Counting Broken Objects:
The Statistics of Ceramic Assemblages

C. R. ORTON & P. A. TYERS
*Department of Prehistory, Institute of Archaeology, University College London,
31–34 Gordon Square, London WC1H 0PY.*

Summary. To estimate and compare the proportions of different types of pottery in ceramic assemblages ('type' being defined according to need, e.g., functional type, chronological type), archaeologists need a measure of the amount of each type. The most suitable proposed measure is the 'eve' (estimated vessel equivalent), for which each measurable fragment is scored as a fraction of a complete vessel.

The 'Pie-slice' project, funded by SERC-SBAC and the British Academy, uses a new statistical transformation (the 'pseudo-count transformation'), which converts eves into 'pies' (pottery information equivalents), with the property that an assemblage of n pies has the same error structure as one of n complete objects. This enables assemblages to be compared by techniques for categorical data, mainly log-linear and correspondence analyses; reduction of the data matrix is usually needed. Case studies of a variety of problems are presented.

1. The need for quantification

1.1 Introduction

An important part of the archaeological study of pottery is the comparison of assemblages, groups of pottery that in some sense belong together. This can be done at many levels: for example, the pottery from a single context or a feature (e.g., a rubbish pit), from a phase of a site, or from a period in a town. An assemblage can be characterised by its composition, i.e., the proportions of different types of pottery of which it is made up; the definition

of 'type' is flexible and can be chosen to meet particular needs (see below). Assemblages can be compared in terms of their compositions. Different levels of assemblage and different definitions of type can be used to answer different questions, the main sorts of which are chronological, spatial and social/functional.

To be able to characterise assemblages in this way, we need a measure of the amounts of the various types of pottery of which they are comprised. If an assemblage consisted of whole vessels there would be no problem—one could just count the pots. In practice whole vessels are relatively rare and assemblages consist of pottery in varying degrees of fragmentation and survival. Measuring the quantities of different types under these circumstances is the core problem. In section 2 we shall see how it has been approached in the past.

1.2 Chronological questions

It is often assumed that the usage of types of pottery follows a simple continuous and unimodal pattern over time—introduction, increasing usage, steady usage, decreasing usage, demise. If this is so, the relative proportions of types in assemblages which form a chronological sequence at a location will also follow a simple pattern (Figure 1). If the proportions are known but the sequence is not, it can be reconstructed by organising the compositions into this sort of pattern. This is the basis of the technique of seriation, which
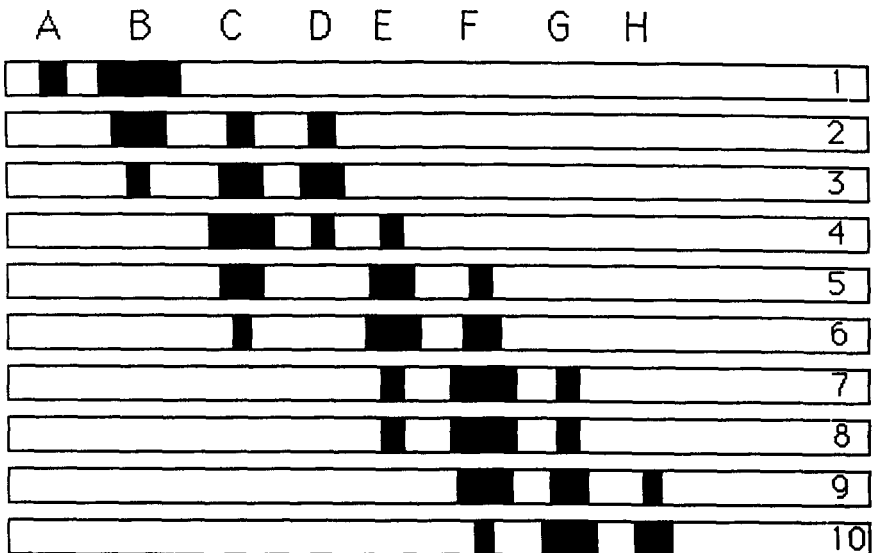

Figure 1. Hypothetical seriation of the proportions of types A-H in contexts 1-10.

is of great value in archaeological research, although its assumptions are sometimes challenged (see Marquardt 1978 for a review; the definitive work is likely to be Laxton, forthcoming).

1.3 Spatial questions

If a type of pottery is traded or distributed from a centre of production, the proportion of it that one would expect to find in assemblages decreases as we move away from the centre. The rate and shape of this fall-off can be interpreted in terms of the methods of distribution and marketing (Hodder and Orton 1976, 104–119). The problem can be studied from either end, (i) looking at the distribution of a particular type or ware across a region (e.g., Lyne and Jefferies 1979, Figures 42–53) or (ii) looking at all the sources of pottery found at a location (e.g., Green 1980, Figures 42–45).

1.4 Social/functional questions

On all but the smallest sites, different parts may have been used for different purposes, e.g., cooking and eating. This may be reflected in the usage of different types of pottery, and, if methods of rubbish disposal were localised, in the archaeological record (e.g., rubbish pits). This approach formed the basis of Millett's (1979a) study of pit-groups from the late Roman fort of Portchester. On a wider scale, social differences within a town might be reflected in differences between assemblages from different parts of the town. There has been surprisingly little work of this nature; a good example is Redman's (1979) work at Qsar es-Seghir.

1.5 Summary

In all these circumstances, which use different levels of assemblage and definitions of types, there is a need for a measure which can be used to determine the proportions of the types, and to compare them between assemblages. In the next section we shall look at various measures that have been proposed, and attempts that have been made to compare them.

2. Previous work

2.1 The initial phase

The first use in the field was seriation (Petrie 1899). Its application to quantified assemblages of pottery started in the USA nearly twenty years later (Spier 1916). Comparisons were made in terms of numbers of fragments of each type (the sherd count), because this was the level at which pottery was

generally studied in the USA at that time. Techniques developed through to the 1960's (e.g., Ford 1962), but the use of sherd counts was not challenged.

