Pergamon

*Journal of Structural Geology*, Vol. 19, No. 12, pp. 1527 to 1536, 1997

PII: S0191-8141(97)00077-1
0191-8141/97 $17.00 + 0.00

# Confidence regions for fabric shape diagrams

TREVOR J. RINGROSE

Applied Mathematics and OR Group, Cranfield University, RMCS Shrivenham, Swindon SN6 8LA, UK

and

DOUGLAS I. BENN

Department of Geography, University of Aberdeen, Aberdeen AB9 2UF, U.K.

(*Received* 2 *December* 1996; *accepted in revised form* 4 *August* 1997)

**Abstract**—Fabric shape is often quantified using the three eigenvalues from the 'orientation tensor' method applied to a sample of directions. Several studies have used eigenvalues plotted on fabric shape diagrams to distinguish sedimentary facies or strain histories. However, such studies seldom consider how well the sample eigenvalues represent the true fabric shape. In this paper, we use 'bootstrapping' techniques to define confidence regions for sample eigenvalues, and show that sample and population eigenvalues may differ substantially. Confidence regions are often very large for small sample sizes, and we recommend that sample sizes should be at least 50. © 1997 Elsevier Science Ltd.

## INTRODUCTION

The 'orientation tensor' or 'eigenvalue method' has been widely adopted as a useful way of summarizing three-dimensional orientation data. It is very similar to the well-known multivariate statistical method principal component analysis (PCA) and resolves sets of observations into three mutually-orthogonal eigenvectors, $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$, in which $\mathbf{v}_1$, the principal eigenvector, is parallel to the axis of maximum clustering in the data, and $\mathbf{v}_3$ is normal to the preferred plane (Scheidegger, 1965; Watson, 1966; Mardia, 1972, pp. 223–226; Mark, 1973, 1974). The method is of course only appropriate when the assumption of such orthogonality is valid. The degree of clustering about the respective eigenvectors is given by the normalized vector magnitudes (eigenvalues) $S_1$, $S_2$ and $S_3$, where $S_1 + S_2 + S_3 = 1$. The eigenvalue method, therefore, can reduce large data sets to simple descriptive statistics, allowing the ready comparison of data from many localities. Taken together, the three eigenvalues define the 'shape' of the data distribution, which varies between three end members (Fig. 1a; Watson, 1966; Woodcock, 1977; Benn, 1994a). Isotropic fabrics (with data points evenly distributed over a sphere) have $S_1 \simeq S_2 \simeq S_3$; planar girdles (with points evenly distributed around a great circle) have $S_1 \simeq S_2 > S_3$; and linear clusters (with all observations approximately parallel) have $S_1 > S_2 \simeq S_3$. In Fig. 1(a), the continuum of fabric shapes is represented on the equilateral or general shape triangle introduced by Benn (1994a,b), which is scaled using an isotropy index $I = S_3 / S_1$ and an elongation index $E = 1 - (S_2/S_1)$. Within this continuum, any fabric can be represented as a point. Another popular plotting method shown in Fig. 1(b) is the biaxial plot of $S_3$ vs $S_1$, where the usable part of the diagram is a triangle which is topologically equivalent to Fig. 1(a). Other alternatives include a triangular plot of $S_1$, $S_2$ and $S_3$ or a plot of $\log (S_1/S_2)$ vs $\log (S_2/S_3)$ (Mark, 1974; Woodcock, 1977; Benn, 1994a).

Fabric eigenvalues are widely used to distinguish sedimentary facies and deformational styles, and progress has been made in interpreting them in terms of depositional processes and strain histories (Lawson, 1979; Domack and Lawson, 1985; Dowdeswell and Sharp, 1986; Hart, 1994; Benn, 1994b, 1995). Such interpretations of eigenvalue data very rarely question the idea that the eigenvalues faithfully represent the 'true' fabric of the material, but fabric eigenvalues are derived from samples and are therefore subject to variability about the true population values. There is no guarantee that sample eigenvalues are particularly representative of the facies under investigation, or that differences between measured eigenvalues for any two samples are due to true differences or merely sampling effects. It is clearly important to be able to define confidence regions around points on fabric shape diagrams to aid in the interpretation of fabrics and to guide choices about sample size and sampling strategy.

There is unfortunately no complete theory that could allow the construction of such confidence regions, and therefore no theoretical basis for discriminating between samples or defining sample sizes. In this paper, we approach this problem by 'bootstrapping' or constructing empirical confidence regions by repeated sampling from known populations. We show that fabric eigenvalues are commonly subject to large variability, and that failure to take this into account can introduce serious errors into the interpretation of orientation data.

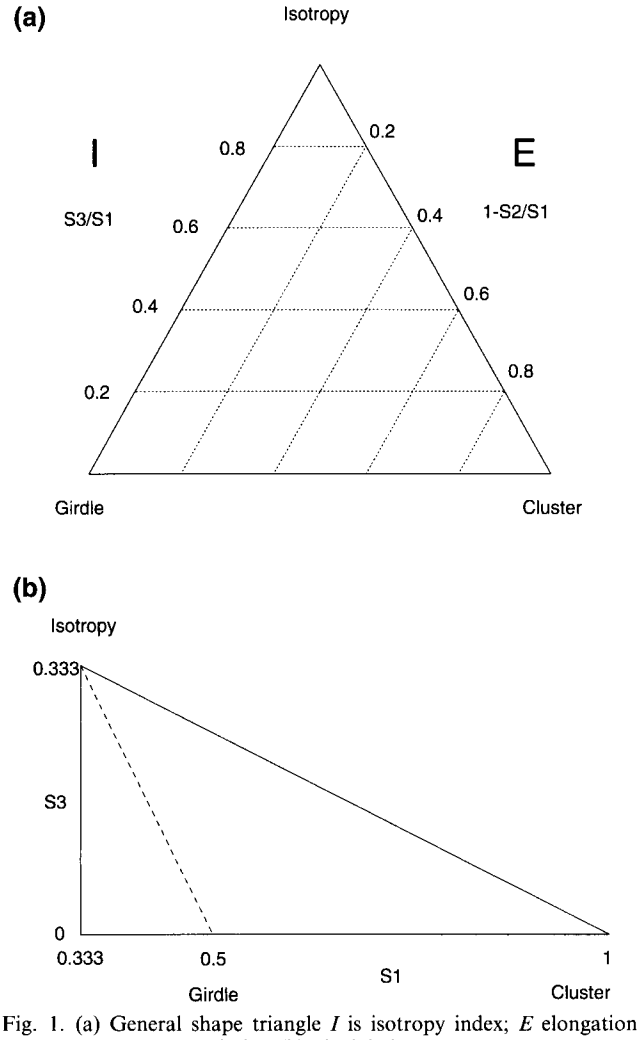


Fig. 1. (a) General shape triangle $I$ is isotropy index; $E$ elongation index. (b) Biaxial plot.

