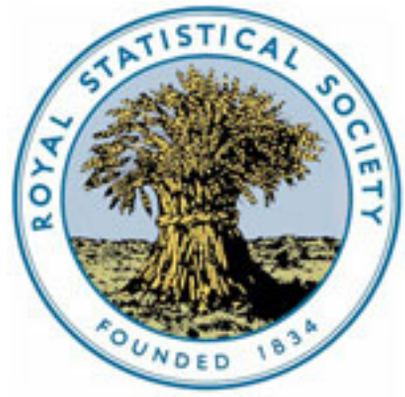

Appl. Statist. (1995)
44, *No.* 4, *pp.* 513–527

Standardization and Transformation in Principal Component Analysis, with Applications to Archaeometry

By M. J. BAXTER†

Nottingham Trent University, UK

[Received April 1994. Revised December 1994]

SUMMARY

Principal component analysis is commonly used in archaeometric applications to identify or display structure in the chemical composition of archaeological artefacts. A recurring topic of debate is whether, and how, data should be transformed and whether, after transformation, standardization should be used. Most discussion has focused on the use of logarithmic transformations. The merits of different approaches are investigated empirically in the paper, using 20 published data sets showing different degrees of structure. The opportunity is taken to examine the merits of the rarely used rank transformation, which has potential attractions when outliers occur or the variables are unusually distributed.

Keywords: Archaeometry; Outliers; Principal component analysis; Rank transformation; Standardization; Transformation

1. Introduction

Principal component analysis involves the linear transformation of correlated variables to pairwise uncorrelated variables and is often used as an exploratory method in which scores based on the first two or three transformed variables are plotted to investigate or display structure in the data (Arnold and Collins (1993) and Jolliffe (1986), pages 64–91).

The method is well known to be scale dependent. Variables are commonly standardized to zero mean and unit variance, and this will usually be necessary if they are measured in different units. If variables are measured in the same units, however, standardization arguably amounts to an arbitrary choice of measurement units (Jolliffe (1986), p. 20). Another choice that needs to be made in practice is whether or not to transform variables before a principal components analysis; for example, data measured on a ratio scale might be logarithmically transformed (Arnold and Collins (1993), p. 384). After a transformation data may or may not subsequently be standardized.

There is no simple prescription for deciding what to do in any particular

†*Address for correspondence*: Department of Mathematics, Statistics and Operational Research, Nottingham Trent University, Clifton Campus, Nottingham, NG11 8NS, UK.
E-mail: mat3baxtemj@ntu.ac.uk

 0035–9254/95/44513

circumstance. The present paper is an empirical investigation of the consequences of different approaches to the issues of transformation and standardization, in the particular context of the analysis of chemical compositions of archaeological artefacts. This is a common problem in archaeometry and is discussed in Section 2. Briefly, the problem is that logarithmic transformation of oxide or element concentrations is often advocated, but not universally accepted. Standardization of data, with or without transformation, is the norm but has been questioned. One aim of the paper is to investigate whether, and to what extent, treatment of the data matters in terms of substantive interpretation.

A particular problem with data of the kind under discussion is that oxides or elements may be present below the minimum level of detection. These can be recorded as 0 values but the use of a logarithmic transformation is then problematic. The distribution of some of the variables in a composition may also be very unusual and not ideally suited to principal components analysis as usually applied. A possible way round these problems is to transform the data to ranks before analysis. This approach has been little used in archaeometry, if at all, and a second aim of the paper is to explore its merits.

2. Principal Component Analysis in Archaeometry

In a typical problem p measurements are made for each of n artefacts such as shards of pottery or fragments of glass. The measurements are of the concentrations of oxides or trace elements in the artefacts, with the former usually recorded as a percentage and the latter as parts per million. With a mixture of the two types of measurement it is possible to convert to the same scale if needed. If only oxides are used p is typically between 7 and 12; with both oxides and trace elements used p can be over 30, though 15–20 or so is more common in published data.

Additional classificatory information is often available: identifying stylistic types or site of origin, for example. Alternatively, some other method of multivariate analysis, such as cluster analysis, may be used to suggest chemically distinct groups within the data for which an archaeological or technological interpretation is then sought.

Principal component analysis is used

(a) to investigate chemical compositional structure in the data, in the absence of other information,
(b) to investigate whether or not archaeological types cluster together on a component plot based on the chemical data only and
(c) to display chemically distinct groups suggested by cluster analysis.

The majority of the numerous applications of principal component analysis in archaeometry fall into one of these categories (Baxter (1994a), chapter 4).

The analyses discussed in this paper are of ceramic or glass compositions. Some of the sources of the data report them in a form that is fully compositional in the sense of Aitchison (1983, 1986). The dominant oxide is invariably silica (SiO_2) with a value typically in excess of 60% by weight. To avoid problems in analysis, of the kind identified by Aitchison with such data, silica has been omitted from the analyses undertaken here. This approach has been described as 'naïve' by

Aitchison but there is evidence that it works well with data of the kind under discussion (Baxter, 1992).

The logarithmic transformation, to base 10, of data before a principal component or other analysis is common. One reason for this is a belief that, within the raw materials of manufacture, elements have a natural log-normal distribution, and that normality of the data is desirable. A second reason is that a logarithmic transformation tends to stabilize the variance of the variables and would thus give them approximately equal weight in an unstandardized principal component analysis. A third reason sometimes cited is that, empirically, the use of a logarithmic transformation leads to more satisfactory results. In advocating the use of logarithmic transformation early work at the Brookhaven Laboratory has been influential (Bieber *et al.*, 1976; Harbottle, 1976). Not everyone has been persuaded and a discussion of the issues involved and debate, with references, was given by Pollard (1986).

Ceramic manufacture involves the addition of a temper to the clays used, and glass manufacture often involves the reuse of old glass (cullet) with the raw materials. From a statistical point of view this means that the distribution of some elements in the finished product will be a mixture of distributions and thus not necessarily (log-)normal, even if natural elements are (log-)normally distributed.