Realisation that even apparently 'coarse' wares could be distributed over wide areas came in the 1930's (e.g., Shepard 1942), but it was not until the 1960's and '70s that quantified distributional studies became common (e.g., Peacock 1977). No great attention was paid to the question of measures; for example, even the often-quoted work of Fulford and Hodder (1974) relied on sherd counts.

2.2 Competing measures

Part of the revolution that occurred in pottery studies around 1960 (Orton *et al.* forthcoming) was the opening-up of the question of measures and the emergence of rivals to the ubiquitous sherd count. The first alternatives were number of vessels represented (Burgh 1959) and weight (Solheim 1960), followed by vessel-equivalents (the idea can be found in Bloice 1971 and Egloff 1973; the term was coined in Orton 1975—see below for definition), surface area (Glover 1972, 93–6; Hulthén 1974) and displacement volume (Hinton 1977). The last two are very similar to weight, and need no explanation; the term 'vessel-equivalent' may be less familiar. Starting from the idea that every sherd is a certain proportion of the whole pot of which it once formed part, we can (in theory) assign these proportions to sherds as 'scores' and add them up to find the total amount of a type. Since a whole pot would give a score of 1, we can say that a group of sherds with a total score of x is equivalent to x pots (x is usually not a whole number). In practice it is not usually possible to assign a score to every sherd, and one is restricted to sherds such as rim sherds whose size can be measured in terms of the proportion of some whole (e.g., a complete rim). Since we are sampling the measurable sherds from an assemblage, we refer to the estimated vessel equivalent (abbreviated to eve).

2.3 Comparisons

Once there was more than one measure, attempts were made to compare them. Glover (1972, 96), comparing sherd count, weight and surface area, concluded that "any one would be quite accurate as a measure of frequency". Hinton (1977) compared sherd count, rim sherd count, weight and displacement volume, concluding that weight was the fastest but sherd count probably the most accurate measure, but of what it is not clear. Millett (1979b) compared sherd count, weight, adjusted weight (an estimate of surface area) and minimum number of vessels; he concluded that they were all highly correlated but, for practical reasons, weight was probably the best.

The development of our view can be traced in a series of papers (Orton 1975; Orton 1982; Orton and Tyers 1990). These studies have ruled out sherd count and number of pots as biased, and favour eves (where practicable) with weight as a respectable but less useful measure.

3. The pie-slice project

3.1 History of the project

Our work on this topic (Orton 1975; 1982) had achieved some results, but had ended in frustration because of its inability to attach standard deviations to estimates of proportions of types, and hence to test the significance of observed differences between assemblages. The publication of the CODA technique (Aitchison 1986) seemed to offer a way out of the impasse, and a two-year project, the Statistical Analysis of Ceramic Assemblages, started in 1988 with the theoretical aims of:

   i) being able to set confidence limits on the proportions of a ceramic assemblage that belong to different types,

   ii) being able to compare, numerically and graphically, the compositions of two or more assemblages in terms of the proportion of each type present in each assemblage, and to assess the statistical significance of the differences between them, and

   iii) the practical aim of applying the theory to assemblages from a wide range of sites, of different types and periods, to assist in their interpretation, and hence the interpretation of the sites themselves. It was expected that the work would lead to recommendations about the recording of ceramic assemblages, and that CODA would be heavily used. The fortuitous non-availability of the CODA package at the start of the project led to the development of a new approach, described below, which made CODA unnecessary for pottery assemblages.

3.2 Basic theory

The numbers of records of the jth type in an assemblage is denoted by m_j, ($j = 1, \ldots, T$), and the total number of records by m.
   The measure of the ith record of the assemblage is denoted by w_i, ($i = 1, \ldots, m$).
   The total measure of a type is denoted by W_j ($j = 1, \ldots, T$), and the overall total by W (note that upper case is used for type and assemblage totals, and lower case for individual values).

The symbol $\sim j$ refers to all types except the jth, and Σ_j means summation over the jth type.

The unadjusted sum of squares $\Sigma_j w_i^2$ is denoted by S_j^2.

3.2.1 Estimates of proportions in a single assemblage

The proportion p_j is estimated by:

$$\hat{p}_j = W_j/W, \text{ for } j = 1, \ldots, T.$$

We define two new variables $x(j)$ and $y(j)$ for all records by:

$$x_i(j) = w_i$$

$$y_i(j) = w_i \text{ if the ith record relates to the jth type,}$$

$$= 0 \text{ otherwise.}$$

Then $\hat{p}_j = W_j/W = \Sigma y_i(j)/\Sigma x_i(j)$, a ratio estimate.

Cochran (1963, 30–1) gives a formula for the variance of a ratio estimate, leading to:

$$\text{var}(\hat{p}_j) \approx (m/(m - 1)W^4)\{W_{\sim j}^2 S_j^2 + W_j^2 S_{\sim j}^2\} \tag{1}$$

and:

$$\text{cov}(\hat{p}_j,\hat{p}_k) = -(m/(m - 1)W^4)\{WW_j S_k^2 + WW_k S_j^2 - W_j W_k S^2\} \tag{2}$$

3.2.2 Comparing proportions in two or more assemblages

Given any type j, we can compare $\text{var}(\hat{p}_j)$ with the variance of an estimate based on a binomial model, i.e., on an assemblage of complete vessels.

In the latter case, the formula is $\text{var}'(\hat{p}_j) = \hat{p}_j \hat{q}_j/n$, for a population of size n, where $\hat{q}_j = 1 - \hat{p}_j$.

So the variances would be the same if $\text{var}(\hat{p}_j) = \hat{p}_j \hat{q}_j/n$.

We can turn this round and define $n_j = \hat{p}_j \hat{q}_j/\text{var}(\hat{p}_j)$, so that n_j is the number of whole vessels that would give the same value of $\text{var}(\hat{p}_j)$ as our sample of m measurable records.