## POPULATIONS, SAMPLES AND EIGENVALUES

In ordinary univariate statistics, the fact that sample characteristics are likely to differ from those of the population is almost always recognised, and is usually allowed for by using hypothesis tests or confidence intervals. However, in multivariate cases such recognition and allowance is much less common. This is partly because of the much greater difficulty in allowing for the variability, since multivariate distributions tend to be less realistic or less tractable, and often both, than their univariate counterparts. Hence multivariate statistics has unfortunately split into two camps. Mathematical statisticians produce tests and confidence regions which are of little practical use and which make the often barely credible assumption of multivariate normality, or something similar. On the other hand applied statisticians and people in other subject areas who use multivariate statistics very often carry out their data analysis blithely ignoring the fact that they are only dealing with a sample. Fortunately this has started to change, as recent papers such as Markus (1994a,b), Milan and Whittaker (1995), and Ringrose (1992, 1994, 1996) covering correspondence analysis, PCA, and canonical variate analysis, show. In these cases observations are projected onto the new axes (eigenvectors) derived by the methods and it is the variability in these coordinates which is assessed. In ordinary PCA the standard errors of the sample eigenvalues are well known if the data are assumed to come from a multivariate normal distribution (Anderson, 1984). In the orientation tensor method it is the variability in the eigenvalues which we must consider, but we cannot assume multivariate normal data and hence the results for PCA do not carry over, despite the similarity in the methods.

Several probability distributions can be assumed for axial data (Fisher *et al.*, 1987; Jupp and Mardia, 1989). However, there are no usable results concerning the orientation tensor eigenvalues except for the isotropic case, so that previous work has always concentrated on tests for isotropy (Bingham, 1974; Woodcock, 1977; Woodcock and Naylor, 1983). While useful, these only tell a very small part of the story. In particular, the effect of gravity means that genuinely isotropic fabrics are unlikely for simple physical reasons. This paper therefore addresses the problem of creating a confidence region for any point on a fabric shape diagram.

In the present case the 'population' means the true fabric shape, and it is assumed that interest is focused on the three population eigenvalues which define this shape. By 'population' eigenvalues we simply mean the eigenvalues we would obtain by collecting an arbitrarily huge sample, and we wish to know how accurate the eigenvalues from a smaller sample will be. In the following it is sometimes assumed that the population eigenvalues are known, thus allowing certain features to be introduced and described more easily. Later we will move onto the realistic case of them being unknown. The population eigenvalues will be denoted $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and the observed sample eigenvalues $S_1 \geq S_2 \geq S_3$.

A central problem when dealing with eigenvalues, which does not seem to have been explicitly recognised in fabric shape studies, is that, in general, the difference between the highest and lowest eigenvalues will be greater for samples than for the parent population, in other words the sample eigenvalues will be more 'spread out' than the population ones. This effect is small if the population eigenvalues are well separated from each other but can be very substantial if they are at all close.

It is easy to see why this feature occurs. Suppose that the population eigenvalues are $\lambda_1 = 0.35$, $\lambda_2 = 0.33$ and $\lambda_3 = 0.32$, so that the fabric is almost isotropic. When a sample is taken, the sample eigenvalues are random variables which will depart from their population values, with the amount of the departure depending on the size and features of this sample. Let $L_j$ be the sample eigenvalue corresponding to $\lambda_j$. Suppose that the sample eigenvalues are such that $L_1 = \lambda_1 - 0.03$, $L_2 = \lambda_2 + 0.05$ and $L_3 = \lambda_3 - 0.02$, so that $L_1 = 0.32$, $L_2 = 0.38$ and $L_3 = 0.3$. Hence the first two axes have 'swopped over' and we in fact have $S_1 = L_2 = 0.38$ and $S_2 = L_1 = 0.32$. The

first sample eigenvalue/eigenvector pair corresponds to the second population eigenvalue/eigenvector pair and vice-versa. When the eigenvalues are close, especially when $n$ is small, such swopovers happen frequently so that, because the first sample eigenvalue $S_1$ is by definition the largest, $S_1$ will nearly always be larger than $\lambda_1$. By the same reasoning we nearly always have $S_3 < \lambda_3$. A similar effect occurs when only two of the eigenvalues are close, so that they will also be more spread out in the sample than in the population.

Therefore the sample eigenvalues will usually tend to make a fabric seem less isotropic, and less girdle-like, than it really is. Hence on both the general shape triangle and biaxial fabric diagrams the sample point will usually plot below and to the right of the true population point, with other plots being affected similarly. The sample points 'drift' towards the 'cluster' vertex, and the closer the population points are to the top and left of the diagrams the larger this drift will tend to be. The general shape triangle is slightly more affected by this than the biaxial plot due to the form of $I$, but the effect applies to all eigenvalue-based plots.

Similarly in all cases the sample points will 'drift' away from the edges of the plot, because these represent sample eigenvalues being exactly equal to each other, or zero, and this almost never happens. This can be compared to the case of a fair coin, where it is almost impossible to obtain exactly 50% heads so that the coin always seems more 'biased' than it really is.

Of course, tests for isotropy partially deal with this problem, showing that a sample point plotting a fair distance away from the 'isotropic' vertex could still have come from an isotropic population. Woodcock and Naylor (1983) discuss various tests, one of which is based on Monte Carlo simulation of $S_1/S_3$, though an easier way to generate the simulated data is given by Fisher *et al.* (1987). Note also that the tests of Mardia (1972) mentioned in Woodcock and Naylor (1983) are the same, of uniformity against the alternative of a Bingham distribution.

However, the feature is more widespread than this, since swopovers and drift can happen even for quite substantially non-isotropic populations. To illustrate this problem, Fig. 2 (a & b) show triangular and biaxial plots for the same one hundred simulated data sets of size 50, with population eigenvalues of 0.46, 0.4348 and 0.1052 so that $I = 0.2287$ and $E = 0.0548$. The population point, plotted as 'P', is close to the left hand 'girdle' side but nearly all of the 100 simulated samples plot further to the right, so that in many of these cases the near-girdle nature of the population would not be apparent. The samples labelled *c*,*d*,*e* are stereoplotted in Fig. 2(c–e), respectively, to illustrate the link between the methods. Note that to compare all of the samples without using the shape triangles we would need to compare 100 stereoplots.

