1

A Bayesian approach to a problem of archaeological site evaluation

Clive Orton

Institute of Archaeology, University College London

1.1 Background

The operation of field archaeology in England today is largely directed by two documents, neither of which has legal force, but which together form a strong framework. They are *Planning Policy Guidance Note 16: Archaeology and Planning* (Department of the Environment 1990), commonly known as PPG 16, and *Management of Archaeological Projects* (Andrews 1991), commonly known as MAP2.

PPG 16 gives policy guidance on the preservation of archaeological remains in the context of rural or urban development. The general philosophy is that much of the conflict of recent years about the preservation of archaeological remains could be avoided if steps were taken, *before* the granting of planning permission, to ascertain the nature and extent of any surviving remains on the site to be developed. It states that 'Where nationally important archaeological remains... are affected... there should be a presumption in favour of their physical preservation' (Department of the Environment 1990, para. 8), but that 'The case for the preservation of archaeological remains must however be assessed on the individual merits of each case,..., including the intrinsic importance of the remains' (Department of the Environment 1990, para. 27). Further, 'Where it is not feasible to preserve remains, an acceptable alternative may be to arrange prior excavation, during which the archaeological evidence is recorded.' (Department of the Environment 1990, para. 24).

Such decisions clearly require the prior knowledge of the survival (or not) of archaeological remains on a development site. This knowledge may be provided by assessment, defined as 'desk-based evaluation of existing information: it can make effective use of records of previous discoveries, ... (Department of the Environment 1990, para. 24), or by field evaluation, defined as 'a rapid and inexpensive operation, involving ground survey and small-scale trial trenching', which is 'quite distinct from full archaeological excavation'. The rationale is that 'Evaluations of this kind help to define the character and extent of the archaeological remains that exist in the area of a proposed development, and thus indicate the weight which ought to be attached to their preservation. They also provide information useful for identifying potential options for minimising or avoiding damage. On this basis, an informed and reasonable planning decision can be taken' (Department of the Environment 1990, para. 21).

A framework for the practical implementation of this policy is provided by MAP2, which breaks down an archaeological field project into five phases (project planning, fieldwork, assessment of potential for analysis, analysis and report preparation, dissemination). It defines the management principles and procedures for each phases, but does not itself offer practical guidance for their implementation. It would not be reasonable to expect it to set out more than general principles, given the wide diversity of archaeological fieldwork.

1.2 Discussion

These two documents, while setting out a quasi-legal framework (and it is important to remember that neither carries the force of law), left something of a vacuum in the practicalities of implementing the policies they put forward. For example, it was not clear how such sites should be selected for evaluation, nor what would constitute a satisfactory level of investigation for a site evaluation. Since one of the aims of PPG 16 is to prevent the sorts of problems that have in the past arisen from the discovery of significant archaeological remains while development is in progress, the criteria for satisfactory evaluation must be that the existence, extent and nature of archaeological remains are established with a degree of certainty that will enable rational decisions to be made about the design of the proposed development, but at a reasonable cost. Clearly, total excavation would provide the required information, but at prohibitive cost (possibly exceeding that of the proposed development in some cases). The problem is essentially one of sample design — choosing the 'best' combination of techniques (geophysical survey, boreholes, excavation, *etc.*) with their respective samples, where 'best' means providing the required information at minimum cost.

The initial, naïve, expectation of archaeologists was that there would be a simple rule, probably expressed in percentage terms, for the definition of a satisfactory sample of a site for purposes of evaluation. A considerable educational effort was made to wean them off this particular dummy; this seems to have been achieved. Archaeologists now seem aware of the difficulties, but as yet no coherent theory or methodology has emerged.

1.3 Aim

The aim of this paper is to look at one small part of this question, the existence (or not) of archaeological remains on a site. The broader questions of assessing their extent, historical significance, likely survival under various schemes of mitigation, *etc.*, will be left to others (*e.g.,* Biddle 1994; Carver 1993). Ours is not such a trivial question as it might at first sight seem to be. While in rural areas valuable evidence may be provided by survey techniques — aerial and geophysical survey, field-walking, *etc.*— these may not be available or useful in urban areas, for instance because of the presence of existing buildings and services, and restrictions on over-flying. For this reason, evaluation in urban areas can be expected to be more speculative than in rural areas, with a higher proportion of 'negative' outcomes (*i.e.,* evaluated sites on which no significant archaeological remains are found). This does appear to be the case: for example, in London in 1992, 134 out of 188 evaluations were negative, and in 1993, 160 out of 226 (71% in each year; McCracken & Phillpotts 1995, pp. 64–5).