The full formula is:

$$n_j = ((m - 1)/m)W_j W_{\sim j} W^2/\{W_{\sim j}^2 S_j^2 + W_j^2 S_{\sim j}^2\} \tag{3}$$

The weakest condition so far found for n_j to be the same for all j is that $S_j^2/W_j = c$ for all j, which is satisfied if all types have the same mean and variance of w. In this case we pool our estimates of the mean and sum of squares of w, obtaining W/m and S^2 respectively, and replace W_j by $W(m_j/m)$, S_j^2 by $S^2(m_j/m)$. So:

$$n_j = ((m - 1)/m)(m_j/m)(m_{\sim j}/m)W^4/\{W^2(m_{\sim j}/m)^2 S^2(m_j/m)$$

$$+ W^2(m_j/m)^2 S^2(m_{\sim j}/m)\}$$

$$= ((m - 1)/m)W^2/\{(m_{\sim j}/m)S^2 + (m_j/m)S^2\}$$

$$= ((m - 1)/m)W^2/S^2 \tag{4}$$

So that if all the types have the same mean and variance of w, we can by pooling estimates obtain a common value $n = n_j$ for all j.

We now have the background theory we need to look at the comparison of several assemblages, say A of them.

We have vectors of measures $\{W_{r1}, \ldots, W_{rT}\}$,

of numbers of observations $\{m_{r1}, \ldots, m_{rT}\}$,

of sums of squares $\{S_{r1}^2, \ldots, S_{rT}^2\}$,

and estimates of proportions $\{\hat{p}_{r1}, \ldots, \hat{p}_{rT}\}$,

and variance-covariance matrices $\| \operatorname{cov}(\hat{p}_{rj}, \hat{p}_{rk}) \|$,

for all $1 \leqslant r \leqslant A$.

We want to compare the vectors of estimated proportions, e.g., to test a hypothesis H_o : all assemblages are 'the same', i.e., can be thought of as samples from the same parent population.

We assume that each assemblage has a single n-value, which we call n_r, for $1 \leqslant r \leqslant A$.

We replace each W_{rj} by $n_r(W_{rj}/W_r)$, for $j = 1, \ldots, T$ and $r = 1, \ldots, A$.

Calling the new numbers $W_{rj}' = (n_r/W_r)W_{rj}$,

we have $W_r' = n_r$ for all assemblages r.

The estimates of proportions are unchanged:

$$\hat{p}_{rj}' = W_{rj}'/W_r' = W_{rj}/W_r = \hat{p}_{rj},$$

and so are their variances and covariances:

$$\operatorname{var}(\hat{p}_{rj}') = (m/(m - 1))(S_r'^2/W_r'^2)(m_{rj}/m_r)(m_{r\sim j}/m_r)$$

$$= (m/(m - 1))(S_r^2/W_r^2)(m_{rj}/m_r)(m_{r\sim j}/m_r)$$

$$= \operatorname{var}(\hat{p}_{rj}),$$

since $S_r'^2 = (n_r/W_r)^2 S_r^2$ and $W_r'^2 = (n_r/W_r)^2 W_r^2$.

And:

$$\operatorname{cov}(\hat{p}_{rj}', \hat{p}_{rk}') = (m/(m - 1))(S_r'^2/W_r'^2)(m_{rj}/m_r)(m_{rk}/m_r)$$

$$= (m/(m - 1))(S_r^2/W_r^2)(m_{rj}/m_r)(m_{rk}/m_r)$$

$$= \operatorname{cov}(\hat{p}_{rj}, \hat{p}_{rk})$$

for the same reason.

Recalling (4) and writing $m_{rj}/m_r \approx \hat{p}_{rj}$, etc., since the assemblages are homogeneous, we have:

$$\operatorname{var}(\hat{p}_{rj}) \approx \hat{p}_{rj}\hat{q}_{rj}/n_r$$

and

$$\text{cov}(\hat{p}_{rj}, \hat{p}_{rk}) \approx \hat{p}_{rj}\hat{p}_{rk}/n_r \ .$$

But these are exactly the same as the variance and covariances we would obtain from a multinomial distribution with parameter p and sample size n.

This is a very important result. It means that, as a large-sample approximation, we can treat the transformed data as a series of samples from multinomial distributions. We can therefore treat the data collectively as a contingency table, and use any of the theory appropriate to contingency tables (e.g., log-linear models or correspondence analysis; Greenacre 1984).

For the first time, this approach enables to make proper statistical comparisons between the proportions of different types in different assemblages. We refer to the transformed values W'_{rj} as pseudo-counts. They are not integers, but can be treated for statistical purposes as if they were. We call the transformation expressed in equation (4) the 'pseudo-count transformation' (pct), the numbers n_j 'pseudo-counts' (because they behave like counts but are not integers), and the total n the 'pseudo-total'. When applied to pottery the pseudo-counts are called 'pies' (pottery information equivalents) because one pie contains as much information (in the statistical sense) as one whole pot.

3.2.3 Log-linear and quasi-log-linear analysis

Suppose we have a three-way (context-by-fabric-by-form) table of pseudo-counts **n**. To follow the standard notation (e.g., Fienberg 1977) we replace n by x, with subscripts i, j and k for three variables, usually context, fabric and form respectively. Context is treated as an explanatory variable, and fabric and form as response variables. We can construct nested models of increasing complexity and archaeological reality, from complete independence at one extreme to the saturated model at the other. The intermediate models correspond to different archaeological needs:-

i) the common (but not universal) situation of different forms being produced in different fabrics,

ii) a functional approach, in which the primary source of variation between contexts is thought to be function, represented by different proportions of the different forms present,

iii) a spatial (inter-site) or chronological approach, in which the primary source of variation between contexts is thought to be the sources of the pottery, representing either geographical or temporal variation (or both).

Within each model, we can calculate the estimates $\hat{m}_{ijk}$ and carry out a

goodness-of-fit test, using the likelihood ratio statistics $G^2(1)$, $G^2(2)$, $G^2(3)$, $G^2(4)$, where (Bishop *et al.* 1975, 125):

$$G^2 = 2\Sigma x \log(x/\hat{m}).$$

This approach enables us to find the simplest model that fits the data reasonably.

Unfortunately, there are many fabric-by-form combinations that cannot exist, and many fabric-by-context and form-by-context ones that do not exist in a particular dataset. The design of the data is incomplete, and the theory of quasi-log-linear models (Bishop *et al.* 1975, 177–228) must be used. This theory raises important and difficult questions about which zeros are structural and which are random, which also have crucial implications for the calculation of the number of degrees of freedom.