These and all subsequent simulated data sets were generated from the angular central Gaussian distribution, defined by Tyler (1987). To do this we simply generate three observations from independent normal distributions in the usual way and then normalize them to have unit sum, thus making the three together define a direction. This is exactly the same as the method for generating isotropic samples in Fisher *et al.* (1987), except that the normal variates can have different variances. To make the data vector axial we simply constrain it further so that one of the three dimensions must have a pre-specified sign. In this case the third dimension was constrained always to be negative, by multiplying all three elements by $-1$ when this was not the case. The variances of the normal distributions define the population structure (eigenvalues) while the independence can be assumed without loss of generality. The relationship between the variances and the population eigenvalues for the orientation tensor method is not a simple one but can be calculated exactly (Ringrose and Jupp, in preparation).

To a certain extent the 'drift' phenomenon is an unavoidable feature of eigenvalues, but it is also a function of the plotting. In the present case, in which the population eigenvalues $\lambda_1$ and $\lambda_2$ are very similar, and so the population point is close to the edge of the plot, consider what happens with various values of $L_1$ and $L_2$. As $L_2$ gets closer to $L_1$ the point plots further left (and down), but when $L_2$ reaches $L_1$ and overtakes it the point is 'reflected' back from the boundary of the diagram and moves right as $L_2$ increases further relative to $L_1$, since $S_1$ is increasing relative to $S_2$. Hence the sample point is only to the left of the population point if $L_1$ and $L_2$ are closer than $\lambda_1$ and $\lambda_2$, so that when sampling produces frequent 'swopovers' the sample point will usually be to the right of the population point.

This problem can be overcome by modifying the general shape triangle to form a six-way plot, in which the assumption that the $i$th sample eigenvalue corresponds to the $i$th population eigenvalue is relaxed (Fig. 3a). The bottom panel of the diagram is the same as the general shape triangle, but in the other panels the ordering of the eigenvalues differs from those of the parent population. The six triangles represent the six possible orderings of the sample eigenvalues with respect to the population ones. For example, a sample in which $L_1$ and $L_2$ have swopped over will plot in the lower left panel, and one in which $L_1$ and $L_3$ have swopped over will plot in the upper panel. Thus, the six-way plot contains more information than the general shape triangle, which does not distinguish between points that have eigenvalue swopovers and those that have not. This diagram is similar to those introduced by Hsu (1966), Hossack (1968), Owens (1974) and Woodcock (1977), but with some differences in scaling. Figure 3(b) shows the six-way plot for the same data as in Fig. 2, with the stereoplotted samples again labelled. In this case, samples with swopovers are clearly identified, and the population point P falls well inside the cloud of sample points. A similar idea would be to plot these points inside the 'unusable' parts of the triangular plot of $S_1$, $S_2$ and $S_3$ (Mark, 1974; Benn,

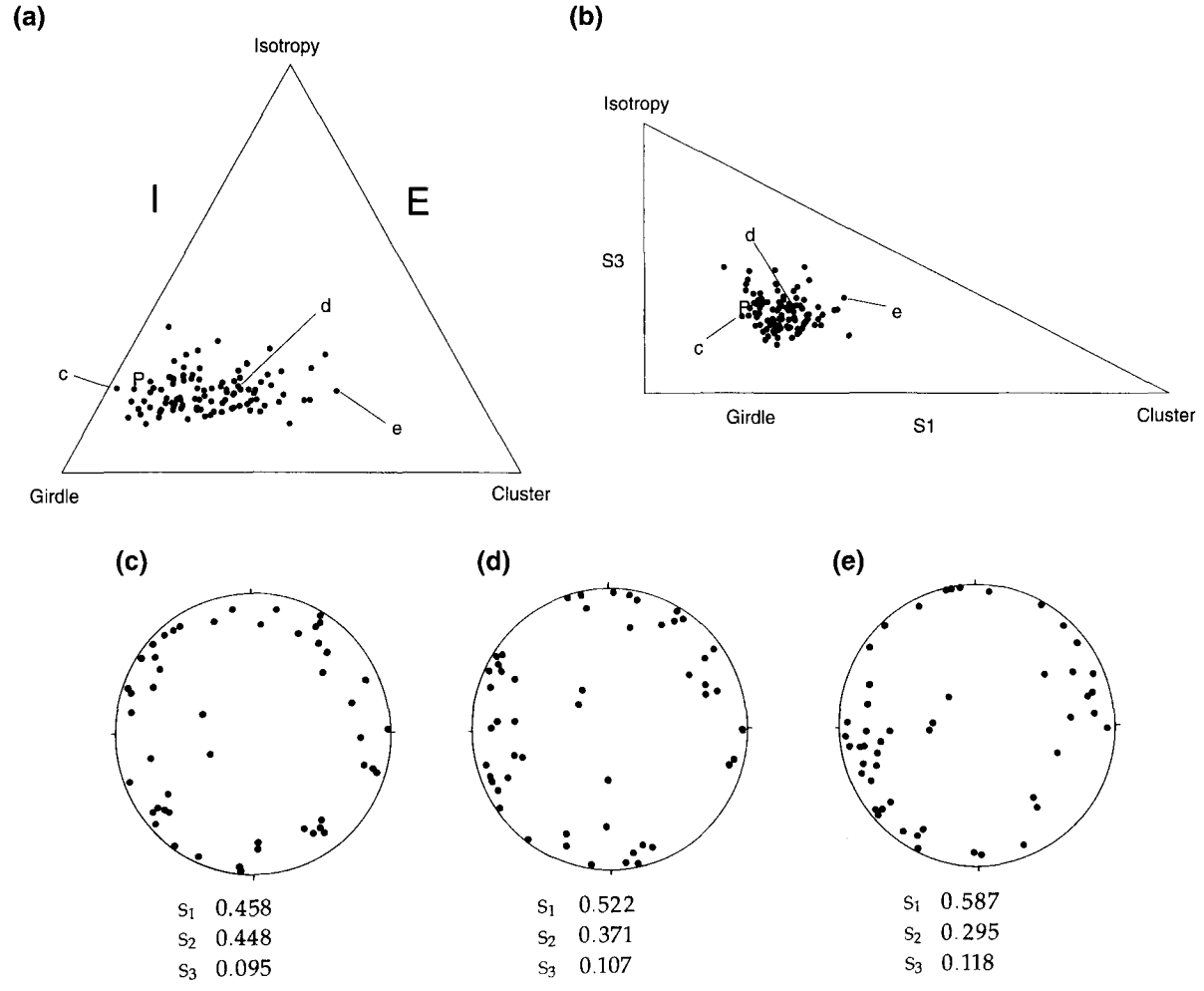


Fig. 2. (a) General shape triangle plot of 100 simulated data sets, $n = 50$. P is the population point $c$, $d$ and $e$ are sample points. (b) Biaxial plot of the same 100 simulated data sets, $n = 50$. (c–e) Stereoplots of the simulated data sets shown in (a) and (b) (c— eigenvalues 0.4572, 0.4486, 0.0941; $I = 0.2060$, $E = 0.0187$; d— eigenvalues 0.5219, 0.3714, 0.1068; $I = 0.2046$, $E = 0.2884$; e— eigenvalues 0.5874, 0.2949, 0.1177; $I = 0.2004$, $E = 0.4980$).