In practice, data have almost invariably been standardized before analysis, even if a logarithmic transformation has been applied. From this point of view the variance stabilizing tendency of the logarithmic transformation is irrelevant. The justification for standardization is that, even after transformation, the variances of variables differ sufficiently for those with the smaller variances to be effectively ignored in an analysis. Bishop and Neff (1989), pages 63–64, criticized the unthinking use of standardization, which they attributed to the fact that it is the default in the software that is typically used, and gave examples where an analysis of transformed but unstandardized data was more fruitful in revealing archaeologically meaningful variation in the data.

3. Rank Transformation

Where the data are unpleasantly distributed—for example, very highly skewed—a component plot can be difficult to interpret. Problems also arise with outliers in the data, or with small, compositionally distinct groups. These can determine the scale of the usual component plot, making any appreciation of structure in the rest of the data difficult. It is possible to repeat analyses after omitting the offending specimens, but this may cause unease among researchers who regard it as an unjustified 'manipulation' of the data. Another possibility is to omit such specimens from the plot.

Wright (1989) suggested that in such circumstances an analysis of rank-transformed data may be helpful. Kamminga and Wright (1988) provided an example of such an approach, applied to anthropological data; I am unaware of any applications to archaeometric data.

The possibility of using rank-transformed data is not discussed in two of the major texts on principal component analysis (Jackson, 1991; Jolliffe, 1986) or in more general texts such as Mardia *et al.* (1979), Seber (1984) and Krzanowski (1988). Jambu (1991), p. 126, noted, but did not exploit, the possibility. Bacon-Shone

(1992) applied various ranking procedures to fully compositional data, to overcome problems in applying the methodology of Aitchison (1983, 1986) to data containing 0 values, but did not subsequently use principal component analysis.

The approach used in this paper is to rank the values for each variable before applying principal component analysis. Standardization is implicit since, in the absence of ties, all variables then have equal variance. In effect, analysis is based on the correlation matrix of the data, using Spearman's rank correlation coefficient rather than the more usual product–moment correlation, with subsequent plotting also using the ranked values.

Clearly a rank transformation applied to bimodal or multimodal univariate data will completely eliminate clear group differences. Whether the same is true for multivariate data in seven or more dimensions remains to be seen. An analysis based on ranks can be expected to be more robust to outliers than the other methods discussed. Its main potential merit, however, would seem to be for the analysis of data where the element distributions are unusual and there are 0s in the data, for which a principal component analysis of untransformed data may be unsatisfactory and a logarithmic transformation is problematic.

4. Data

Summary details of 20 published data sets, used in the present investigation, are given in Table 1. The $n \times p$ data matrices X have n ranging from 19 to 146 and p from 7 to 22. The data are ordered according to the similarity of analyses based

TABLE 1
Data sets used in analyses—summary details†

Identifier	*Source*	*n*	*p*	*O*	*S*	*M*
A	Matthers *et al.* (1983)	58	15	N	Y	C
B	Wisseman *et al.* (1987)	36	20	N	N	C
C	Cox and Gillies (1986)	27	12	N	Y	G
D	Mirti *et al.* (1990)	45	12	N	Y	C
E	Krywonos *et al.* (1982)	146	11	N	Y	C
F	Baxter *et al.* (1990)	47	11	N	N?	G
G	Foy (1985) (a)	27	9	Y	Y	G
H	Topping and MacKenzie (1988)	36	8	N	Y?	C
I	Tubb *et al.* (1980)	48	9	N	Y	C
J	Baxter (1994a)	105	22	N	Y?	G
K	Christie *et al.* (1979)	19	8	Y?	Y	G
L	Pollard and Hatcher (1986)	133	8	Y	Y	C
M	Cracknell (1982)	56	7	Y?	N	C
N	Wolff *et al.* (1986)	58	15	Y	Y	C
O	Sanderson *et al.* (a) (1984)	45	8	Y	Y	G
P	Calamioutou *et al.* (1984)	47	11	Y	N	C
Q	Foy (1985) (b)	46	9	Y	Y?	G
R	Velde and Gendron (1980)	58	8	Y	N	G
S	Alvey and Laxton (1978)	54	8	Y	Y?	C
T	Sanderson *et al.* (b) (1984)	35	8	Y	N	G

†n and p indicate the number of specimens and variables; O, S and M indicate outliers, structure and material respectively with Y and N corresponding to 'yes' and 'no' and '?' indicating possible doubt. In the 'material' column C and G indicate ceramic and glass.

on the standardized raw and rank-transformed data (see Section 5) with A, B, C, ... the most similar.

An indication is given of whether previous analyses, typically cluster analysis and/or principal component analysis, have found structure in the data. For some data sets, structure, in the sense that there are clear disjoint chemically distinct groups, is evident from a plot of the first two components. For other data sets, grouping is less evident but labelling points according to type, for example, will show that different types occupy distinct regions of the graph.

The appearance of structure is not, in itself, an indication of a successful analysis, since it may occur for purely mathematical reasons and need have no substantive interpretation (Baxter *et al.*, 1990). Different analyses may also indicate different, but equally valid, structure. This is considered at greater length in Section 7.

The presence or absence of outliers in the data sets is also indicated. These were identified, subjectively, from component plots based on the standardized raw data. The identification of outliers, and group structure, in this way is subjective; analyses where there is doubt are indicated in Table 1. Outliers identified in this way were usually evident from other methods, with analyses of unstandardized logarithmically transformed data sometimes revealing additional unusual data (e.g. Fig. 2, later).

For data sets C, F, G, K, L, O, Q, R, S and T the concentration of silica is listed in the original source but not used here for the reasons stated earlier. Data sets C, G, J, K and Q contained 0s. To allow comparisons between the methods investigated 0s were either set to a number just less than the next smallest value for the variable or, with many 0s, the variable was not used in the analysis. Had the data been used as given the methods based on logarithmic transformation would not be applicable, whereas the modifications have little or no effect on the analyses with untransformed data.