We have looked at three approaches to the problem of which zeros should be treated as structural:-

i) treat all zeros as random, i.e., use log-linear models [2.5]. This leads to high numbers of degrees of freedom, and a situation in which almost any model would fit the data [3.4] (references in square brackets are to sections of the archival report *Statistical Analysis of Ceramic Assemblages*; see Section 3.4).

ii) treat zeros as structural if they correspond to a zero in a marginal two-way table [2.6.5]. This approach ('conditioning on all the data') gives realistic degrees of freedom but obscures the points which may be of most interest archaeologically.

iii) treat zeros as structural if they correspond to a zero in a marginal two-way table which has already been shown to relate to a significant interaction ('conditioning on the model' [2.6.2]). This gives rise to fewer degrees of freedom than (i) but more than (ii); more importantly, it seems to correspond most closely to archaeological reality. It does however seem prone to trouble from sparse datasets (see below).

3.3 Problems

Two main problem areas which require a theoretical response have been encountered: (i) inhomogeneity and (ii) sparseness.

3.3.1 Inhomogeneity

Inhomogeneity occurs when not all types in an assemblage have the same distribution of w. There are two situations in which it is likely to arise: the first depends on site formation processes and the second on the nature of the types.

If the post-depositional history of pottery is seen as a series of 'events', at each of which further breakage and dispersal may occur, then $\bar{w}$ decreases (weakly) at each event. Types which have had a longer post-depositional history are likely to have a smaller $\bar{w}$, which in principle identifies them. This has long been known intuitively in archaeology and is known as residuality; for most comparative purposes it is useful to be able to isolate any residual types.

This argument rests on the assumption that all types break equally readily and at the same rate. In practice this is not the case, and we can identify types which are more resistant to breakage than the usual run of types. They are usually small and/or thick-walled or (if we are measuring rims) have a small rim diameter, e.g., flagons. We call them 'chunky' types. Their inclusion in an assemblage boosts both $\bar{w}$ and S^2 unrealistically; we have therefore developed a technique for adjusting for chunkiness [2.4].

Inhomogeneity seems to be a manageable problem. The latter sort is identified by unusually high values of $\bar{w}$ and can be allowed for statistically. The former is identified by unusually low values. We have devised an approach which should accommodate this effect (the 'hinged' contingency table) and intend to implement it in phase 2 of the project. Identifying either sort requires a multiple comparison method; unexpectedly, the most satisfactory was one of the earliest: Fisher's least significant difference (LSD) test (1935).

3.3.2 Sparseness

Even when using quasi-log-linear analysis, we encounter problems because of the many 'small' cells in the table. They are:

i) if there are many such cells, they may contribute greatly to the number of degrees of freedom, but little to the overall X^2 or G^2 statistic, thus masking any potential significance of other parts of the table,

ii) the presence of a very small expected value together with a positive observed value can give an abnormally high contribution to the X^2 or G^2 statistic.

The answer seems to be merge or delete rows and/or columns of the table to remove small cells. We adopted the criterion that all cells should have an expected value of at least 1.0 (see Craddock and Flood 1970), and devised a technique called 'simultaneous reduction of dimension' (srd) to achieve this aim (Orton and Tyers 1991). At present, use of this technique may cause problems in the interpretation of any subsequent analysis because it affects significance levels; this will be tackled in phase 2 of the project.

3.4 The present situation

In the interval between the two phases of the project (September 1990 to April 1991) we can stop and take stock. We have available:

i) a computer package written in C and running in 'command-line' style over Unix on 80386-based micros. It has three main elements—the pseudo-count transformation (pct), the simultaneous reduction of dimension (srd) and the quasi-log-linear analysis (qlla). It makes use of the correspondence analysis program (ca) in the iastats package (Duncan *et al.* 1988).

ii) computer-based catalogues of ceramics from Roman, medieval and post-medieval sites in Chelmsford*, Lambaesis (Algeria)*, London, Silchester basilica, Winchester, Worcester (urban sites), Usk (military site), Ewell (Surrey), Leicestershire*, Witham (Essex)* (rural sites) and Southwark* (kiln site). Funding for the computerisation of catalogues from the sites marked with a * was provided by the British Academy.

iii) an archival report on the project, including case-studies on sites in London, Silchester, Winchester, Usk and Ewell.

Dr. Tyers worked in France from January to April 1991, applying Pie-slice to assemblages of Roman pottery from various sites.

3.5 Future work

Three main theoretical tasks and one practical one face us in phase 2. The theoretical ones are:

i) extension of our current 'large-sample' theory to 'small-sample' problems. This is likely to require monte carlo methods and should yield information on the minimum size of a viable dataset.

ii) the treatment of inhomogeneous assemblages of the first sort defined in Section 3.3.1.

iii) most seriously, the interface between srd and qlla and ca needs attention. The use of srd to merge categories carries the risk of altering the real significance levels of tests carried out in subsequent qlla, making the differences between groups appear more significant than they actually are. There are also problems in combining archaeological and statistical criteria in decisions about merging categories. The answer to both is probably to make srd more flexible and to move away from an over-riding strictly hypothesis-testing approach.

The practical problem is to make this work accessible to ordinary archaeologists working on pottery, through the creation of a 'friendly front-end' running in a windowing environment such as Open Desktop.

4. Case studies

4.1 Introduction

The uses made of the datasets have been at four levels:

i) checking the operation of the programs,

ii) answering specific technical questions, e.g., on structural zeros,

iii) searching for the appropriate level of aggregation of fabrics, forms and assemblages for different questions,

iv) producing interpretable results for discussion with the originators of the datasets.

Artificial datasets have also been used to help with i) and ii). Questions at levels ii) and iii) are discussed elsewhere in this paper, as they arise. Here we concentrate on level iv) but are inhibited by the need to maintain confidentiality of unpublished data. We shall look at the results thematically rather than site-by-site.