1994a). However, in such a diagram girdles would plot as intermediate points rather than at vertices, making it less generally useful than the six-way plot above.

Eigenvalue swopovers can be detected by inspecting the eigenvectors. When, say, the first two eigenvalues swop places, so, by definition, will the first two eigenvectors. Hence the first sample eigenvector will look like the second population eigenvector and the second like the first. Therefore swopovers are detected by comparing sample to population eigenvectors and matching the former to the latter, as outlined in Ringrose (1994, 1996). The closest fit of the sample to the population eigenvectors is ascertained and this is taken to be the ordering of the eigenvalues. This is confused somewhat by the fact that, when two eigenvalues are very close, the individual eigenvectors become ill-defined and, technically, only the plane covered by the pair together is well-defined. If the sample eigenvalues respect the order of the population eigenvalues then the sample point plots in the bottom triangle as normal. If they do not then they plot in one of the other six triangles as shown in Fig. 3(a). Some of the 'drift' problem remains, but this seems to be unavoidable.

The extent of eigenvalue/eigenvector swopovers is illustrated for a grid of points in Fig. 4. For each location of a population point on the triangular diagram we calculate the corresponding eigenvalues and generate 1000 samples from the appropriate angular central Gaussian distribution. These samples are of size 25, 50 or 100, and the pictures show the number of these 1000 simulations where the first simulated eigenvalue does not correspond to the first population eigenvalue. As expected, for a girdle distribution this happens half of the time and, for isotropy, two-thirds of the time.

Clearly the main determinants of the likelihood of swopover are the index of elongation $E$ and the sample size $n$. For $n = 100$, swopovers present a substantial problem where $E < 0.2$; for $n = 50$, swopovers are a problem for $E < 0.3$, and for $n = 25$, swopovers are problematical for $E < 0.4$. Thus, when the sample size is small, 'swopovers' may introduce spurious or misleading results in over half of the area of fabric plots, and this is even when we are ignoring swopovers between just the second and third eigenvalues.

Of course in practice we do not know the population

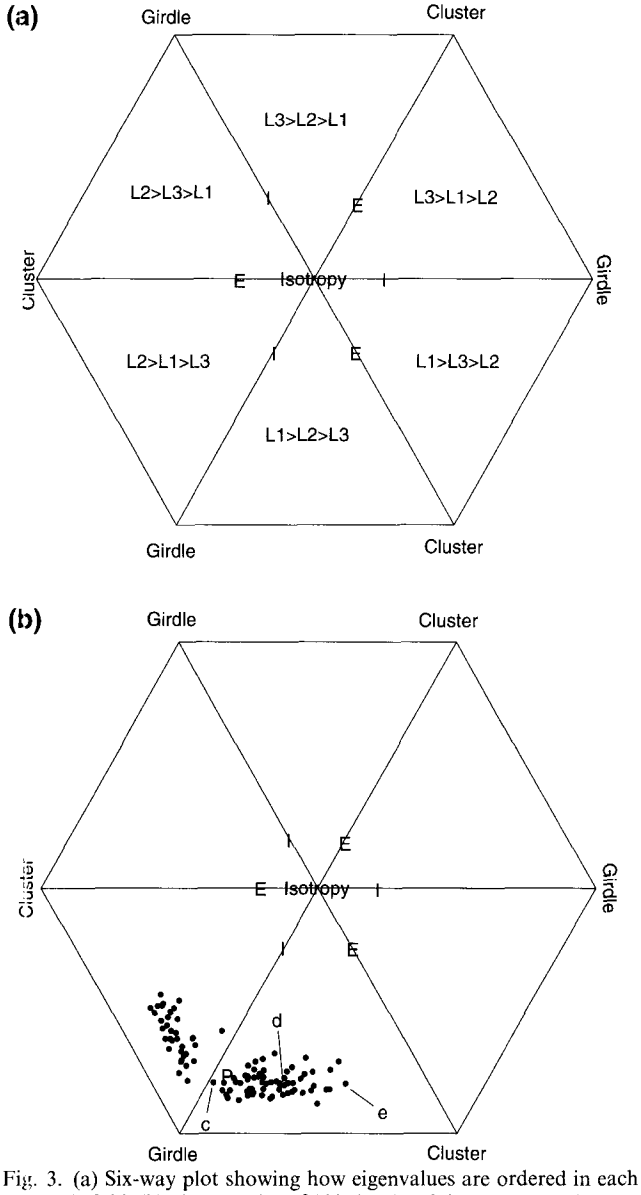


Fig. 3. (a) Six-way plot showing how eigenvalues are ordered in each sub-field. (b) Six-way plot of 100 simulated data sets, $n = 50$.

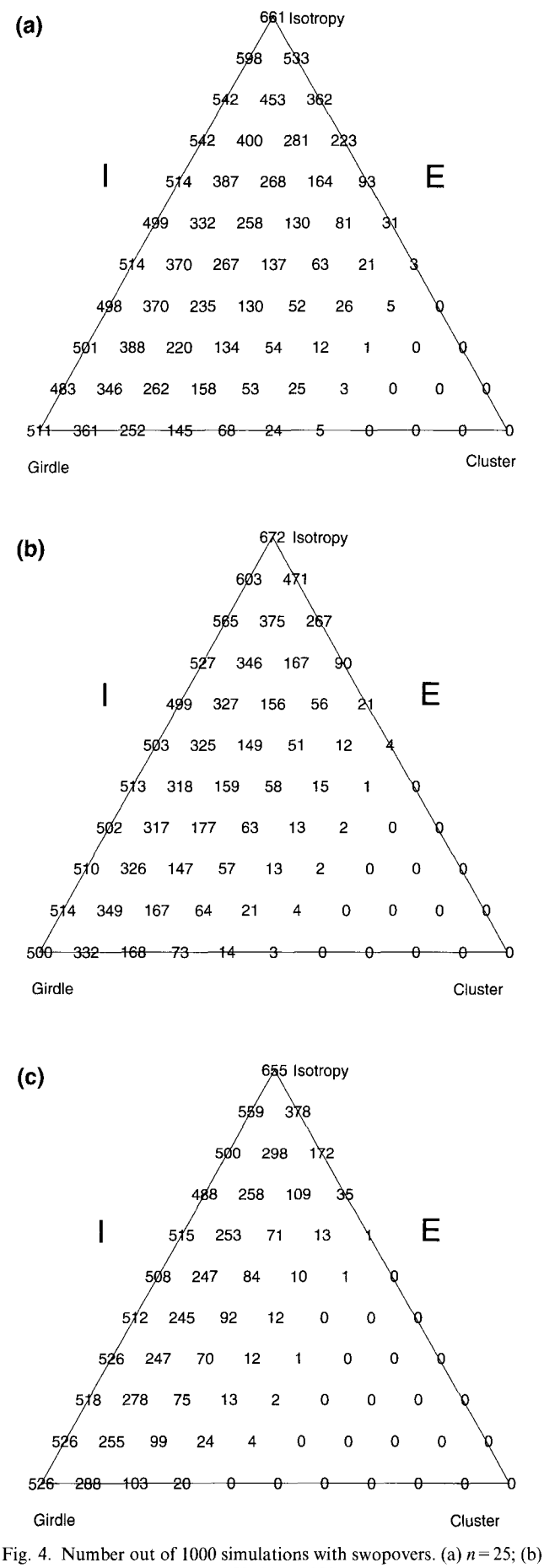


Fig. 4. Number out of 1000 simulations with swopovers. (a) $n = 25$; (b) $n = 50$; (c) $n = 100$.