5. Analysis

Four methods of analysis were applied to each data set with principal component analysis being based on

(a) standardized raw data,
(b) rank-transformed data,
(c) standardized logarithmically transformed data and
(d) unstandardized logarithmically transformed data.

In the first instance, plots based on the leading two components were compared for each method and data set, labelling points to facilitate the assessment of similarities and differences. The coefficients of the leading components for each method and data set were also compared.

In nearly all analyses the leading two components account for 50% or more of the variation in the data (on the scale used), with four components usually sufficient to account for at least 80% of the variation. Methods (a)–(c) are generally similar in terms of variation accounted for, with method (d) typically returning higher values. This last phenomenon arises because the data are unstandardized and the transformed variables still have markedly different variances in many cases.

The procedures described so far are subjective; more objectively, a coefficient derived by Sibson (1978) may be used to compare configurations. Let Z_i be an $n \times k$ matrix containing the co-ordinates of a configuration of points in k-dimensional space. Two such configurations may be rotated, reflected and expanded or contracted uniformly to obtain as close a match as possible. A symmetric measure of the similarity between two configurations $i = 1, 2$ is

$$\gamma = 1 - \{\mathrm{tr}(Z_2^{\mathrm{T}} Z_1 Z_1^{\mathrm{T}} Z_2)^{1/2}\}^2 / \mathrm{tr}(Z_1^{\mathrm{T}} Z_1)\,\mathrm{tr}(Z_2^{\mathrm{T}} Z_2)$$

where tr( ) is the trace operator, $A^{1/2}$ is the unique positive semidefinite symmetric square root of a positive semidefinite symmetric matrix A and $\gamma = 0$ for identical configurations.

In Table 2 the measure 100γ is used to assess the similarity of the component scores for the analyses undertaken for each data set. Results for $k = 4$ and $k = 2$ are given. The data sets are ordered according to the similarity of methods (a) and (b) in two dimensions.

In the rest of this section some generalizations based on Table 2 are made. More specific and detailed discussion of some of the results is given in Section 6. Past experience (Baxter, 1992) suggests that a value of γ less than 0.10 (i.e. $100\gamma < 10$)

TABLE 2
Four-dimensional and two-dimensional comparisons of component scores using Sibson's coefficient for all pairwise method combinations†

Identifier	*Scores for four dimensions and the following method pairs:*						*Scores for two dimensions and the following method pairs:*					
	12	*13*	*23*	*14*	*24*	*34*	*12*	*13*	*23*	*14*	*24*	*34*
A	14	4	8	12	18	8	2	1	1	44	44	43
B	10	9	17	37	44	44	3	2	3	29	29	27
C	6	4	5	24	20	16	3	3	4	27	21	17
D	12	6	13	23	22	15	3	5	5	12	11	6
E	6	2	3	5	6	3	4	1	2	4	6	2
F	4	1	2	27	27	26	4	1	2	14	18	14
G	9	8	14	20	28	8	4	5	7	25	28	15
H	8	4	10	13	15	9	4	9	10	22	24	34
I	11	11	20	33	34	21	4	10	10	26	27	16
J	13	5	8	31	31	27	5	2	5	23	24	17
K	14	6	13	46	44	34	7	9	13	49	52	30
L	17	12	8	38	28	30	10	9	7	27	19	13
M	16	8	12	18	22	12	12	7	8	16	16	8
N	17	7	9	35	29	25	14	5	6	26	24	21
O	17	8	20	46	47	35	14	11	17	76	75	60
P	20	17	18	28	26	19	17	12	15	23	23	11
Q	12	6	9	18	21	10	19	12	5	63	61	54
R	18	10	9	46	35	34	26	17	7	63	47	45
S	28	6	24	44	52	34	36	5	28	89	85	88
T	28	15	30	50	55	29	59	33	34	50	71	48

†Entries show the value of 100γ, where γ is Sibson's coefficient. The closer to 0 this is the better the agreement between two configurations of points. Methods 1–4 are based on standardized raw, rank-transformed, standardized logged and unstandardized logged data.

implies good agreement between configurations. On this basis inspection of Table 2 suggests the following.

(a) Standardizing the raw and logarithmically transformed data leads, in most cases, to very similar results with the most obvious exception—data set T—containing very clear outliers.
(b) Except for data set G the value of γ for comparing the standardized raw and rank-transformed data neatly divides the data sets into those with and those without outliers. In the absence of obvious outliers there is generally good agreement between the methods; with outliers agreement can be poor.
(c) It follows from (a) and (b), and is evident from Table 2, that the rank and standardized logarithmically transformed data produce similar results if there are no outliers. In the presence of outliers the ranked data usually produce results nearer to those for the logarithmically transformed than for the raw data. This is because the logarithmic transform tends to downweight the influence of outliers.
(d) In general, differences between the three methods appear to depend on the presence or absence of unusual values in the data, not on the presence or absence of structure.
(e) There are in general, and as might be expected, clear differences between results for the first three methods and for the unstandardized logarithmically transformed data. These are discussed in more detail in Sections 6 and 7.
(f) Inspection of values of γ for $k = 4$ tends to confirm the foregoing observations. For the first three methods agreement is actually better for data sets Q–T. For the fourth method for $k = 4$ agreement with the other three methods is better for data sets A, C, G, H, O and Q–T.

6. Examples

The results for data sets D and P are given in detail by Baxter (1994b). The examples given here have been selected to display particular aspects of the analysis.

6.1. *Example 1—Data Set I, Romano-British Pottery*

Data set I relates to Romano-British pottery from five kiln sites in three separate localities, the nine variables being measured in percentage terms. The analysis of the unstandardized log-transformed data in Fig. 1 suggests three clear groups corresponding to the three localities Wales (bottom left), Gloucester (top left) and the New Forest (right). This is similar to the principal co-ordinate analysis in the original paper. The same three groups are evident in the plot based on ranked data (top right, bottom right, left), with a single point, number 12, that seems not to belong to a group. The Wales group on this plot divides into two subgroups which do not correspond to distinct kiln sites.