4.2 Chronological patterns

So far, these have dominated our analyses, especially of the late pre-Roman and Roman periods. It is well known (see Madsen 1988, 24) that a chronological sequence should be represented by a 'horse-shoe' shaped curve (approximately a parabola) on a correspondence analysis plot. We have observed such patterns at Usk (AD 55–70), Lime Street, London (AD 70–160), and Silchester (*c.*15 BC–AD 60). Perhaps more interesting than the expected horse-shoe were the deviations from it:

i) at Lime Street, 'rag-bag' categories (e.g.,'fine imported wares') occupy locations well off the curve, towards the centre, because they are an amalgam of types of different dates,

ii) at Silchester, context-groups with apparently high proportions of residual material also lie off the inside of the curve, towards the 'early' end.

4.2.1 Lime Street [7.3]

Comparison of fabrics with phases gave an apparently horse-shoe-shaped curve (Figure 2), with fabric SHEL (shell-tempered ware) early in the sequence and BB1, BB2 (black-burnished wares) and MORT (mortaria) late in the sequence indicated by the ordering of the phases, but with most of the points bunched in the apex of the curve.

The removal of SHEL opens up the curve (Figure 3). Phase 1 (which has
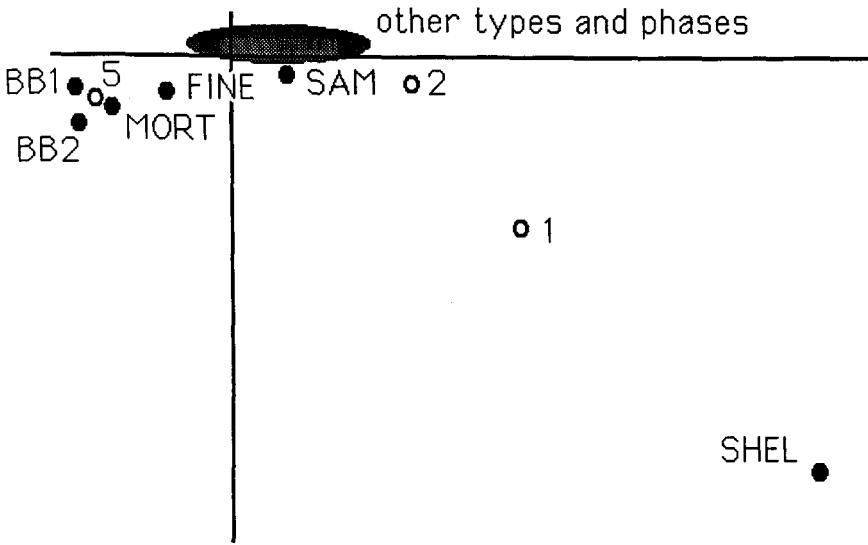
Figure 2. Correspondence analysis plot of fabric against phase, Lime Street site, City of London.

become very small by the removal of SHEL) and phase 6 (always very small and possibly residual) are out of sequence, but the major phases (3, 4 and 5) are in the 'right' order. The fabric FINE IMP now stands out from the curve; its removal would open out the curve.
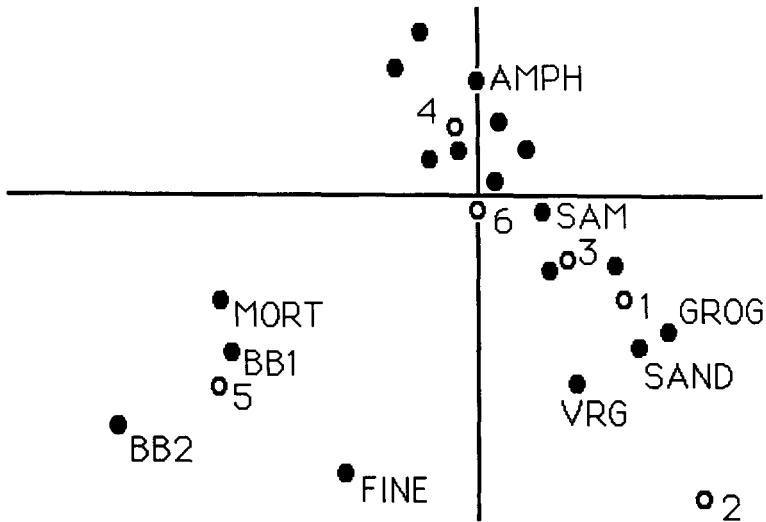

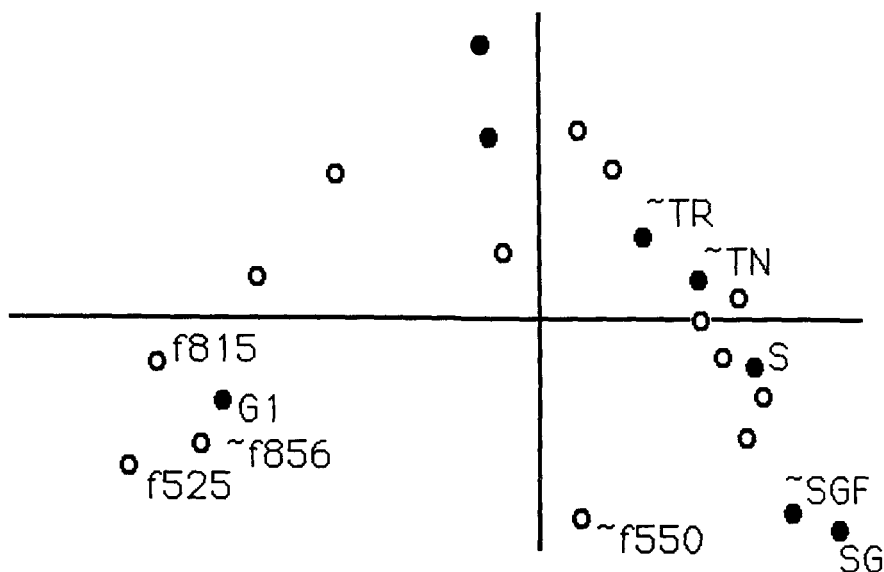**Figure 3.** As figure 2, with fabric SHEL removed.

Figure 4. Correspondence analysis plot of fabric against feature-group, Silchester basilica site.

