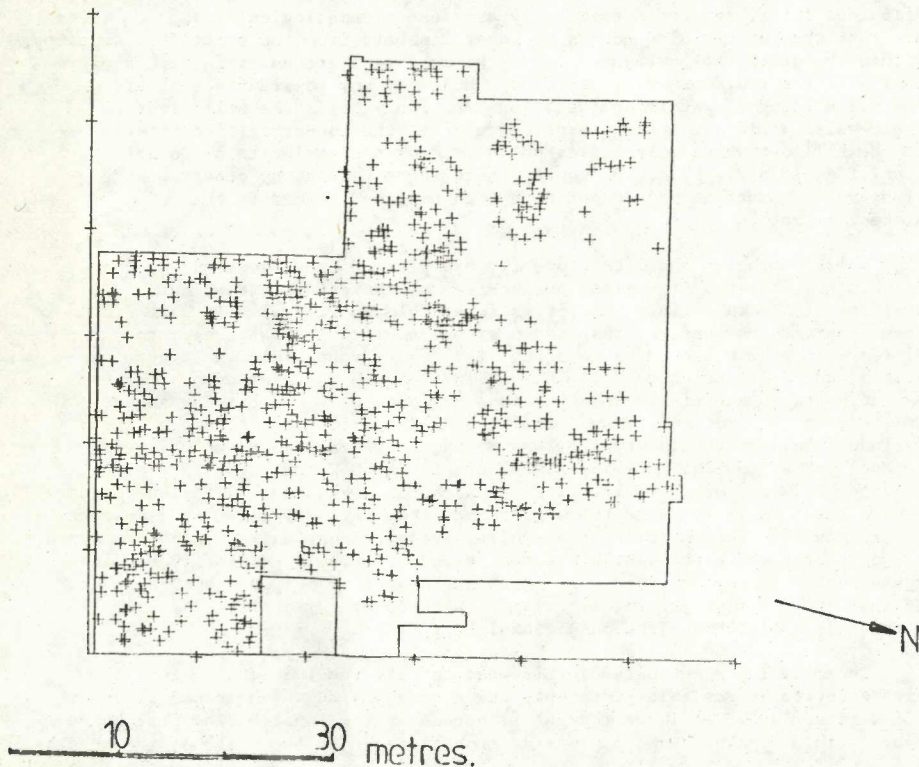


FIGURE 1

2. (i) THE BASIC TECHNIQUE

The problem is to identify possible or probable circles of post holes from a plan of a site. Consider N post holes over an area of A cm², giving a density of

$$\rho = \frac{N \cdot 100^2}{A} \text{ post holes/m}^2.$$

Clearly because of measurement errors in drawing the plan and reading from it, errors in placing a post in a post hole and errors the builder made, it is unrealistic to expect several post holes to lie exactly on a circle. Because of these errors we look for post holes which lie in an annulus of thickness D cm with average radius R cm as in Figure 2.

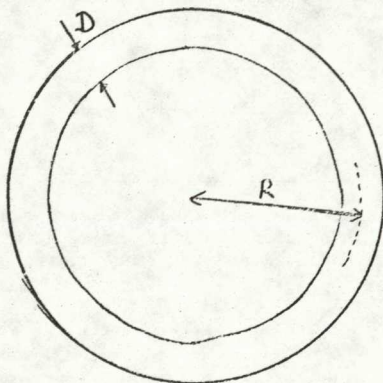


Figure 2.

In order to scan the whole area we would have to vary both the radius and the centre in such a way that all possible 'circles' were investigated, and since this would be very tedious, an alternative method is used. Two post holes a suitable distance apart (1.75 m to 2.25 m) are selected to start the search and then assuming that they lie on a circle the radius of this circle is varied from 3m to 8m in steps of D cm and the maximum number of post holes which are within $\pm \frac{D}{2}$ of any one of these circles is found, as in Fig. 3.

After this search the radius and co-ordinates of the centre of the annulus containing the most post holes is output. The whole search is then repeated with the circles on the other side of the initial pair of post holes, then a new pair of post holes is selected and the search repeated.

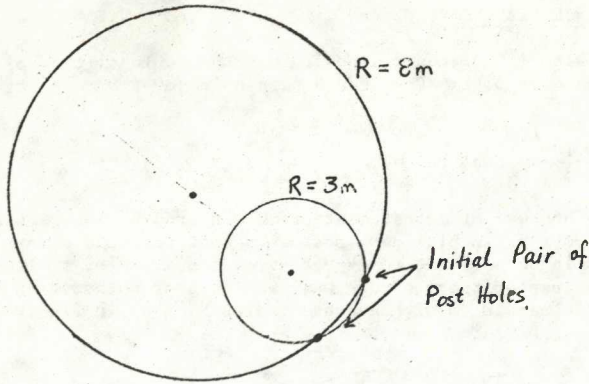


Figure 3.