A knowledge of localities allows the groups to be seen in the plot based on the standardized raw data, though it is slightly less clear cut than the other two. In the plot for the correlation matrix of the logged data the New Forest group separates out to the right. To the left, no clear groups can be seen, and there is some mixing of specimens from the other two localities.

In this example the results of analysis of the ranked data or the unstandardized

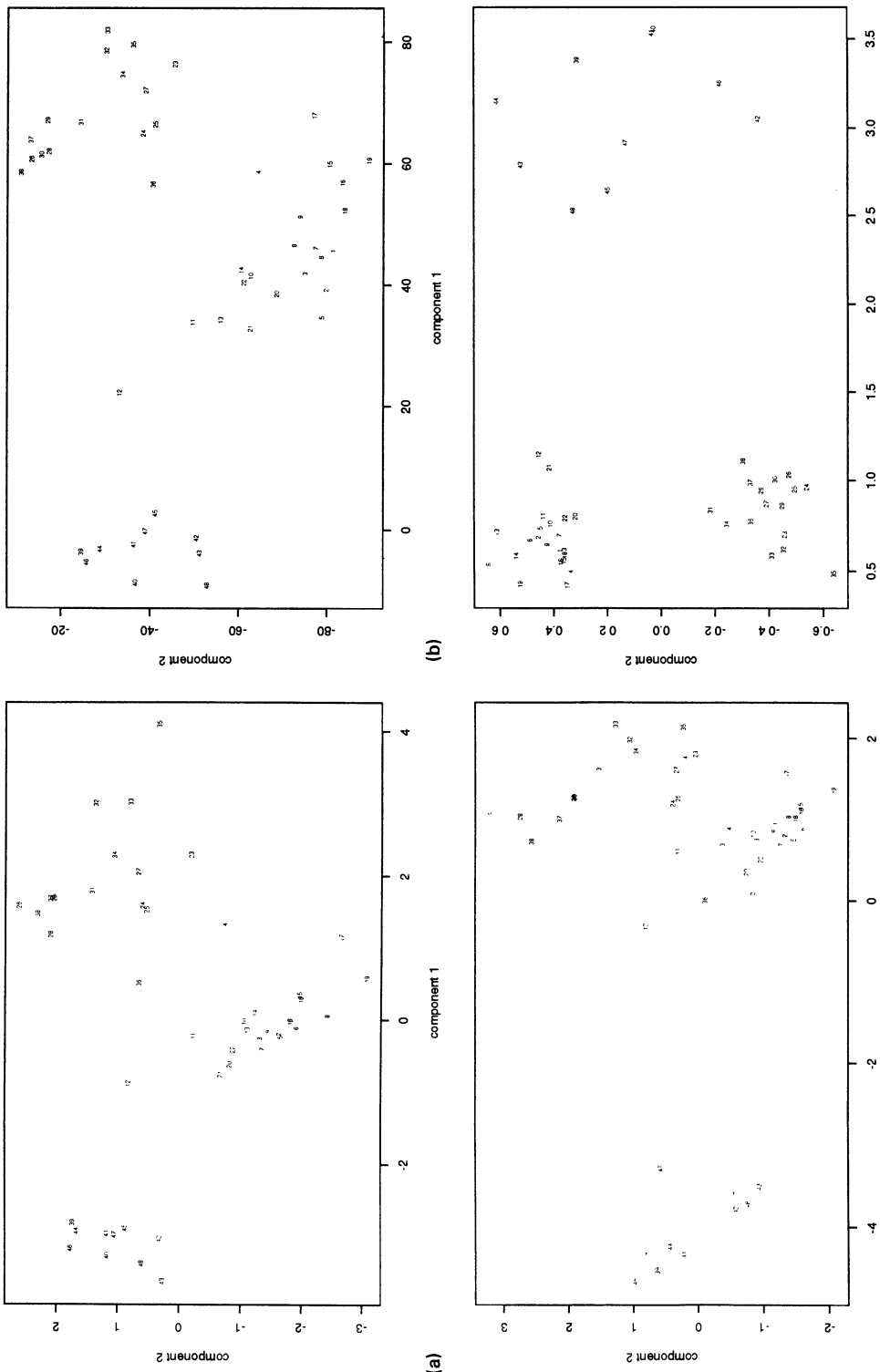

Fig. 1. Principal component analysis for data set I—Romano-British pottery; (a) analysis of correlation matrix; (b) analysis of correlation matrix—ranked data; (c) analysis of correlation matrix—logged data; (d) analysis of covariance matrix—logged data

logged data are most satisfactory in archaeological terms, with the former suggesting potential extra structure in the data not evident from the other approaches. Analyses of the ranked data for some of the other data sets resulted in a similar phenomenon. The 'reality' or otherwise of such additional subgroupings would usually be a matter for substantive judgment, and in the present example the subgroups do not have a coherent archaeological interpretation. Analysis of the standardized logged data is least satisfactory in substantive terms.

6.2. *Example 2—Data Set O, Eighth- and Ninth-century Funnel Beakers*

Fig. 2 shows the component plots for data set O. There are very clear outliers in the data set, but those specimens identified as such from the plots depend on the particular method used. Observation 7 is identified as an outlier in the analysis of the standardized raw and logged data, but not by the analysis of the unstandardized logged data. The analyses of the logged data suggest that observations 1, 23 and possibly 25 are unusual. Observations 1 and 7 are slightly apart from the rest in the analysis of the ranked data, but not in an extreme fashion.

In this analysis the observations suggested as outliers clearly depend on the scale of measurement used, and whether or not standardization is used. For three of the plots the scale of the first component is largely determined by the outliers, and this makes it difficult to perceive any structure in the bulk of the data. The exception to this, the analysis of the ranked data, suggests, perhaps, three groups, separated out at about 10 and 45 on the first component.