4.2.2 Silchester basilica [7.2]

In the period to which the data relate, both forms and fabrics changed rapidly, with many introductions of new types. In the ca plot (Figure 4) the features and fabrics are arranged in a horse-shoe-shaped curve suggesting a broadly chronological progression. The 'early' end of the curve would be fabric G1 (the 'Belgic' grog-tempered wares) and feature groups ~f856, f525 and f815. They include most of the phase 1 and phase 2 deposits. The 'later' part of the curve includes the sequence ~TR, ~TN, and SG which would be the expected order of introduction for the fine ware fabrics terra rubra, terra nigra and South Gaulish samian.

The group ~f550 lies towards the centre of the curve and has a slight positive residual on fabric G1—the 'early' grog-tempered ware. This may suggest a higher proportion of residual material in these contexts.

4.3 Spatial patterns

The opportunity to investigate spatial patterns between sites has not yet arisen, but we intend to compare Chelmsford and Witham (Essex), and possibly some of the Leicestershire sites, in phase 2 of the project. It might be possible to compare Ewell (Surrey) with a London site, but it would be difficult to find a site that matches chronologically. Within-site analyses are considered under functional/social patterns (see below).

4.4 Functional/social patterns

We had hoped to observe such patterns on the Silchester and Aldgate, London (17th-18th century) sites. At Silchester we have so far been unsuccessful, because any such pattern seems to be masked by chronological changes in the forms available.

At Aldgate [7.4] there are four main types of feature—cesspits, cellars, make-up layers and structural features (e.g., foundation trenches)—at least one of which could be expected to correlate with the pottery. Although the obvious association of chamber pots with cesspits was detected, and a further one (cups and dishes with some make-up layers) was suggested, the results on the whole were difficult to interpret. Contexts tended to be grouped with contexts from other features rather than with other contexts from the same feature. While this could arise if several pits or cellars were open simultaneously and receiving contemporary material, this interpretation seems to be optimistic. Examination of the published functional typology (Orton and Pearce 1984, 63) shows that in this example broad categories of form cannot be simply equated with function, since the function of (for example) a bowl depends on whether it is decorated or not—information that is not available in this analysis.

It was the medieval tenements at Brook Street, Winchester [7.5], that gave the clearest indications of this sort of patterning. A preliminary analysis of forms by 'final phases' (phases within buildings) showed a three-way opposition between cooking-pots, jugs and lamps, with bowls (including bowl/dish and bowl/jar) occupying a central, roughly neutral, position (Figure 5). The final phases that can be linked with these forms through the ca show an association of lamps with industrial activity (dyeing, metal-working), jugs with a stone-built house and cooking-pots with less substantial houses. The interpretation of these results is at an early stage and must be seen as provisional.

4.5 Discussion

It is clear that the statistical analyses are not a panacea, and make careful archaeological preparation and interpretation more, rather than less, necessary. The definitions of fabrics, forms and assemblages, and their grouping into larger units for specific purposes (see below) have to be carefully thought out. But provided this is done, there does seem to be scope for the detection of patterns which might otherwise have gone unnoticed.
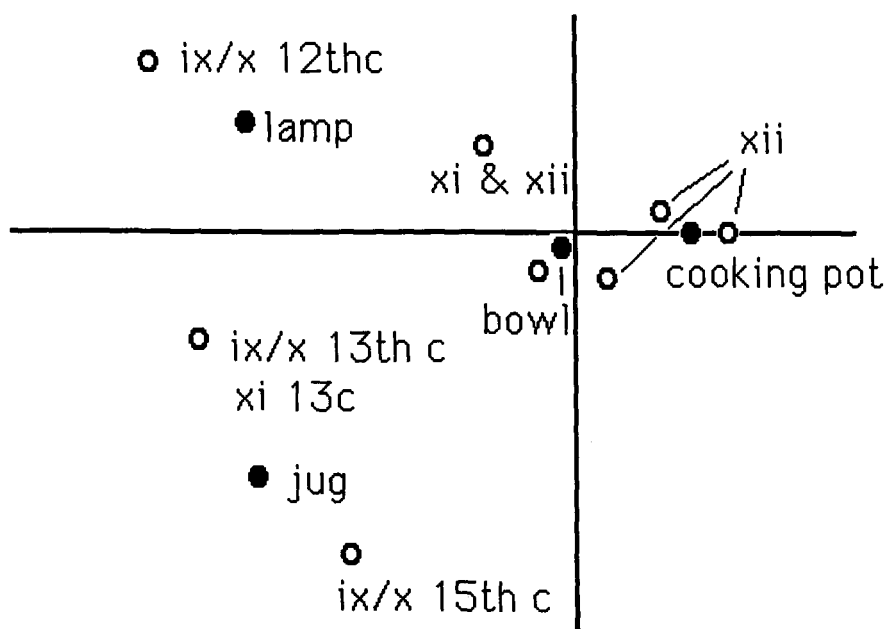
Figure 5. Correspondence analysis plot of form against final phase, Brook Street site, Winchester.

5. Implications for pottery studies

5.1 Recording methods

In Section 3.2, we saw that the weakest condition for the existence of a pseudo-total n was that:

$$S_j^2 / W_j = c \text{ for all types j;}$$

this seems to mean in practice that all types should have the same mean and variance of the measure w. Of the four measures considered in Section 2.2, only vessel-equivalents and number of vessels represented meet this requirement; weight could do so if it were possible to scale all types to a common weight (thus becoming an alternative estimate of the vessel-equivalent—the standardised weight approach). Earlier work (Orton 1975; see Section 2.3) has shown that number of vessels represented has serious and unpredictable biases; the only suitable measure is the vessel-equivalent or its estimate, the eve.

There remains the question of the best way to estimate the vessel-equivalent. The most commonly-used is the rim-eve, but other approaches should always be considered. When the necessary information is available, the standardised weight (see above) is likely to give a very good estimate.