eigenvalues. The above shows the kind of variability we should expect in our samples, so that if we fail to take it into account we may make serious mistakes. The following section considers the practical situation of having a sample and wishing to learn as much as possible about the population from which it was drawn.

Pal (1993) gives a comprehensive review of methods, allowing for the problems noted above, which estimate the eigenvalues of the population covariance matrix of multivariate normal data. However, although these are essentially the PCA eigenvalues, and the orientation tensor case is very similar to this, again the requisite distributional results do not carry over.

## CONFIDENCE REGIONS

Eigenvalues are notoriously difficult to produce confidence regions for because the special cases where some

eigenvalues are identical have drastically different properties to those where all are distinct. Hence, while it is often quite easy to develop tests for the special cases by assuming that they are true and then assessing the compatibility of the data, it is very difficult to produce a confidence region allowing for all possibilities. Add to this the relative paucity of appropriate theoretical results for axial data, and the commonly-expressed suspicion of the distributional assumptions in the first place, and it quickly becomes clear that theory-based confidence regions are unlikely to be practical.

Therefore we use the statistical technique of 'bootstrapping', first proposed by Efron (1979) and reviewed by Efron and Tibshirani (1993) and Young (1994). The basic idea is to use the observed data as a surrogate for the population and to draw 'bootstrap samples', or resamples, from the data. The bootstrap resamples should vary about the true sample in a similar way to that in which the possible samples vary about the population. Theoretical results exist to show that this works well in a wide variety of situations and allows variances and confidence intervals to be produced in cases where the theory is intractable. Unfortunately it is well-known that bootstrapping also has problems in dealing with eigenvalues when we cannot assume that all are distinct, an assumption that certainly cannot be made in the present case, as noted above. This is essentially due to the swopover problem. For more details see, for example, Beran and Srivastava (1985), Eaton and Tyler (1991) and Hall *et al.* (1993).

The six-way diagram introduced above allows much of the problem to be avoided. The idea is very similar to that outlined above, except that the sample takes the place of the unknown population and the bootstrap resamples replace the clouds of sample points. The spread of these resamples then gives information on how far the sample might have diverged from the population.

The eigenvalues and eigenvectors of each bootstrap resample are calculated in the usual way and in each case the correspondence between sample and resample eigenvalues and eigenvectors is ascertained, using the method noted in Ringrose (1996, pp. 580–581). In the present case this is done as follows. Let the three sample eigenvectors be $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ and the bootstrap eigenvectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$, so that we wish to see which $\boldsymbol{b}_j$ corresponds to each $\boldsymbol{a}_i$. There are six possible permutations, and the 'goodness of fit' of each permutation can be judged simply by calculating

$$\sum_{i=1}^{3} |\boldsymbol{a}_i^T \boldsymbol{b}_j| \qquad (1)$$

where $\boldsymbol{b}_j$ is the bootstrap eigenvector corresponding to the $i$th sample eigenvector in this permutation. If the $\boldsymbol{a}_i^T \boldsymbol{b}_j$ terms are all large in magnitude then we have found the permutation where the pairs of vectors all point in roughly the same direction, so that this is probably the correct one. Hence we choose the permutation of the $\boldsymbol{b}_j$s where (1) is largest.

The scheme to produce a bootstrap confidence region for the population point on a fabric shape diagram, based on a sample of size $n$, is as follows:

(1) Draw a random sample *with replacement* of size $n$ from the sample. In the original sample each of the observations occurs exactly once, of course, while in a bootstrap resample each can appear zero, one, two or more times.

(2) Repeat the above a large number $B$ of times to produce $B$ bootstrap resamples.

(3) Perform the orientation tensor method as normal on each resample.

(4) Plot each resample on the six-way fabric shape diagram, along with the original sample point, except that (1) is used to determine which segment of the plot each bootstrap resample appears in. The sample point is plotted in the bottom panel.

Stage 1 above makes this a 'non-parametric' bootstrap. An alternative is the parametric bootstrap where the data are assumed to come from a specified distribution, the parameters of which are then estimated from the data, and the bootstrap resamples are generated from this distribution.

Unfortunately there is a problem that even the six-way plot cannot fully solve. The likely amount of variability and 'drift' from population to sample and from sample to bootstrap decreases as the population and sample eigenvalues, respectively, get further from isotropy. Thus if a population is nearly isotropic then the possible samples will have very high variability, with most of them appearing much less isotropic than the population. If a sample from such a population is one of the extreme non-isotropic ones, then the variability of its bootstrap resamples will be fairly small. In contrast, if the sample is very similar to the population then its bootstrap resamples will have much larger variability. Thus, a sample close to the population point is likely to have a large confidence region while a sample far from the population point is likely to have a small region, just the opposite of what we would hope for. It is in the nature of confidence regions that, in extreme samples, the population value will be outside, but the problem is greater than usual here if the population is close to isotropy.

Similarly, it is not possible to use the bootstrapping to correct for the 'drift'. A common technique, used in Markus (1994a,b), is to use the difference between sample and bootstrap parameter estimates as an estimate of the bias in the sample value, and then subtract this from the sample value to obtain a better estimate of the population value. The problem described above means that this is very unreliable since a large difference between sample and bootstrap often follows from a small difference between population and sample.