The data relate to compositions of eighth- and ninth-century funnel beakers from three geographically distinct regions. In the publication from which the data were taken, Sanderson *et al.* (1984) concluded that there were clear chemical differences between the groups. They used, but do not report in detail, discriminant analysis. In the plot based on the ranked data all the points in the leftmost group come from the same region. Other points from this region occur in the rest of the figure. Points from the second region tend to occur centrally, and points from the third region to the right, but there is no clear-cut regional grouping in the plot. This can be seen in Fig. 3 which is the same as Fig. 2, but with points labelled by region of origin. If points 1, 7 and 23 are omitted then an analysis of the standardized raw data leads to results that are similar to those for the analysis of the ranked data.

6.3. *Roman Glass Found in Norway*

Fig. 4 shows the results for the small data set K, based on eight oxides measured for 19 specimens of Roman glass found in Norway. Using a principal component analysis of the standardized raw data Christie *et al.* (1979) identified two groups separated out at about 0 on the first component.

These groups are also evident on the plot for the standardized logged data and, perhaps less clearly, in the analysis of the ranked data. The first of these two plots suggests two outliers, 9 and 17. The first of these outliers is the dominant feature of the analysis of the unstandardized logged data; the offending point has a (relatively) very low value for one of the oxides that is emphasized by the logarithmic transformation. The picture obtained by this last analysis is different from the other three, even if the effect of the outlier is discounted.

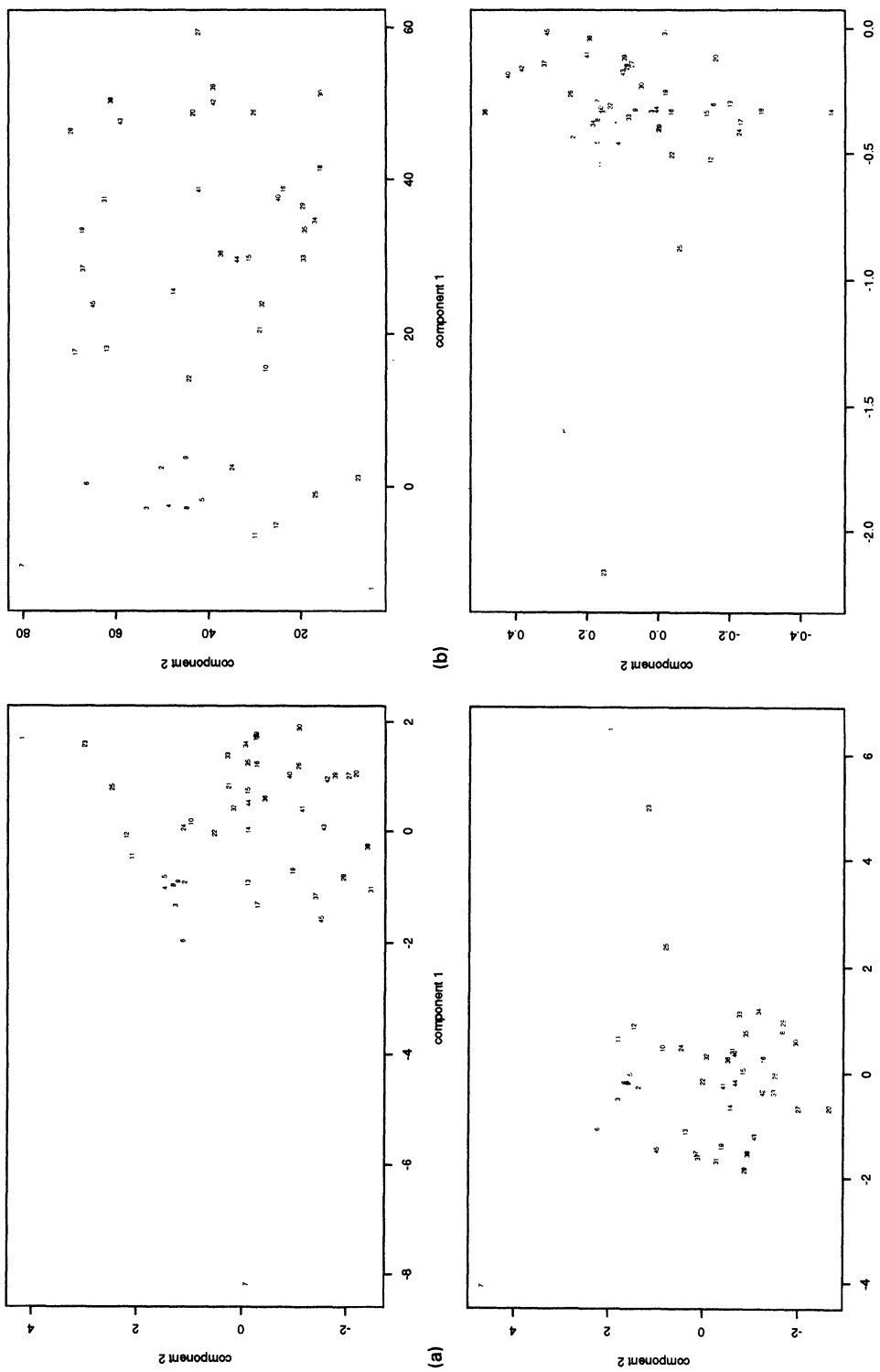

Fig. 2. Principal component analyses for data set O—eighth- and ninth-century funnel beakers: (a) analysis of correlation matrix; (b) analysis of correlation matrix—ranked data; (c) analysis of correlation matrix—logged data; (d) analysis of covariance matrix—logged data