(ii) Significance of the number of post holes in an annulus

Assume that the N post holes are randomly scattered over the area A so that every square metre has equal probability of containing a post hole. Then, for a fixed annulus of radius R and tolerance D , the probability of a particular post hole being in it is

$$p = \frac{2\pi RD}{A} \dots\dots\dots (1)$$

and the mean number of post holes in it is

$$\mu = \frac{2\mu RDN}{A} \dots\dots\dots (2)$$

Let the random variable X denote the number of post holes in the annulus. Since p is small and N large, X will approximately have a Poisson distribution and so the probability of having at most C post holes in the annulus will be

$$\text{prob}(X \leq C) = \sum_{x=0}^C \frac{e^{-\mu} \mu^x}{x!} \dots\dots\dots (3)$$

Since for any single search the radius R is varied from $3m$ to $8m$ in steps of D cm, the probability that the maximum number of post holes found in any annulus is at most C will be

$$\text{prob}(X_{\text{MAX}} \leq C) = \prod_{\text{All } \mu} \left\{ \sum_{x=0}^C \frac{e^{-\mu} \mu^x}{x!} \right\} \dots\dots\dots (4)$$

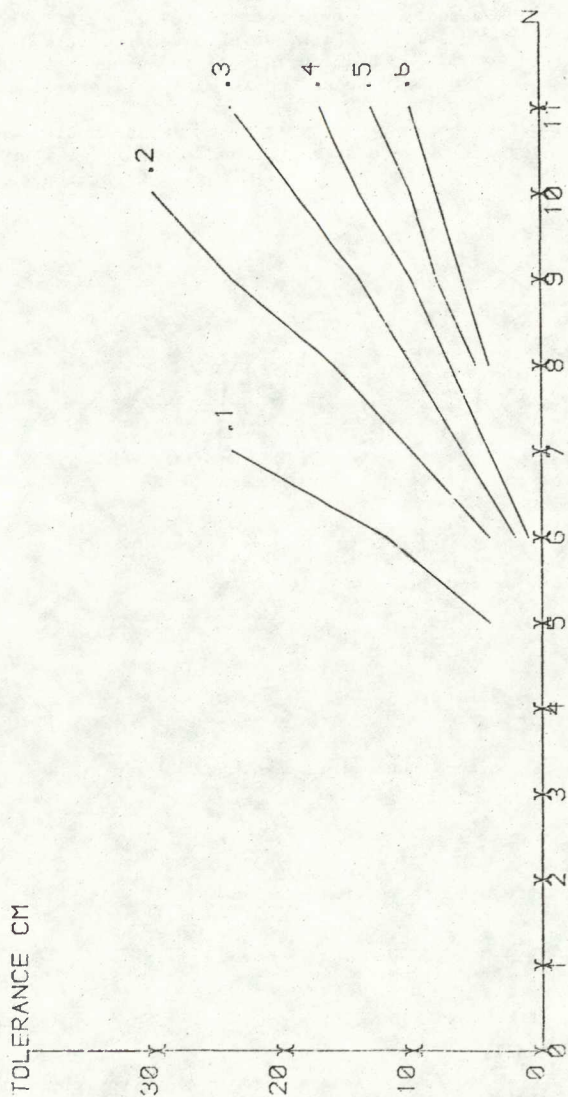


Figure 4.

This has been calculated for various values of the density ρ and the results are shown in Figure 4. Here D is the tolerance in cm and for each value of ρ the curve gives the value of n so that there is only a 5% chance of finding n or more post holes in an annulus, including the two initial post holes used to start the search.

These results allow a suitable choice of the tolerance D to be made. These graphs show that if the density is high, using a reasonable tolerance of 10 or 20 cm would yield too many too often, as was the case with Danebury, and a ridiculously low tolerance would be needed to make 9 a rare number.

As well as measuring the number of post holes in an annulus we decided to measure how evenly they were spread around the circle. For n points distributed around a circle there will be n angles θ_i at the centre, as in Figure 5.

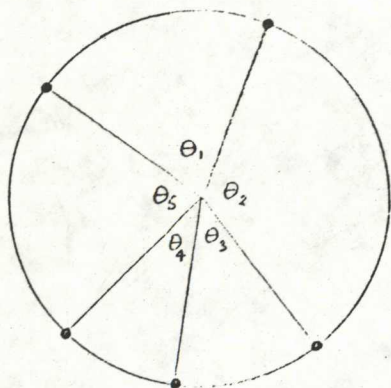


Figure 5.