It might be thought that this represents a waste of effort, and that zoning should be applied, with evaluations carried out only in areas of high archaeological potential. This would lead to a situation in which fieldwork merely confirmed what archaeologists thought they already knew, and since our knowledge is based on previous fieldwork (itself reflecting the density of archaeological activity in an area, see Hodder & Orton 1976, pp. 20–24), there would be a cycle of self-reinforcing bias. In practice, important discoveries have resulted from speculative evaluation in areas not known for their archaeological potential, *e.g.,* in London at Clapham (Densem & Seeley 1982) and Tulse Hill (Greenwood & Maloney 1995, p. 343).

This argument provides the rationale for this paper. A methodology is needed for the evaluation of sites which will probably contain no archaeological remains, but which may contain unexpected remains of unknown nature and importance. The approach must be cost-effective (to avoid suspicion of archaeologists keeping themselves in work by investigating 'dead' sites), but must carry an element of quality assurance, *i.e.,* a negative evaluation must really be negative, otherwise (i) valuable new information may be lost, or (ii) embarrassing discoveries may be made in the course of subsequent development.

1.4 Method

1.4.1 Classical statistics

The problem can be stated thus: given a particular site, we want to design a sampling scheme, consisting of trial trenches and/or boreholes, that will enable us to say with a chosen level of confidence, that there are no significant archaeological remains on the site, if our sample reveals no such remains. The definition of 'significant' is crucial for sample design. Clearly, we do not mean the presence of isolated artefacts, or even a small individual feature (*e.g.,* an isolated post-hole), but the existence of either a substantial assemblages of artefacts (*e.g.,* a flint scatter), or a functional feature or group of features (*e.g.,* a group of post-holes comprising a hut). It seems reasonable therefore to initially define 'significant' in terms of size, and, to a lesser degree, shape (*e.g.,* 'artefacts or features occupying an area more than 2m in maximum diameter'), while acknowledging that much more work needs to be done on this topic.

The impetus for this paper came from an outstanding example of statistical serendipity. I attended a talk given to the General Applications Section of the Royal Statistical Society by Mike Nicholson of the Directorate of Fisheries Research of the Ministry of Agriculture, Fisheries and Food, Lowestoft. His work was concerned with the 'problem of making inferences about the distribution of a species [of shellfish] when it is not observed in a survey area. Absence of the species implies no more than that the sampling points did not coincide with the occurrence of the species, not necessarily that the species is absent from the area' (Nicholson & Barry 1995, p. 74; see also Barry & Nicholson 1993). This seems to me to be mathematically analogous to the archaeological situation: for 'area' read 'site', for 'sampling point' read 'trial trench' or 'borehole' (his 'points' are not literal points, but may be small quadrats or circles), for 'species' read 'archaeological remains' (*e.g.,* flints). He distinguishes two reasons for such surveys:

1. when the species is beneficial, and we want to be reasonably sure that we have not missed a significant (*e.g.,* commercially exploitable) patch of it, as in the example of cockles on Holbeach Sands, The Wash (Barry & Nicholson 1993, pp. 359–61).

2. when the species is undesirable, and we want to be reasonably sure that it has not invaded an area, as in the case of Manila clams in Poole Harbour (Nicholson & Barry 1995, pp. 76–7).

Paradoxically, some archaeological situations may be of the latter type, since the aim may be to be reasonably sure of being right if we say that there are no significant remains on the site.

In their first paper, Barry & Nicholson (1993) consider the probability of detecting a single circular patch, of radius r, in an area where there is one sampling point per area d^2. That is, if there are N sampling points in an area of size A, d^2 is defined by the equation

$$d^2 = A/N.$$

They discover that the key statistic is what they call the 'standardized patch radius' R, defined by $R = r/d$, *i.e.*, the ratio of the patch radius to a notional distance between sampling points. They compare the average probabilities of patch detection for four sampling designs — square and triangular lattices, random design and transect design. They plot the values of this probability for values of R from 0 to 1.0, showing that the lattice designs are superior to the others (with the triangular slightly superior to the square, and the random slightly superior to the transect; Barry & Nicholson 1993, Fig. 4). This conclusion appears to conflict with archaeological expectations, which are perhaps unduly influenced by the experience of Flannery (1976), that lattice designs are inferior to random designs because their sampling points may fall between regularly-spaced archaeological features. The difference in perception arises because Barry and Nicholson's work concerns the probability of detecting a single patch in the area.