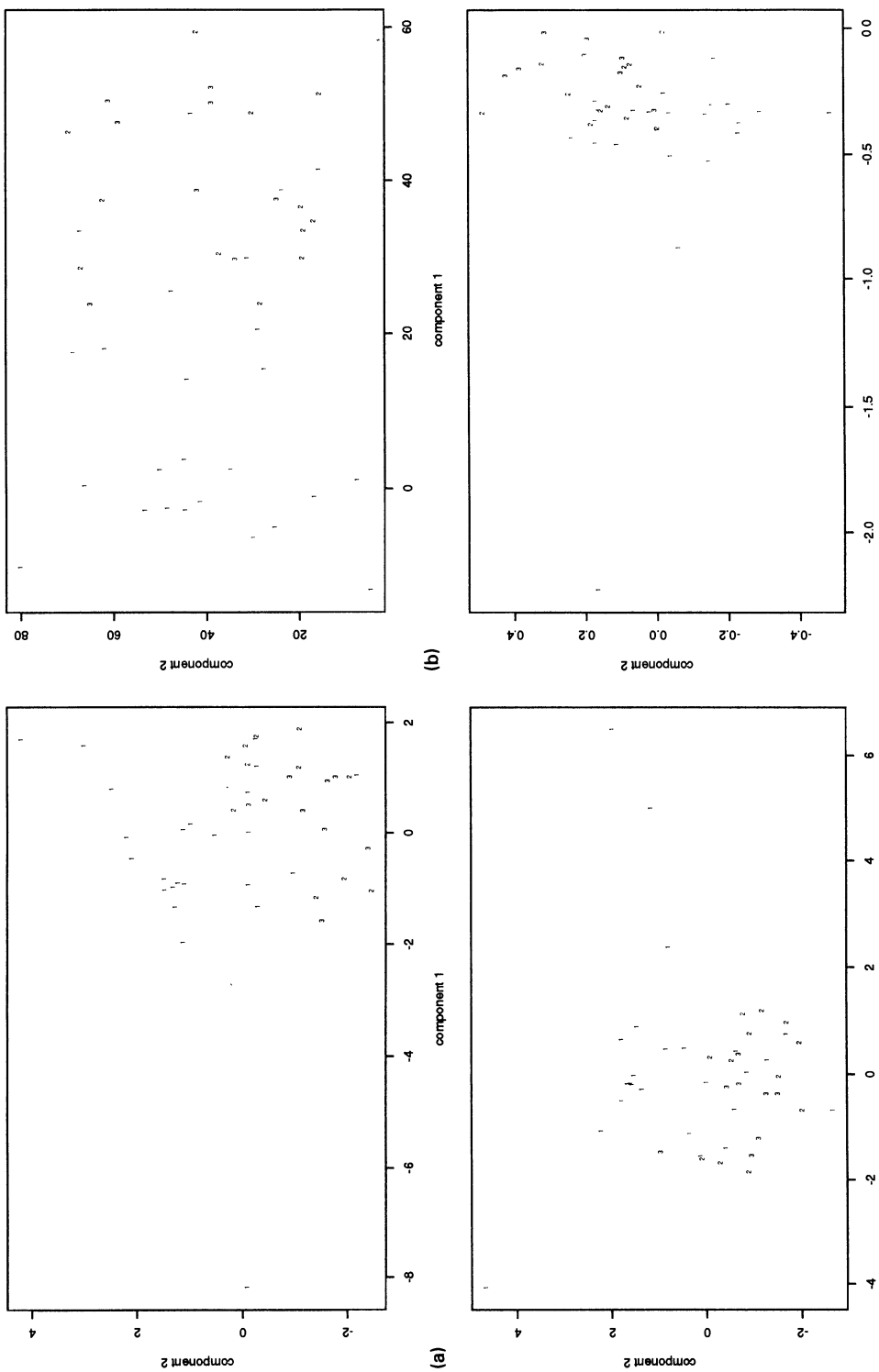

Fig. 3. Principal component analyses for data set O – eighth- and ninth-century funnel beakers labelled by region of origin: (a) analysis of correlation matrix; (b) analysis of correlation matrix – ranked data; (c) analysis of correlation matrix – logged data; (d) analysis of covariance matrix – logged data