One measure of how evenly the points are distributed in the variance of θ ,

$$s_{\theta}^2 = \frac{\sum \theta_i^2}{n} - \left(\frac{2\pi}{n}\right)^2.$$

However, since it can be shown that

$$E(s_{\theta}^2) = \sigma_{\theta}^2 = \frac{4\pi^2(n-1)}{n^2(n+1)} \dots\dots\dots (5)$$

we define a 'clumping coefficient' k by

$$k = \frac{s_{\theta}^2 \cdot n^2(n+1)}{4\pi^2(n-1)} \dots\dots\dots (6)$$

If the points are randomly distributed, the expected value of k is $+1$, with k distributed between 0 and $n+1$, with $k = 0$ implying that the points are equally spaced. This distribution of k has been investigated and the following empirical equation giving the value of k^* so that $\text{prob}(k \leq k^*) = 0.05$ obtained

$$k^* = 0.48 - \frac{1.20}{n} \dots\dots\dots (7)$$

The value of k can be used to choose which of the annuli found by the search method are of special interest by looking for those with low values of k (near k^*) indicating an even distribution, or those with a high value for k indicating perhaps part of a circular hut.

(iii) RESULTS

The programs, written in Fortran, were run on a GEC4082 with a plotter for graphical output from GINO routines. We chose pairs of post holes to initiate the searches which were between 1.75m and 2.25m apart, and we used a 5% sample of all such pairs (for our data it can be shown that we would expect 925 such pairs). We had 810 post holes in an area of 2,224 m² as shown in Figure 1, and Figure 6 shows two possible circles together with their clumping coefficients.

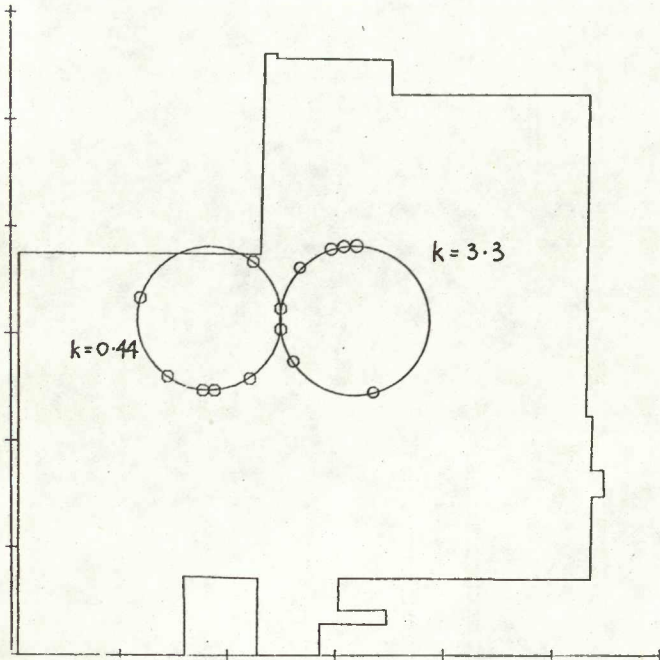


Figure 6.

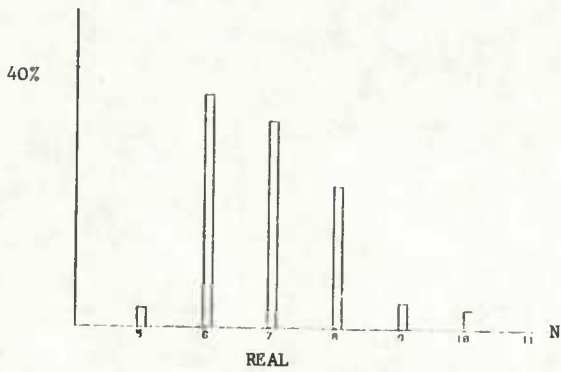
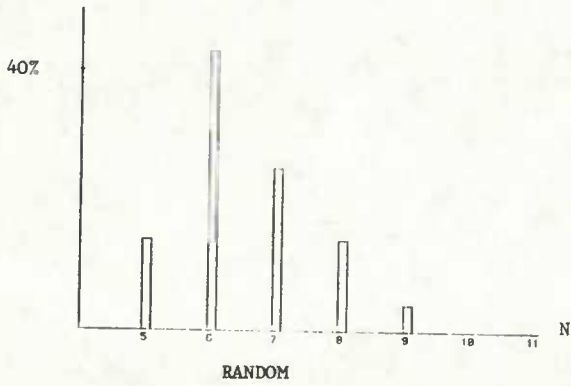
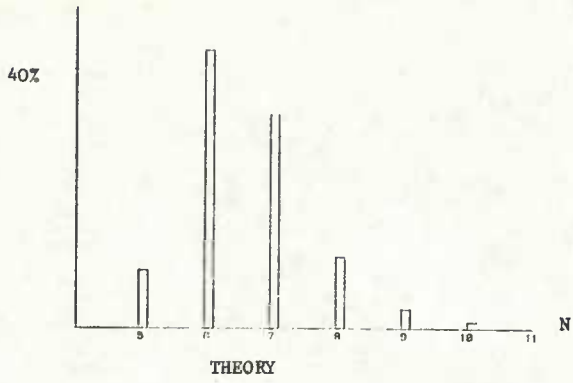