Archaeologically, one could use this work to calculate a size of sample needed to detect (with specified probability) a single 'patch' of archaeological remains on a site. The higher this probability is, the higher the certainty that no patch is present if one is not detected.

In their second paper, Nicholson & Barry (1995) show how the incorporation of prior information (*e.g.*, about the likely density of a species in the area) can, by means of Bayesian statistical analysis, improve the efficiency of a sample design. They start in classical fashion with the case of 'an effectively infinite area in which $\theta (0 < \theta < 1)$ is the probability that the species is present at a randomly spaced sampling point' and that a sample of N randomly spaced sampling points fails to detect any examples of the species. The estimate of θ is then zero, but they show that the upper $100p\%$ confidence limit for θ can be constructed by finding the largest value θ_p such that

$$1 - p < \begin{array}{c}\text{probability that the species}\\ \text{is not detected when}\end{array} \theta = \theta_p$$

i.e.,

$$1 - p < (1 - \theta_p)^N$$

which by taking logarithms becomes

$$N > \frac{\log(1-p)}{\log(1-\theta_p)}$$

They quote McArdle (1990), who derived from this inequality the equation:

$$N = \frac{\log(1-p)}{\log(1-\theta_p)}$$

as a formula (I) for the minimum sample size needed to meet specified values of p and of θ_p.

Some values of N, calculated for chosen values of θ and of θ_p, are shown in Table 1.1. Note that this table shows the sample size (N) needed for us to be $p\%$ certain that the proportion of a site occupied by archaeological deposits is less than a specified amount (θ_p), in the case when no deposits are actually found in the sample. For example, if we want there to be a 95% probability that the proportion is less than 2%, we need a sample size of about 114.

So far, this is all straightforward Classical statistics. It reflects work done in the USA in the 1980s (*e.g.*, Nance 1981; Read 1986; Shott 1987) and 1990s (Sundstrom 1993), and in England in the 1990s (Champion *et al.* 1995). Nicholson and Barry's formula (I) is essentially the same as that quoted by Read (1986, p. 484) and Shott (1987, p. 365) on regional and intrasite scales respectively. Shott considered the detection of a rare class of feature on a multi-feature site, rather than the detection of a 'patch' of archaeological deposits on a 'sparse' site, but the statistical argument is, so far, the same.

1.4.2 A Bayesian approach

The advance that Nicholson and Barry make is to suppose that we may have prior information about likely values of the parameter θ, which can be incorporated into the analysis by the use of Bayesian methods. They model the prior distribution of θ as a Beta distribution with parameters a and b. The Beta is a standard statistical distribution with the useful properties that (a) if no archaeological deposits are found, then the posterior distribution of θ is also a Beta distribution (Cox & Hinkley 1974), and (b) as Nicholson & Barry (1995, p. 75) point out, appropriate values of a and b lead to an intuitive understanding of the role of the prior distribution and its parameters, and a simple solution for the posterior distribution.

They show that if a is set to the value 1, different prior beliefs about θ can be represented by choosing different values of b. For example, choosing $b = 1$ corresponds to a prior belief that all values of θ between 0 and 1 are equally likely. They next show that with $b > 1$, values of $\theta > 0.5$ become progressively less likely, while with $b < 1$, values of θ above 0.5 are more likely (see Fig. 1.1). Prior belief that archaeological remains are unlikely to be found therefore corresponds to high values of b.

They go on to show that the value of b acts as the 'sample size of a hypothetical survey from which the species was absent' and give the example that 'a prior belief expressed as being 95% ($100p$) sure that θ is less than 0.1 (θ_{prior}) gives $b = 28.4$, equivalent to a hypothetical prior survey of 29 sampling points

p	θ_p					
	0.01	0.02	0.05	0.10	0.20	0.50
0.90	229	114	45	22	10	3.3
0.95	298	148	58	28	13	4.3
0.99	458	228	90	44	21	6.6

Table 1.1: Minimum size of sample N needed to achieve specified values of p and of θ_p.