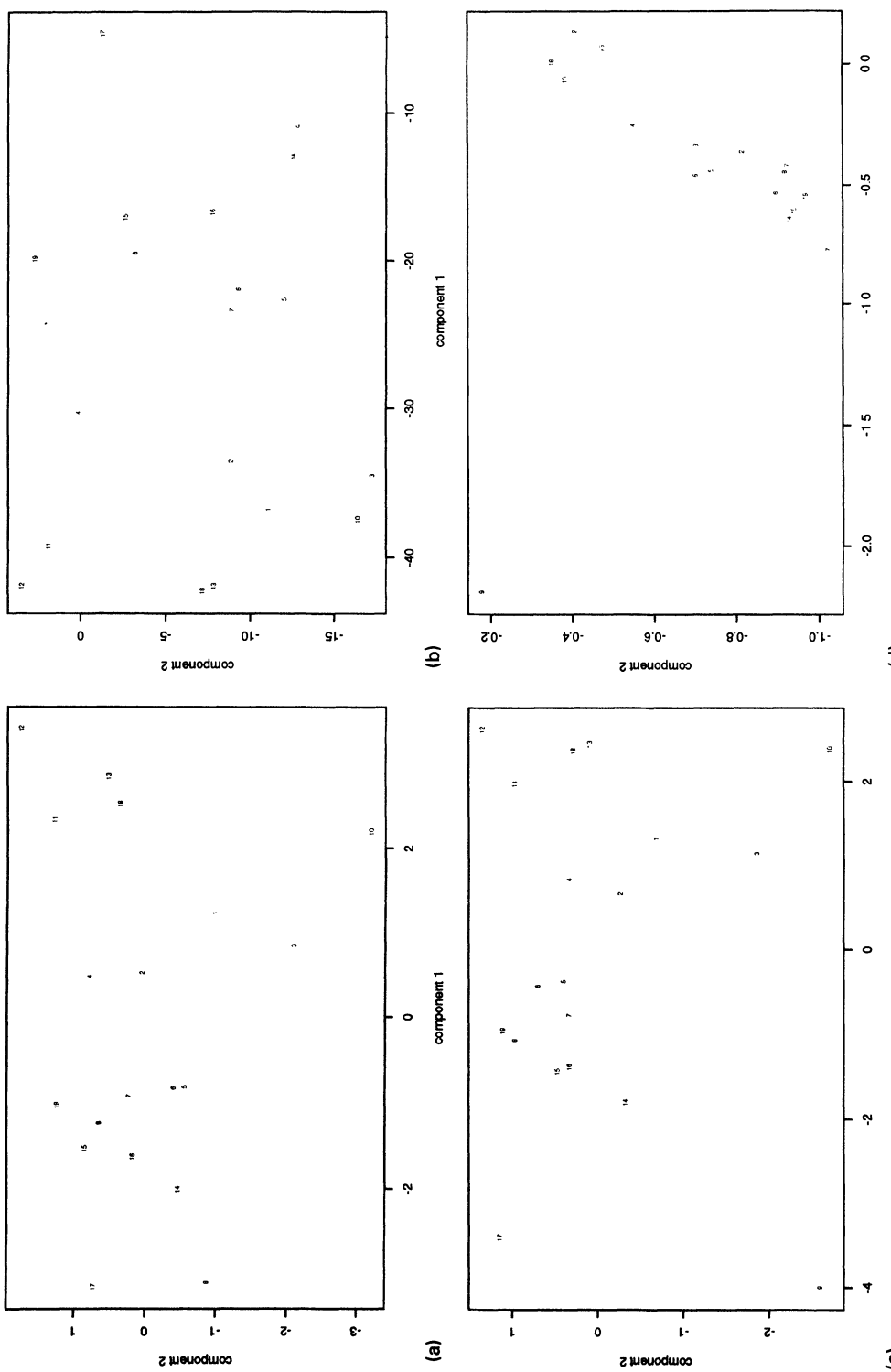

Fig. 4. Principal component analyses for data set K—Roman glass from Norway: (a) analysis of correlation matrix; (b) analysis of correlation matrix—ranked data; (c) analysis of correlation matrix—logged data; (d) analysis of covariance matrix—logged data

7. Discussion

The analyses reported in this paper were motivated by a wish to investigate different approaches to issues concerning the transformation and standardization of data, in the context of the analysis of artefact compositions in archaeometric applications using principal component analysis. The empirical study undertaken here, supplemented by the detailed analysis of data sets D and P reported by Baxter (1994b), suggests the following.

(a) The results obtained by analysing the standardized raw and log-transformed data will frequently be very similar in terms of the configurations obtained, particularly if there are no outliers in the data. These are the two methods that are most often used in practice. Where the results do differ, in the form of the clarity of a component plot, there is no obvious reason for preferring one approach to the other.

(b) When measured in formal terms, the difference between results using the unstandardized logged data and the other approaches can be large. This may also manifest itself in component plots that appear markedly different. In terms of substantive conclusions, however, the examples of the previous section suggest that similar conclusions may often be reached. Neff and Glascock (1992) made this point, noting that 'standardization and log-transformation lead to equivalent results' in many cases. This is not, however, inevitable so both plots may be worth inspection.

(c) Different approaches can lead to the identification of different outliers in a data set, and these in turn complicate the visual and formal comparison of results from different analyses. When outliers are identified and subsequently omitted from an analysis the approaches used here will often lead to similar substantive conclusions.

(d) How best to identify and deal with outliers is a matter for debate, and it is suggested here that analysis of ranked data may have some merit. In the absence of outliers results are usually similar to those for analyses of standardized data, whether logarithmically transformed or not. In the presence of outliers the plots based on ranked data may be easier to read and can suggest additional structure. Unfortunately this is not invariably so and structure, evident on plots from the other methods, is sometimes obscured when using ranked data. Analyses of ranked data downweight the effect of outliers but may also conceal their existence and this is not necessarily desirable.

(e) Discussion so far begs, to some extent, the question of how exactly analyses should be compared. Tangri and Wright (1993) argued forcefully that the appearance of structure in a component plot does not in itself imply that it is superior to a plot without structure, since it may be a purely mathematical artefact having no substantive meaning. The issue raised is problematic. The merits of different analyses need to be judged in context. If all analyses lead to similar and sensible substantive conclusions there is usually not a problem. It is also quite possible for different approaches to lead to different, and complementary, substantive conclusions. For example, in glass studies, one approach may identify groups based on the

major oxides, reflecting the technology of manufacture, whereas another approach may give more weight to minor oxides and trace elements that influence glass colour. In some of the examples reported in Section 6, compositionally distinct subgroups are sometimes suggested that have no obviously distinct archaeological interpretation. There is, in other words, no simple way of determining the relative merits of different approaches to transformation and standardization. In particular analyses substantive considerations must take precedence and no single approach invariably leads to 'superior' results.

(f) In summary, no single approach can be recommended for data of the kind under consideration. Analyses of the standardized raw and log-transformed data and ranked data will often give similar results. The last method is more robust to outliers and gives sensible results, which may suggest structure additional to that shown by other methods. Analysis of unstandardized logged data can lead to apparently very different results which may, nevertheless, have a similar substantive interpretation. Such data should almost certainly be analysed in several ways, when extra insight should be gained about the data structure if the conclusions to be drawn from the analyses differ.

Acknowledgements

I am grateful to Ian Jolliffe and Richard Wright for information on the use (or its lack) of ranked data for principal component analysis. Initial analyses on some of the data sets were undertaken by Tracey Viles as part of an undergraduate dissertation under the direction of the author. An Editor and referee are thanked for suggestions leading to an improved presentation of the paper.

References

Aitchison, J. A. (1983) The statistical analysis of compositional data (with discussion). *J. R. Statist. Soc.* B, **44**, 139–177.

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

Alvey, R. C. and Laxton, R. R. (1978) A note on the chemical analysis of some Nottingham clay tobacco pipes. *Archaeometry*, **20**, 189–196.

Arnold, G. M. and Collins, A. J. (1993) Interpretation of transformed axes in multivariate analysis. *Appl. Statist.*, **42**, 381–400.