The amount of pottery comprising the individual record is very important. Strictly, if we are to use the pct and subsequent statistical theory, each record should contain the measure of all measurable sherds of the same pot in the same assemblage (the 'measurable sherd family'). If one family is divided between two or more records (the 'over-detailed record'), the theory is not applicable because the observations are not independent. If two or more families make up the same record (the 'conflated record') the theory may be used but the variances are inflated. In practice, we can tolerate a low level of the occurrence of either of these problems [4.2.2]. Nevertheless, when recording pottery archaeologists should strive to achieve the 'one measurable sherd family per record' rule.

5.2 Definitions of fabrics, forms and assemblages

Throughout this work, there has been a tension between the need to aggregate data to make datasets acceptable for statistical purposes, and the need to maintain a fine enough level of detail for useful archaeological interpretation. In general, it is likely that some grouping of both fabrics and forms, as defined by conventional archaeological methods, will be needed before statistical analysis can be undertaken. It is probably better, if it is possible, to form preliminary groupings on archaeological criteria before starting the statistical analysis.

In a chronological study it will usually be necessary to incorporate other sorts of evidence, especially stratigraphy and direct dating evidence, such as coins and C14 dates. The combining of different sorts of data in a chronological study is a topic in itself, and is beyond the scope of this project. Nevertheless, we offer some general guidelines, while aware that they may need to be over-ruled in some practical situations:

i) forms should where possible be grouped according to style or decoration, as these are the aspects most likely to reflect chronological change.

ii) it may make sense to group fabrics according to source and, if possible, phases within sources.

If, however, we are looking for spatial (inter-site) differences, we should concentrate on groups of fabrics based on source. Groupings of forms may not be possible unless forms distinctive of sources can be identified.

A search for functional or social differences demands a third approach. A grouping of fabrics according to technological aspects might be more appropriate, e.g., fine and coarse wares, or perhaps a finer division based on the degree of tempering. Forms should be grouped into functional types, e.g., cooking pots, drinking vessels.

This discussion shows that no one typology, of fabrics or of forms, will

serve for all purposes. The recorder is faced with a dilemma—which to use for the basic recording of the pottery? The ultimate uses of the data will not be known at the time of recording, and it seems undesirable to strait-jacket the data by immediately-perceived needs. The answer is to record in as fine a level of detail as is possible within the resources available, and to indicate ways in which types may be grouped for different purposes. The same data can then be analysed in different ways according to the groupings employed.

Just as for fabrics and forms, there is scope for choosing different groupings of assemblages to meet different needs. If chronology is the main concern, grouping contexts into stratigraphic phases will make sense. For inter-site spatial analysis, aggregation to site-groups is an obvious choice, but has a pitfall if sites are not exactly contemporaneous. Different proportions of different fabrics on the sites may then represent chronological as well as spatial differences. Grouping by phase within site may then be a safer option. To look for functional differences, groupings should be based on the supposed 'function' of contexts, though there is a danger of circular argument here, and a finer level of detail may be safer. Social differences may be marked by differences between assemblages at the level of individual buildings or features (e.g., pits or associated groups of pits).

On any site, there is likely to be more than one such need. Pottery should therefore be recorded according to the finest level of stratigraphic detail (usually the context), with indications of which groupings of contexts would be appropriate for particular purposes. It may be desirable to sub-divide extensive layers spatially (e.g., by grid squares), but this should not be seen as an endorsement of 'digging by spits', which can wreck an attempt at ceramic analysis.

To merge assemblages is to run the risk of breaking the rule set out in Section 5.1, if the same vessels are present in two or more of the groups. However, the situations in which the extent of such 'cross-joining' is so great as to cause serious problems seem to be rare [4.4.1].

6. The broader picture

6.1 Application to other classes of find

In principle the theory could be applied to any class of find which occurs in quantity and usually in broken form—to so-called 'bulk finds' such as animal bones and building materials (brick, tile, etc.). Classes which are not usually broken can be studied as categorical data without the need for a transformation.

The question is most often raised in the context of animal bone studies. It is tempting to equate the whole animal with the whole pot and to look for

'animal equivalents' to correspond to eves in the way that the MNI (minimum number of individuals) statistic corresponds to the 'minimum number of vessels' statistic sometimes used as a lower bound for the number of vessels represented. The analogy, we believe, breaks down because the use of eves depends on the implicit assumption that different parts of the same broken pot are not selectively discarded—which is known not to apply to different parts of the same animal. However, an analogy might be possible at the level of the individual bone, which itself is often broken. One could then use the pct to set up a three-way table of bone name, species and context, and use qlla as before.

Building material, especially tile, seems to be a more promising candidate. Tiles are almost always found broken and can readily be "eve'd", either by standardised weight or counting corners. The standardised nature of the product means that there may be some difficulty in establishing sherd families. However, it seems a worthwhile approach to an otherwise rather intractable class of find.

6.2 Comparison of pottery with other classes of find

Use of the pct enables us to compare and combine pottery assemblages with other classes of finds for which counts are the appropriate level of data, e.g., coins or 'small finds', and to perform for example a joint correspondence analysis. It would be very interesting, for example, to try to integrate pottery into the work done on medieval small finds from Winchester (Biddle *et al.* 1990).

Acknowledgements

We are very grateful to the following, who have supplied data and much valuable background information:

Aldgate and Lime Street—Department of Urban Archaeology, Museum of London;

Chelmsford, Essex—Mr C. Going, Chelmsford Archaeological Unit;

King William IV, Ewell, Surrey—Nonsuch Antiquarian Society;

Lambaesis, Algeria—Ministry of Culture, Algeria;

Leicestershire rural sites—Leicestershire Museums Service;

Mark Brown Wharf, Southwark—Department of Greater London Archaeology, Museum of London;

Silchester Basilica excavations—Professor M.G. Fulford and Dr J. Timby, Department of Archaeology, University of Reading;

Usk, Gwent—Prof. W.M. Manning, Department of Archaeology, University College Cardiff;

Winchester—Professor M. Biddle and Ms K. Barclay, Winchester Research Unit;

Witham, Essex—Mr C. Wallace, Essex County Council Archaeology Service;

Worcester—County Museum Archaeology Service, Hereford and Worcester Archaeology Service.