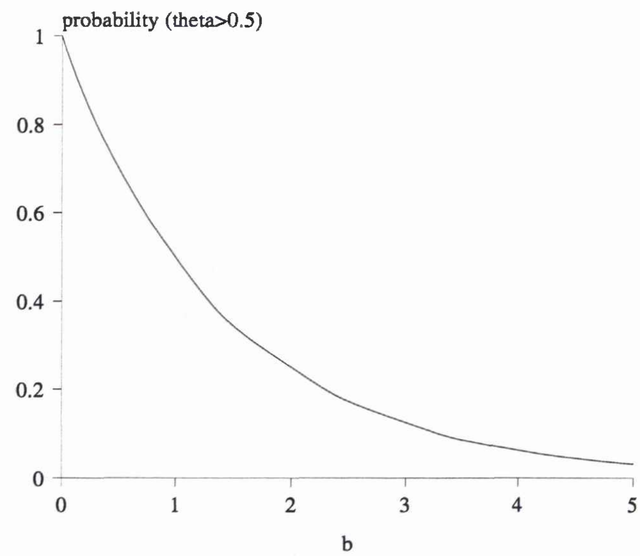

Figure 1.1: Prior probability that $\theta > 0.5$, as a function of b (after Nicholson & Barry 1995, Fig. 1).

in which the species was not observed.' (Nicholson & Barry 1995, p. 76).

This may seem to the archaeologist like mathematical trickery, in that several sampling points have been created out of thin air (or our own prior beliefs). As a counter to such suspicions, Nicholson and Barry point out that the Classical formula (I) is identical to the corresponding Bayesian formula for N only when $b = 0$, which corresponds to a belief that $\theta = 1$. Since archaeologists would not carry out an evaluation if $\theta = 1$ (which means that archaeological remains exist across the entire site), it is reasonable to accept a prior belief in a lower value of θ, and thus in a value of b which reduces the required sample size.

Whatever prior belief we choose to employ, we can calculate a sample size that corresponds to chosen levels of p (our confidence level) and of θ_p (the upper acceptable limit for the proportion of the site on which archaeological remains survive). The parameter θ_p must be chosen to make the area of any surviving archaeological patches (*i.e.*, $A\theta_p$) 'insignificant'.

1.5 A worked example

Suppose a development site of 1 ha. (100m × 100m = $10{,}000\text{m}^2$) is to be evaluated for possible archaeological remains. None have been found on the site, but other work in the locality suggests that there may be some. How does the archaeologists design his/her project? For purposes of illustration, we suppose that site conditions make the use of geophysical prospection either impossible or inconclusive.

1.5.1 Setting the parameters

The archaeologist must first decide on (i) the appropriate definition of 'significant archaeological remains' and (ii) the acceptable posterior probability that there are no such remains on the site, if none are found in the evaluation. Such parameters may well be the subject of future guidelines, but for the time being (s)he is on his/her own.

The choice for (i) depends on the type of remain that might be found. Well-defined types would suggest precisely-defined 'patches', for example if *grubenhäuser* were anticipated, rectangular patches of 3m × 2m might be specified, while if Iron Age roundhouses were anticipated, circular patches of 10m diameter might be suspected. If there were no strong hints towards a particular type of remain, the preference might be to specify an area or proportion of the site. This might be particularly relevant to urban sites, where the question might be not so much whether remains of a particular type once existed on the site, but how much (if any) has survived more recent development. For purposes of illustration, this approach is adopted, and 'significant archaeological remains' are defined as remains with a total extent of 100m^2. or more on the site. Since $100/10{,}000 = 0.01 = 1\%$, this is equivalent to setting the critical value of θ to 0.01.

The choice for (ii) depends on the likely cost of being 'wrong': *i.e.*, what will be the cost if the archaeologist says that there are no significant archaeological remains, but they are encountered in the development (the situation that PPG 16 is supposed to prevent occurring). The greater the cost, the higher the posterior probability should be. This raises all sorts of questions — is the cost that of the delay to the developer caused by the investigation of unexpected remains, or of redesigning the development around them (both of which are likely to be high) or the cost to archaeology of the loss of unrecorded remains (which is so far unquantifiable)? This is a very wide and difficult subject; for the time being the archaeologist avoids it and makes an arbitrary decision, say 90%. To sum up, the archaeologist has decided that (s)he wants to be 90% certain that, if (s)he says there are no significant archaeological remains on the site, the total extent of any remains is less than 100m^2.

1.5.2 Devising the strategy

The archaeologist must now design a sampling procedure that will meet these aims at the lowest possible cost. The cost includes not only person-hours worked, but also the total length of the time spent on site (and hence the potential delay to the developer). The aspects that must be decided are:

1. the size and shape of the interventions (*e.g.*, 1m^2 test pits; 2m-wide trenches)
2. their number and location
3. the means of excavating them (by hand, mechanical excavator, *etc.*).