In a univariate bootstrap confidence interval the basic procedure is to order all of the $B$ resamples and hence use them to define the confidence limits, so that the range covered by the middle 95% of the resamples defines the

95% confidence interval. (In fact numerous improvements to this can be made.) This is similar to Woodcock and Naylor's (1983) Monte Carlo test, except that it is not restricted to testing just for a fully-specified special case. However, this is much less easy in a multivariate case since there is no unique way of ordering multidimensional values. A natural multivariate equivalent of this is to show the convex hulls of the point cloud, as in Greenacre (1984), Ringrose (1992) and Markus (1994a,b). The convex hull of a cloud of points is simply the set of lines joining the 'outer' points. If an elastic band is stretched around the outside of the point cloud and allowed to contract, then the points it touches are the outer points. Convex hulls are not perfect since they are highly sensitive to a small number of unusual points. However, by using the 'peeling' algorithm described in Green and Silverman (1979) we can also consider the more stable inner hulls. Peeling is simply the process of successively removing the points on the current hull and recalculating the hull for the remaining points. There is also the minor problem of the concavity of some point clouds, as in Fig. 3(b). In that case it might be better to produce separate hulls for each panel, though such an approach would fail when points appeared in more than two or three panels.

The method is illustrated in Fig. 5. Samples of size 50 have been generated from four different populations: (a) close to isotropic: eigenvalues are 0.3554, 0.3338 and 0.3108 ($I=0.875$, $E=0.0601$); (b) no special features: eigenvalues are 0.5192, 0.3564 and 0.1245 ($I=0.240$, $E=0.314$); (c) close to a girdle: eigenvalues are 0.4600, 0.4348 and 0.1052 ($I=0.229$, $E=0.055$); (d) close to a cluster: eigenvalues are 0.6425, 0.1845 and 0.1730 ($I=0.269$, $E=0.713$).

In each case the population eigenvectors were taken, without loss of generality, to be (1,0,0), (0,1,0) and (0,0,−1), corresponding to the first, second and third population eigenvalues, respectively.

In each diagram the population point is plotted as 'P' while ten sample points are plotted as 'S'. Since we are considering the situation in which we do not know the population eigenvalues, the sample points all plot in the lower panel. For each sample, the second convex hull based on $B=1000$ bootstraps is shown. Only the hulls are shown, the actual bootstrap points are omitted for clarity.

In Fig. 5(a) the population is almost isotropic but some of the sample points plot a long way down the bottom panel, all plotting below the population point. However, the confidence regions of all but the two bottom-left sample points include the population point. The fact that most regions include the isotropic point means that, for most samples drawn from this population, the geologist would conclude that the fabric was effectively isotropic. In Fig. 5(b) the population point is fairly central but most of the sample points drift towards the cluster vertex, though only one region fails to include the population point. Note that one region includes a swopover between $S_2$ and $S_3$ while most of the others include one between $S_1$ and $S_2$. In Fig. 5(c) the population is close to being a girdle so that all confidence regions include the possibility of a swopover between $S_1$ and $S_2$. All the regions include the population point. In Fig. 5(d) the confidence regions are noticeably smaller due to the much smaller variability in near-cluster populations, and all but one sample plots close to the population point.

It would be possible, though computationally awkward, to 'fold' the six-way plot back onto the ordinary one-way picture, with the confidence region comprising all areas of the plot where the equivalent area in any of the six panels was included in the confidence region as derived above. However, the only real gain here would be to save paper, so that the loss of information does not seem worthwhile.

The size of the regions should not be too surprising. For $n=50$ the Monte Carlo test of Woodcock and Naylor (1983) will reject the hypothesis of isotropy only for $I<0.5$, so that the sample point would have to plot halfway down the general shape triangle before we would conclude that the fabric was non-isotropic. The results here show that this variability is not confined merely to isotropic fabrics.

## PERFORMANCE OF THE CONFIDENCE REGIONS

A simulation study was performed to assess the accuracy of the proposed confidence regions. The four population structures used in Fig. 5, the cases $n=25$, 50, 100 and the cases $B=100$, 1000 are considered. For each combination of population, $n$ and $B$, 1000 samples were generated and the confidence regions constructed for each one as described in the previous section. Then for each sample it was assessed whether the population point plotted within the generated confidence region. If the population point plots within the region around 95% of the time, then the regions are roughly equivalent to 95% confidence regions and so on. Note that if the population point plots inside the region every time then this probably means that they are much too big, whereas if it plots inside less than about 90% then they are probably not very useful.

The use of a particular parametric distribution to test the proposed regions should not be misleading since the angular central Gaussian distribution fits precisely the kind of data to which the orientation tensor method can validly be applied. Starkey (1993) criticises theoretical distributions for being unable to cope with real data, but the angular central Gaussian should be better than most as it can deal with any distribution based on three orthogonal axes, in other words the situation that the orientation tensor method assumes. For example neither can cope with paired maxima (Starkey, 1993, p. 1357), since the first eigenvalue and eigenvector will then be averages of the two maxima and hence probably be