Bacon-Shone, J. (1992) Ranking methods for compositional data. *Appl. Statist.*, **41**, 533–537.

Baxter, M. J. (1992) Statistical analysis of chemical compositional data and the comparison of analyses. *Archaeometry*, **34**, 267–277.

——— (1994a) *Exploratory Multivariate Analysis in Archaeology*. Edinburgh: Edinburgh University Press.

——— (1994b) Principal component analysis in archaeology. *Arch. Calc.*, **5**, 23–38.

Baxter, M. J., Cool, H. E. M. and Heyworth, M. P. (1990) Principal component and correspondence analysis of compositional data: some similarities. *J. Appl. Statist.*, **17**, 229–235.

Bieber, Jr, A. M., Brooks, D. W., Harbottle, G. and Sayre, E. V. (1976) Application of multivariate techniques to analytical data on Aegean ceramics. *Archaeometry*, **18**, 59–74.

Bishop, R. L. and Neff, H. (1989) Compositional data analysis in archaeology. In *Archaeological Chemistry IV* (ed. R. O. Allen), pp. 57–86. Washington DC: American Chemical Society.

Calamiotou, M., Fillipakis, S. E., Jones, R. E. and Kassals, D. (1984) X-ray and spectrographic

analyses of terracotta figurines from Myrina: an attempt to characterise workshops. *J. Arch. Sci.*, **11**, 103–117.
Christie, O. H. J., Brenna, J. A. and Straume, E. (1979) Multivariate classification of Roman glasses found in Norway. *Archaeometry*, **21**, 233–241.
Cox, G. A. and Gillies, K. J. S. (1986) The X-ray fluorescence analysis of medieval durable blue soda glass from York Minster. *Archaeometry*, **28**, 57–68.
Cracknell, S. (1982) An analysis of pottery from Elgin and north-east Scotland. *Med. Ceram.*, **6**, 51–65.
Foy, D. (1985) Essai de typologie des verres medievaux d'après les fouilles provençales et languedociennes. *J. Glass Stud.*, **27**, 18–71.
Harbottle, G. (1976) Activation analysis in archaeology. *Radiochemistry*, **3**, 33–72.
Jackson, J. E. (1991) *A User's Guide to Principal Components*. New York: Wiley.
Jambu, M. (1991) *Exploratory and Multivariate Data Analysis*. New York: Academic Press.
Jolliffe, I. T. (1986) *Principal Component Analysis*. New York: Springer.
Kamminga, J. and Wright, R. V. S. (1988) The upper cave at Zhoukoudian and the origins of the mongoloids. *J. Hum. Evoln*, **17**, 739–767.
Krywonos, W., Newton, G. W. A., Robinson, V. J. and Riley, J. A. (1980) Neutron activation analysis of Roman coarse-ware from Cyrenaica. *Archaeometry*, **22**, 189–196.
Krzanowski, W. J. (1988) *Principles of Multivariate Analysis*. Oxford: Oxford Science.
Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979) *Multivariate Analysis*. London: Academic Press.
Matthers, J., Liddy, D. J., Newton, G. W. A., Robinson, V. J. and Al-Tawel, H. (1983) Black-on-red ware in the Levant: a neutron activation analysis study. *J. Arch. Sci.*, **10**, 369–382.
Mirti, P., Zelano, V., Aruga, R., Ferrara, E. and Appolonia, L. (1990) Roman pottery from Augusta Praetoria (Aosta, Italy): a provenance study. *Archaeometry*, **32**, 163–175.
Neff, H. and Glascock, M. D. (1992) Neutron activation analysis of formative period pottery from Huaca Prieta, Peru. Unpublished. University of Missouri, Columbia.
Pollard, A. M. (1986) Data analysis. In *Greek and Cypriot Pottery: a Review of Scientific Studies* (ed. R. R. Jones), pp. 56–83. Athens: British School at Athens.
Pollard, A. M. and Hatcher, H. (1986) The chemical analysis of oriental ceramic body compositions: Part 2—greenwares. *J. Arch. Sci.*, **13**, 261–287.
Sanderson, D. C. W., Hunter, J. R. and Warren, S. E. (1984) Energy dispersive X-ray fluorescence analysis of first millenium AD glass from Britain. *J. Arch. Sci.*, **11**, 53–69.
Seber, G. A. F. (1984) *Multivariate Observations*. New York: Wiley.
Sibson, R. (1978) Studies in the robustness of multidimensional scaling. *J. R. Statist. Soc.* B, **40**, 234–238.
Tangri, D. and Wright, R. V. S. (1993) Multivariate analysis of compositional data: applied comparisons favour standard principal components analysis over Aitchison's loglinear contrast method. *Archaeometry*, **35**, 103–112.
Topping, P. G. and MacKenzie, A. B. (1988) A test of the use of neutron activation analysis for clay source characterization. *Archaeometry*, **30**, 92–101.
Tubb, A., Parker, A. J. and Nickless, G. (1980) The analysis of Romano British pottery by atomic absorption spectrophotometry. *Archaeometry*, **22**, 153–171.
Velde, B. and Gendron, C. (1980) Chemical composition of some Gallo-Roman glass fragments from central western France. *Archaeometry*, **22**, 183–187.
Wisseman, S. V., Hopke, P. K. and Schindler-Kaudelka, E. (1987) Multi-elemental and multivariate analysis of Italian terra-sigillata in the World Heritage Museum, University of Illinois at Urbana-Champaign. *Archaeomaterials*, **1**, 101–117.
Wolff, S. R., Liddy, D. J., Newton, G. W. A., Robinson, V. J. and Smith, R. J. (1986) Classical and Hellenistic black glaze ware in the Mediterranean: a study by epithermal neutron activation analysis. *J. Arch. Sci.*, **13**, 245–259.
Wright, R. V. S. (1989) *Doing Multivariate Analysis and Prehistory*. Sydney: University of Sydney Press.