These questions are inter-related, and the aim is to find the 'best' combined answer to all three. For purposes of illustration, we start from the mathematically simplest situation, in which the interventions are so small that they may be regarded as mathematical points on the site (*e.g.*, boreholes, or perhaps 30cm-square 'shovel tests', *e.g.*, Krakker *et al.* 1983). Application of the Classical formula (I), with $p = 0.90$ and $\theta_p = 0.01$, gives $n = 229$ as the number of interventions that would be needed. But, as we have seen in 1.4.2, this corresponds in Bayesian terms to the unlikely situation of a prior belief that $\theta = 1$. If we adopt a position of vagueness, *e.g.*, that all values of θ are equally likely, then $b = 1$ and the difference to n is trivial. But if we are prepared to make stronger statements about θ we can reduce the value of n appreciably.

The prospect of undertaking a large number of very small interventions does not appeal (Champion *et al.* 1995, p. 39), so we consider larger ones, say (for example) 2m-square test pits. Mathematically, increasing the size of the intervention is equivalent to increasing the area of the 'target patch'. For example, a 10×10m patch would be hit by points that actually fall within it, but it would be hit by 2m squares whose centres lie up to 1m from its edge, *i.e.*, within a total area of $12 \times 12\text{m} = 144\text{m}^2$. Thus, in this simple case, the effective value of θ, θ_e, is increased from 0.1 to 0.144 by the use of 2m-square test pits, and the corresponding value of n is reduced to 158, which appears to be much more expensive than 229 shovel-tests. But two points inflate the comparison:

1. the assumption that the archaeological remains form a single square patch, and
2. that the remains are more likely to be seen in a large intervention than in a small one.

We look at each in turn.

1. the effective area of the target patch is increased by a proportion that depends crucially on its shape and nature (one patch? several? square? linear?). For example, a linear patch, say 50 $\times$ 2m, has an effective area of 52×4m, *i.e.*, 208m^2, so that θ_e is increased to 0.0208, and n is reduced to 110. Considering several small patches, for example ten of 4×2.5m each, gives a larger increase in the effective area, in this case to $10 \times 6 \times 4.5 = 260\text{m}^2$, corresponding to $\theta_e = 0.026$ and leading to $n = 87$. The process could be carried to ridiculous extremes, for example 100 patches of 1×1m gives an effective area of $100 \times 3 \times 3\text{m}^2, \theta_e = 0.09$, and $n = 24$. Clearly archaeological opinion about the likely size and shape of 'significant archaeological remains' is very important at this stage. It should be noted that the area excavated in even as few as 24 such test pits ($24 \times 2 \times 2 = 96\text{m}^2$) is more than the total area of 229 shovel tests ($220 \times 0.3 \times 0.3 = 20.6\text{m}^2$), suggesting that the most efficient strategy is a very large number of very small interventions. This would be the case, but for the next point.
2. a second important point is that, if significant archaeological remains are present, they are more likely to be detected in a large intervention than in a small one. This is called the 'visibility' of the remains, or the 'site detection probability' and has been discussed by Champion *et al.* (1995). Its effect is to make small interventions relatively less efficient, since they have a higher probability of not detecting remains even when they are present in them.

It may be that at this stage the cost of the outcome of these considerations is seen as too high, perhaps in relation to the development value of the site. If so, it can be reduced only by reducing the value of p (*i.e.*, of increasing the risk of not detecting significant archaeological remains), or of increasing the value of θ_p (*i.e.*, of increasing the size of the remains that the archaeologist is prepared to 'write off').

Thus the choices listed at the start of this section depend on a complex interplay between the size and shape of the sorts of patches of significant archaeological remains that might be expected, with their visibility on interventions of various shapes and sizes and dug by various means, and what is perceived as a reasonable cost of evaluation.

1.5.3 Designing the sample

After the archaeological and mathematical complexities of 1.5.1 and 1.5.2, this is a relative formality. The choice lies between a purely random and a more systematic layout of the chosen number of trenches and/or test-pits; there seems to be agreement between archaeologists and biologists that the systematic is probably better, and that within the systematic designs a triangular (also called hexagonal) lattice is probably the best (Champion *et al.* 1995, p. 39; Barry & Nicholson 1993).

1.6 Conclusion

The design of archaeological field evaluations has already passed beyond the stage at which it was believed that there was a minimum sample size (*e.g.*, 2%) that was needed for a successful evaluation (Champion *et al.* 1995, p. 36). It is now accepted that design must be based on a complex interplay of factors — type and extent of expected remains, definition of 'archaeological significance', size, shape and method of archaeological interventions, and the probability of recognising different types of remains in different types of intervention.