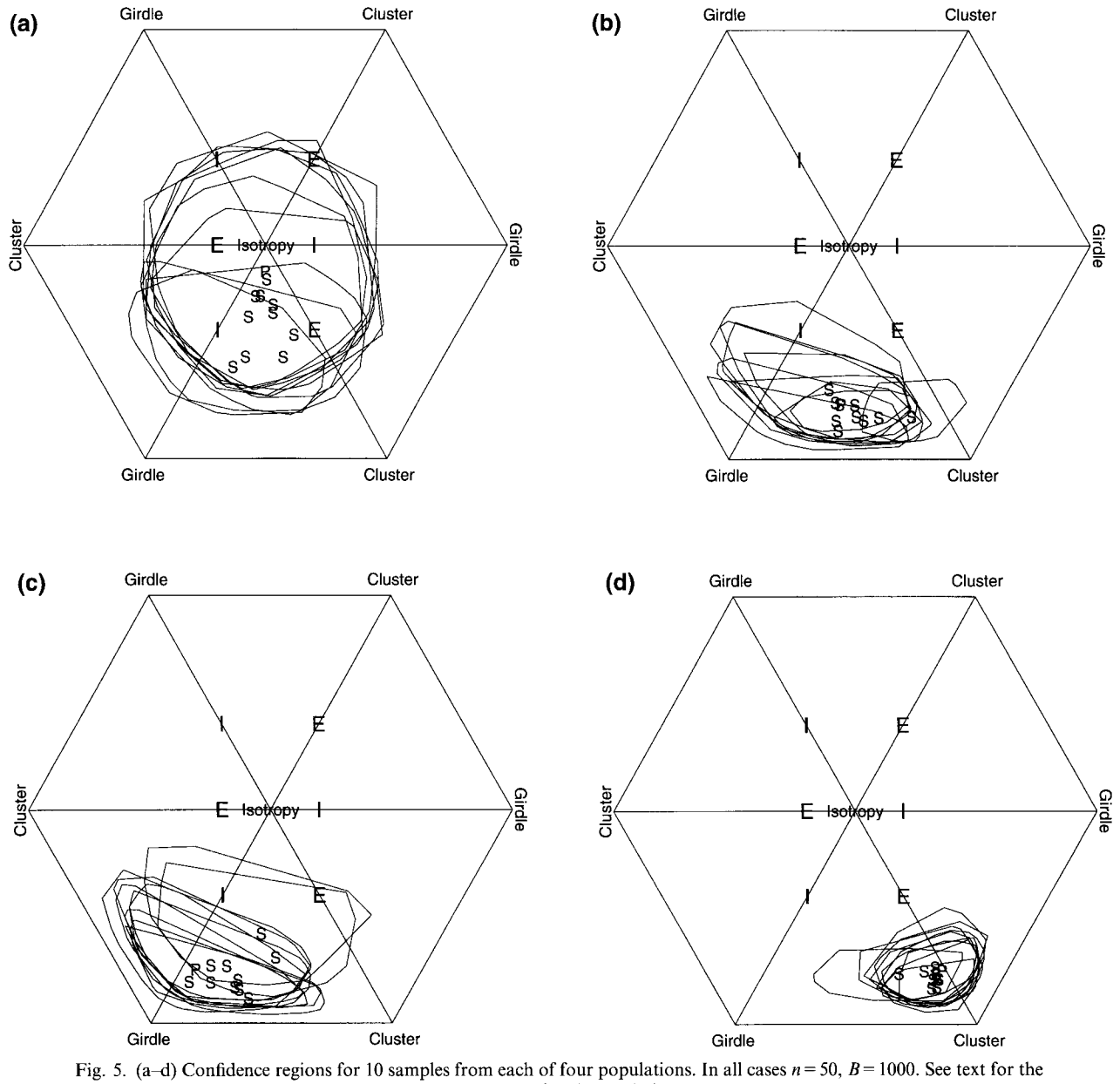


Fig. 5. (a–d) Confidence regions for 10 samples from each of four populations. In all cases $n = 50$, $B = 1000$. See text for the eigenvalues of each population.

misleading. This distribution has received relatively little attention in the literature simply because it is rather intractable mathematically, but it has many appealing properties. Hence conclusions drawn here should be approximately valid in most practical situations where it is appropriate to use the orientation tensor method in the first place.

Table 1 shows the percentages of the samples in which the population points plotted inside the outermost three convex hulls of the bootstrap regions, for each combination of factor levels considered above. For example, in the near-isotropic case with $n = 50$ and 1000 bootstraps, the population point was within the 3rd hull in 85.3% of cases, between the 2nd and 3rd in 3.9%, between the outer and 2nd in 6.0% and outside the outer hull in 4.8% of cases. When $B = 100$, in all cases except one, more than 10% of outer hulls failed to include the population point, while when $B = 1000$, in all cases except one, fewer than 10% of outer hulls failed to include the population point. These results strongly suggest that 1000 bootstraps are necessary to obtain reasonable coverage.

Clearly samples of size $n = 25$ seem to be too small, with sometimes twice as many confidence regions failing to include the population point as for $n = 50$. However, $n = 100$ offers a rather smaller advantage over 50. Hence it seems that taking samples of size 50 is fairly reasonable. This is gratifying as many geologists adopt a sample size of 50 anyway.

As previously noted, the outermost convex hull is highly sensitive to unusual points so that it may be more sensible to consider the 2nd hull. From Table 1, for $B = 1000$ and $n \geq 50$ the second hull contains the population point in at least 89% of the simulations. Hence the second hull can be regarded as being in effect a confidence

Table 1. Percentage of samples with population point inside and outside bootstrap convex hulls in six-way triangular plot

| | Eigenvalues | $I$ | $E$ | $B=$ $n$ | 100 out | 1000 out | 100 in 1st | 1000 in 1st | 100 in 2nd | 1000 in 2nd | 100 in 3rd | 1000 in 3rd |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0.3554 | | | 25 | 30.1 | 7.5 | 23.6 | 7.5 | 18.5 | 7.2 | 27.8 | 77.8 |
| *a* | 0.3338 | 0.87 | 0.06 | 50 | 26.0 | 4.8 | 21.6 | 6.0 | 16.3 | 3.9 | 36.1 | 85.3 |
| | 0.3108 | | | 100 | 20.9 | 3.9 | 19.9 | 3.6 | 16.0 | 4.7 | 43.2 | 87.8 |
| | 0.5192 | | | 25 | 15.2 | 4.5 | 12.0 | 3.4 | 9.3 | 3.9 | 63.5 | 88.2 |
| *b* | 0.3564 | 0.24 | 0.31 | 50 | 13.9 | 3.8 | 10.2 | 2.8 | 11.8 | 2.8 | 64.1 | 90.6 |
| | 0.1245 | | | 100 | 13.1 | 3.1 | 9.8 | 2.9 | 11.8 | 3.1 | 65.3 | 90.9 |
| | 0.4600 | | | 25 | 12.9 | 2.6 | 14.2 | 2.8 | 15.0 | 2.3 | 57.9 | 92.3 |
| *c* | 0.4348 | 0.23 | 0.05 | 50 | 11.5 | 1.6 | 16.0 | 2.7 | 14.7 | 2.8 | 57.8 | 92.9 |
| | 0.1052 | | | 100 | 9.6 | 1.0 | 14.7 | 2.1 | 14.1 | 1.6 | 61.6 | 95.3 |
| | 0.6425 | | | 25 | 21.6 | 10.9 | 12.5 | 4.7 | 14.1 | 3.3 | 51.8 | 81.1 |
| *d* | 0.1845 | 0.27 | 0.71 | 50 | 16.8 | 4.7 | 14.1 | 4.2 | 13.1 | 5.1 | 56.0 | 86.0 |
| | 0.1730 | | | 100 | 14.7 | 3.0 | 14.8 | 2.4 | 15.3 | 2.9 | 55.2 | 91.7 |

region with coverage percentage somewhere between 90% and 95%. Alternatively, we could use even more bootstraps and take the third or fourth hull.

Similar studies were also performed for the simple triangular and biaxial plots, but these confidence regions performed disastrously in all cases except that of (b), the 'central' point. This was as expected, given the problems with swopovers and 'reflection' noted earlier. Regions based on using the whole of the triangular plot of $S_1$, $S_2$ and $S_3$ would probably behave reasonably, though as noted above such plots represent the data rather less informatively than those described here, and hence have not been investigated. In cases where the samples are mostly clusters and interest is in showing different levels of clustering then the plot of $\log(S_1/S_2)$ vs $\log(S_2/S_3)$ is probably more useful. Confidence regions could easily be constructed using the same basic ideas as here, though the six-way plot is likely to be unnecessary.

## CONCLUSIONS

The most important conclusion to be drawn from this study is that the variability in sample eigenvalues is very large. This can be seen from both the plots of multiple samples (Fig. 2) and those of the proposed confidence regions (Fig. 5). This suggests that, for many cases in the literature, interpretations of the spread of sample points may be unnecessary since the spread can be explained just as easily as being due to random variation. Although confidence regions do not deal directly with pairwise comparisons between samples, it is reasonable to conclude that, if a pair of sample points are each contained within the confidence region from the other sample, then there is little evidence for a difference between the populations from which they were drawn.