We are particularly grateful to Professors M. Biddle and M. Fulford for permission to publish the results relating to Winchester and Silchester respectively. The views and interpretations expressed here are, however, entirely our own.

We have benefited from many discussions with colleagues at the Institute of Archaeology and at Computer Applications in Archaeology Conferences, particularly Nick Fieller and Andy Scott.

References

Aitchison, J. 1986: *The Statistical Analysis of Compositional Data* (London, Chapman and Hall).

Biddle, M., Barclay, K. and Orton, C. 1990: The chronological and spatial distribution of the objects. In Biddle, M. (editor), *Object and Economy in Medieval Winchester* (Oxford, Clarendon Press) 42–73.

Bishop, Y.M.M., Fienburg, S.E. and Holland, P.W. 1975: *Discrete Multivariate Analysis* (Cambridge, Massachussetts Institute of Technology Press).

Bloice, B.J. 1971: Note. In Dawson, G.J. Montague Close Part 2. *London Archaeologist* 1, 250–1.

Burgh, R.F. 1959: Ceramic profiles in the Western Mound at Awotovi, Northeastern Arizona. *American Antiquity* 25, 184–202.

Cochran, W.G. 1963: *Sampling Techniques* (New York, Wiley).

Craddock, J.M. and Flood, C.R. 1970: The distribution of the chi-squared statistic in small contingency tables. *Applied Statistics* 19, 173–81.

Duncan, R.J., Hodson, F.R., Orton, C.R., Tyers, P.A. and Vekaria, A. 1988: *Data Analysis for*

Archaeologists: the Institute of Archaeology Programs (London, University College Institute of Archaeology).

Egloff, B.J. 1973: A method for counting ceramic rim sherds. *American Antiquity* 38, 351–3.

Fienberg, S.E. 1977: *The Analysis of Cross-classified Categorical Data* (Cambridge, Massachusetts Institute of Technology Press).

Fisher, R.A. 1935: *The Design of Experiments* (Edinburgh, Oliver and Boyd).

Ford, J.A. 1962: *A Quantitative Method for Deriving Cultural Chronology* (Washington, Pan American Technical Manual 1).

Fulford, M.G. and Hodder, I.R. 1974: A regression analysis of some late Romano-British fine pottery: a case study. *Oxoniensia* 39, 26–33.

Glover, I.C. 1972: *Excavations in Timor* (Canberra, Australian National University PhD Thesis).

Green, C. 1980: Roman Pottery. In Jones, D.M. *Excavations at Billingsgate Buildings 'Triangle' Lower Thames Street, 1974.* (London, London and Middlesex Archaeological Society Special Paper 4).

Greenacre, M.J. 1984: *Theory and Applications of Correspondence Analysis* (London, Academic Press).

Hinton, D.A. 1977: 'Rudely made earthen vessels' of the twelfth to fifteenth centuries AD. In Peacock, D.P.S. (editor), *Ceramics and Early Commerce* (London, Academic Press) 221–38.

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

Hulthén, B. 1974: On choice of element for determination of quantity of pottery. *Norwegian Archaeological Review* 7, 1–5.

Laxton, R.R. forthcoming: *Seriation—the Theory and Practice of Chronological Ordering in Archaeology* (Chichester, Wiley).

Lyne, M.A.B. and Jefferies, R.S. 1979: *The Alice Holt/Farnham Roman Pottery Industry* (London, Council for British Archaeology Research Report 30).

Madsen, T. 1988: Multivariate statistics and archaeology. In Madsen, T. (editor), *Multivariate Archaeology* (Aarhus, Jutland Archaeological Society Publications 21) 7–27.

Marquardt, W. 1978: Advances in archaeological seriation. In Schiffer, M.B. (editor), *Advances in Archaeological Method and Theory 1* (New York, Academic Press) 257–314.

Millett, M. 1979a: An approach to the functional interpretation of pottery. In Millett, M. (editor), *Pottery and the Archaeologist* (London, Institute of Archaeology Occasional Publication 4) 35–48.

Millett, M. 1979b: How much pottery? In Millett, M. (editor), *Pottery and the Archaeologist* (London, Institute of Archaeology Occasional Publication 4) 77–80.

Orton, C.R. 1975: Quantitative pottery studies: some progress, problems and prospects. *Science and Archaeology* 16, 30–5.

Orton, C.R. 1982: Computer simulation experiments to assess the performance of measures of quantity of pottery. *World Archaeology* 14, 1–20.

Orton, C.R. and Pearce, J.E. 1984: The Pottery. In Thompson, A., Grew, F. and Schofield, J. Excavations at Aldgate, 1974. *Post-medieval Archaeology* 18, 34–68.

Orton, C.R. and Tyers, P.A. 1990: Statistical analysis of ceramic assemblages. *Archeologia e Calcolatori* 1, 81–110.

Orton, C.R. and Tyers, P.A. 1991: A technique for reducing the size of sparse contingency tables. In Lockyear, K. and Rahtz, S.P.Q. (editors), *Computer Applications and Quantitative Methods In Archaeology 1990* (British Archaeological Reports International Series 565) (Oxford, Tempus Reparatum) 121–126.

Orton, C.R., Tyers, P.A. and Vince, A.G. forthcoming: *Pottery in Archaeology* (Cambridge, Cambridge University Press).

Peacock, D.P.S. (editor) 1977: *Pottery in Early Commerce* (London, Academic Press).

Petrie, W.M.F. 1899: Sequences in prehistoric remains. *Journal of the Royal Anthropological Institute* 29, 295–301.

Redman, C.L. 1979: Description and inference with the late medieval pottery from Qsar es-Seghir, Morocco. *Medieval Ceramics* 3, 63–79.

Shepard, A.O. 1942: *Rio Grande Glaze Paint Ware; a Study Illustrating the Place of Ceramic Technological Analysis in Archaeological Research* (Washington, Carnegie Institute).

Solheim, W.G. 1960: The use of sherd weights and counts in the handling of archaeological data. *Current Anthropology* 1, 325–9.

Spier, L. 1916: An outline for the chronology of Zuni ruins. *Anthropological Papers of the American Museum of Natural History* 18, 207–231.