A valuable next step would be to require designs of evaluations to carry formal statements of quality assurance, the most fundamental of which would be a lower limit on the probability of a site containing no significant archaeological remains, given that none are found in the evaluation. This would raise archaeological (almost political) questions about the definition of significant archaeological remains, and about an acceptable probability level of faling to detect them. These problems have always existed, hidden beneath the cloak of professional judgement; it is better hat they be made explicit and discussed openly.

At this point, the use of a Bayesian approach has much to offer. First, archaeologists seem to find arguments or requirements based on subjective probabilities easier to grasp intuitively than ones based on Classical hypothesis testing. Second, by incorporating prior knowledge, a Bayesian approach can reduce the sample size needed to meet a specification, and thus reduce fieldwork costs. There are technical questions to be answered, such as the suitability of the Beta distribution (other than its sheer convenience), and the choice of its parameters, but these are not the province of the archaeologist. However, archaeologists will need to gain experience in articulating their prior beliefs.

A two-fold approach is needed to advance the subject: (i) advocacy of the practice of designing evaluations so that such probability statements can be made, (ii) provision of the infrastructure (software and training) that will enable archaeologists to achieve (i). This paper is a first step towards (i); funding for (ii) is currently being sought.

Acknowledgements

My prime thanks must go to Mike Nicholson, whose excellent talk was the starting point of this work. I was encouraged to pursue it by Peter Hinton and Kris Lockyear, and this paper has been greatly improved through the perceptive comments of a referee.

References

ANDREWS, G. 1991. *Management of Archaeological Projects.* English Heritage, London.

BARRY, J. & M. NICHOLSON 1993. 'Measuring the probability of patch detection for four spatial sampling designs.' *Journal of Applied Statistics* 20(3): 353–361.

BIDDLE, M. 1994. *What future for British archaeology?* Oxbow Lecture 1. Oxbow Books, Oxford.

CARVER, M. O. H. 1993. *Arguments in Stone: Archaeological Research and the European Town in the First Millenium.* Oxbow Books, Oxford.

CHAMPION, T., S. J. SHENNAN & P. CUMING 1995. *Planning for the Past Volume 3: Decision-making and field methods in archaeological evaluations.* University of Southampton and English Heritage, Southampton and London.

COX, D. R. & D. V. HINKLEY 1974. *Theoretical Statistics.* Chapman and Hall, London.

DENSEM, R. & D. SEELEY 1982. 'Excavations at Rectory Grove, Clapham, 1980–81.' *London Archaeologist* 4(7): 177–184.

DEPARTMENT OF THE ENVIRONMENT 1990. *Planning Policy Guidance: Archaeology and Planning.* Department of the Environment, London.

FLANNERY, K. V. (ed.) 1976. *The Early Mesoamerican Village.* Academic Press, London.

GREENWOOD, P. & C. MALONEY 1995. 'London fieldwork and publication round-up 1994.' *London Archaeologist* 7(13).

HODDER, I. & C. R. ORTON 1976. *Spatial Analysis in Archaeology.* Cambridge University Press, Cambridge.

KRAKKER, J. J., M. J. SHOTT & P. D. WELCH 1983. 'Design and evaluation of shovel-test sampling in regional archaeological survey.' *Journal of Field Archaeology* 10: 469–480.

McArdle, B. H. 1990. 'When are species not there!' *Oikos* 57: 276–277.

McCracken, S. & C. Phillpotts 1995. *Archaeology and planning in London. Assessing the effectiveness of PPG 16.* Standing Conference on London Archaeology, London.

Nance, J. D. 1981. 'Statistical fact and archaeological faith: Two models in small site sampling.' *Journal of Field Archaeology* 8: 151–165.

Nicholson, M. & J. Barry 1995. 'Inferences from spatial surveys about the presence of unobserved species.' *Oikos* 72: 74–78.

Read, D. W. 1986. 'Sampling procedures for regional surveys: a problem of representativeness and effectiveness.' *Journal of Field Archaeology* 13: 477–491.

Shott, M. J. 1987. 'Feature discovery and the sampling requirements of archaeological evaluations.' *Journal of Field Archaeology* 14: 359–71.

Sundstrom, L. 1993. 'A simple mathematical procedure for estimating the adequacy of site survey strategies.' *Journal of Field Archaeology* 20: 91–96.

Clive Orton
Institute of Archaeology
University College London
31–34 Gordon Square
London WC1H 0PY

c.orton@ucl.ac.uk