A sample size of 50 seems to be the absolute minimum necessary to obtain usable results. The size of the confidence regions and the extent of swopovers indicate that, unless the population is an obvious cluster, results from smaller samples will be of minimal use. The results of the simulation study show that a sample size of 100 leads to only slightly higher rates of inclusion for the confidence regions. However, it must be emphasized that a larger sample size will always reduce the variability in the sample eigenvalues, thus reducing the spread of sample points on the six-way plot and, usually, reducing the size of the confidence region around each of these sample points. A larger sample size will also reduce the effect of eigenvalue 'drift'. Hence a larger sample will always lead to more accurate conclusions, and the size should always be at least 50.

## COMPUTATIONAL DETAILS

All calculations were performed in a SUN UNIX environment using a PASCAL program calling FORTRAN and NAG subroutines and the plots were produced using the statistics package S+. Copies of the programs can be obtained free of charge from the first author by e-mail at ringrose@rmcs.cranfield.ac.uk.

*Acknowledgements*—The authors would like to thank Peter Jupp and Ben Nasatyr for help with the topology.

## REFERENCES

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

Benn, D. I. (1994) Fabric shape and the interpretation of sedimentary fabric data. *Journal of Sedimentary Research* **A64**, 910–915.

Benn, D. I. (1994) Fluted moraine formation and till genesis below a temperate glacier: Slettmarkbreen. Jotunheimen, Norway. *Sedimentology* **41**, 279–292.

Benn, D. I. (1995) Fabric signature of subglacial till deformation. Breidamerkurjokull, Iceland. *Sedimentology* **42**, 735–747.

Beran, R. and Srivastava, M. S. (1985) Bootstrap tests and confidence regions for functions of a covariance matrix. *Annals of Statistics* **13**, 96–115.

Bingham, C. (1974) An antipodally symmetric distribution on the sphere. *Annals of Statistics* **2**, 1201–1225.

Domack, E. W. and Lawson, D. E. (1985) Pebble fabric in ice-rafted diamicton. *Journal of Geology* **93**, 577–591.

Dowdeswell, J. A. and Sharp, M. J. (1986) Characterization of pebble fabrics in modern terrestrial glacigenic sediments. *Sedimentology* **33**, 699–710.

Eaton, M. L. and Tyler, D. E. (1991) On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix. *Annals of Statistics* **19**, 260–271.

Efron, B. (1979) Bootstrap methods: Another look at the jackknife. *Annals of Statistics* **7**, 1–26.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

Fisher, N. I., Lewis, T. and Embleton, B. J. J. (1987). *Statistical Analysis of Spherical Data*. Cambridge University Press, Cambridge.

Green, P. J. and Silverman, B. W. (1979) Constructing the convex hull of a set of points in the plane. *The Computer Journal* **22**, 262–266.

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

Hall, P., Hardle, W. and Simar, L. (1993) On the inconsistency of bootstrap distribution estimators. *Computational Statistics and Data Analysis* **16**, 11–18.

Hart, J. K. (1994) Till fabric associated with deformable beds. *Earth Surface Processes and Landforms* **19**, 15–32.

Hsu, T. C. (1966) The characteristics of coaxial and non-coaxial strain paths. *Journal of Strain Analysis* **1**, 216–222.

Hossack, J. R. (1968) Pebble deformation and thrusting in the Bygdin area (Southern Norway). *Techtonophysics* **5**, 315–339.

Jupp, P. E. and Mardia, K. V. (1989) A unified view of the theory of directional statistics, 1975–1988. *International Statistical Review* **57**, 261–294.

Lawson, D. E. (1979) A comparison of the pebble orientations in ice and deposits of the Matanuska Glacier, Alaska. *Journal of Geology* **87**, 629–645.

Mardia, K. V. (1972). *Statistics of Directional Data*. Academic Press, London.

Mark, D. M. (1973) Analysis of axial orientation data, including till fabrics. *Geological Society of America Bulletin* **84**, 1369–1374.

Mark, D. M. (1974) On the interpretation of till fabrics. *Geology* **2**, 101–104.

Markus, M. T. (1994a). Bootstrap confidence regions for homogeneity analysis; the influence of rotation on coverage percentages. In Dutter, R. and Grossmann, W. (Eds) *COMPSTAT: Proceedings in Computational Statistics: 11th Symposium held in Vienna, Austria, 1994*. Physica-Verlag, Heidelberg. pp. 337–342.

Markus, M. T. (1994b). *Bootstrap Confidence Regions in Non-Linear Multivariate Analysis*. DSWO Press, Leiden.

Milan, L. and Whittaker, J. (1995) Application of the parametric bootstrap to models that incorporate a singular value decomposition. *Applied Statistics* **44**, 31–50.

Owens, W. H. (1974) Representation of finite strain by three-axis planar diagrams. *Geological Society of America Bulletin* **85**, 307–310.

Pal, N. (1993) Estimating the normal dispersion matrix and the precision matrix from a decision-theoretic point of view: a review. *Statistical Papers* **34**, 1–26.

Ringrose, T. J. (1992) Bootstrapping and correspondence analysis in archaeology. *Journal of Archaeological Science* **19**, 615–629.

Ringrose, T. J. (1994). Bootstrap confidence regions for canonical variate analysis. In Dutter, R. and Grossmann, W. (Eds) *COMPSTAT: Proceedings in Computational Statistics: 11th Symposium held in Vienna, Austria, 1994*. Physica-Verlag, Heidelberg. pp. 343–348.

Ringrose, T. J. (1996) Alternative confidence regions for canonical variate analysis. *Biometrika* **83**, 575–587.

Scheidegger, A. E. (1965) On the statistics of the orientation of bedding planes, grain axes, and similar sedimentological data. *United States Geological Survey Professional Paper* **525**(C), 164–167.

Starkey, J. (1993) The analysis of three-dimensional orientation data. *Canadian Journal of Earth Sciences* **30**, 1355–1362.

Tyler, D. E. (1987) Statistical analysis for the angular central Gaussian distribution on the sphere. *Biometrika* **74**, 579–589.

Watson, G. S. (1966) The statistics of orientation data. *Journal of Geology* **74**, 786–797.

Woodcock, N. H. (1977) Specification of fabric shapes using an eigenvalue method. *Geological Society of America Bulletin* **88**, 1231–1236.

Woodcock, N. H. and Naylor, M. A. (1983) Randomness testing in three-dimensional orientation data. *Journal of Structural Geology* **5**, 539–548.

Young, G. A. (1994) Bootstrap: more than a stab in the dark? *Statistical Science* **9**, 382–395.