FIGURE 7

The whole procedure was repeated using 810 simulated post holes randomly distributed over the area.

Figure 7 compares the relative frequency of the maximum number of post holes found to fit an annulus for the real data, random data and the theoretical results predicted by equation (4).

It can be seen that the results for random data and those from the theory agree well, whilst those for the real data show a tendency towards more post holes fitting a circle than expected. This can be explained partly by the obvious clumping of the post hole distribution (see Figure 1).

3. ARCHAEOLOGICAL CONCLUSIONS

An important aspect of this technique, for the archaeologist, is that it can be used on two levels. The results at one level have been seen to give the statistical significance of any particular circle occurring. At a less sophisticated level the technique can be used purely as a circle finding tool using the clumping coefficient, tolerance and number of post holes to identify circles with specific attributes.

Future work will involve the complete run on the whole of the Muntham Court distribution using a bigger machine, an ICL 2960. This will output only those annuli of interest, perhaps those with 9 or more post holes and with either high or low values for the clumping coefficient, which can then be analysed and interpreted archaeologically. It is also hoped to change the search procedure to look for 4 and 6 post structures which could be more useful within the overall context of Iron Age studies than looking for circles.

Bibliography

- | | | |
|--------------------------------|-------|--|
| CUNLIFFE. B.W., | 1971. | 'Danebury, Hampshire: first interim report on the excavations, 1969-70'.
Antiquaries Journal. 51:
240-52. |
| CUNLIFFE. B.W., | 1976. | 'Danebury, Hampshire: second interim report on the excavations, 1971-5'.
Antiquaries Journal. 56.2:
198-216. |
| DIXON, P.W., | 1976 | 'Crickley Hill, 1969-1972', in Harding. D.W.
1976
161-75. |
| DORAN. J.E. and
HODSON. F.R | 1975 | 'Mathematics and computers in archaeology'.
Edinburgh. |
| GUILBERT. G.C. | 1976 | 'Moel-y-Gaer (R-hosesmor) 1972-73: an area excavation in the interior', in Harding. D.W. 1976
303-17. |
| HARDING. D.W. (ed) | 1976 | 'Hillforts: later prehistoric earthworks in Britain and Ireland'. |

- HARDING. D.W. and 1963 'An early iron age settlement in Dorset'.
BLAKE. I.M. Antiquity. 37. 63-4.
- HODSON. F.R., KENDALL. D.G. 'Mathematics in the archaeological and
and TAUTU. P. 1971 historical sciences'.
Edinburgh.

Appendix

Abstracts of papers presented at the conference for which no written version is available. For further details, please contact the authors.

- Esmee Webb "Model Building and Paradigm Testing - with a Typological Basis"
Institute of Archaeology, London.
It is suggested that a clarification of the nature of the basic typological material employed is necessary before high-order computer based methods of statistical analysis are imposed upon archaeological data, which are then manipulated in sophisticated models which in themselves provide a basis for extended theorising. A superstructure is only as secure as its foundations and these in archaeology are often very shaky.
- Ian Graham "Studies in Spatial Analysis"
Institute of Archaeology, London.
Many methods have been proposed for the analysis and comparison of 2-dimensional distributions of archaeological objects. This paper describes 3 such methods:
nearest neighbour analysis.
local density analysis.
spectral analysis.
The methods are illustrated by their application to the distribution of objects in graves in the cemetery at Hallstatt, Austria.
- John Wilcock "Microcomputer Graphics in Archaeology"
North Staffordshire Polytechnic
Computer graphics using light pens, cathode ray tubes and similar devices have proved too expensive in the past for much use by archaeologists.
Micro-computer low and high-resolution graphics are, however, proving much cheaper and these methods may be within archaeological budgets. The paper will demonstrate a number of applications of micro-computer graphics in archaeology and make suggestions for future development.