JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 96, NO. B7, PAGES 11,723–11,740, JULY 10, 1991

Bootstrap Statistics for Paleomagnetic Data

LISA TAUXE, NIELS KYLSTRA, AND CATHY CONSTABLE

Scripps Institution of Oceanography, La Jolla, California

The power and utility of paleomagnetic analyses stem largely from the ability to quantify such parameters as the degree of rotation of a rock body, or the orientation of an anisotropy axis. Until recently, estimates for uncertainty in these paleomagnetically determined parameters derived from assumptions concerning the underlying parametric distribution functions of the data. In many geologically important situations, the commonly used parametric distribution functions fail to model the data adequately and the uncertainty estimates so obtained are unreliable. Such essentials as the test for common mean require data sets consistent with a spherically symmetric underlying distribution; their application in inappropriate circumstances can result in flawed interpretations. Moreover, the almost universally used approximation for a cone of 95% confidence for the mean of a sample drawn from a Fisher distribution is quite biased even for moderate dispersions (κ = 25). The availablity of inexpensive, powerful computers makes possible the empirical estimation of confidence regions by means of data resampling techniques such as the bootstrap. These resampling schemes replace analytical solutions with repeated simple calculations. We describe a bootstrap approach for the calculation of uncertainties for means or principal directions of paleomagnetic data. The method is tested on means of simulated Fisher distributions with known parameters and is found to be reliable for data sets with more than about 25 elements. Because a Fisher distribution is not assumed, the approach is applicable to a wide range of paleomagnetic data and can be used equally well on directions or associated virtual poles. We also illustrate bootstrap techniques for the discrimination of directions and for the fold test which enable the use of these powerful tests on the wider range of data sets commonly obtained in paleomagnetic investigations.

INTRODUCTION

Since *Fisher* [1953] first gave a detailed account of the probability distribution named after him, paleomagnetists have relied on it for the statistical analysis of directional data. The assumption is that the distribution underlying the data is symmetric about a mean direction; this effectively prohibits bimodal or elliptically distributed data sets. Thus, many data sets must be subjectively edited in order to satisfy the constraints of Fisher's distribution. Probability distributions less restrictive than Fisher's have been proposed; in some ways these are more appropriate for paleomagnetic data (see *Fisher et al.*, [1987] for a comprehensive review). For example, *Bingham* [1964] introduced a distribution that allows both bimodal and elliptical data. The modes are assumed to be antipodal and to have the same dispersion. However, the calculation of the confidence regions is computationally taxing and most implementations rely on interpolation of tabulated parameters limiting the data sets for which Bingham parameters can be calculated. The Kent distribution (FB_5 of *Kent* [1982]) is also elliptical in shape and can be either unimodal or bimodal, but a method for the determination of confidence regions is available only for the unimodal case.

The very nature of the Earth's magnetic field produces directional data that are not amenable to standard parametric analysis. Records of reversals in field polarity are characterized by bimodal data with intermediate directions which often depart from spherical symmetry. Such behavior is therefore intrinsic even to data of very high quality. Moreover, recent compilations of paleomagnetic data by *Schneider and Kent* [1988] suggest that there may be long term differences between the average normal and reversed fields, producing nonantipodal modes and violating the basic premise of the Bingham approach. Even if directional data are distributed in a spherically symmetric fashion (although many data sets are not), the associated virtual geomagnetic poles cannot be, owing to the mapping function relating the two. This distortion leads many paleomagnetists to calculate confidence ellipses for the associated paleomagnetic poles by stretching the circular confidence regions derived from Fisher statistics. These confidence ellipses often do not reflect the actual distribution of the virtual poles. Therefore, it is not just "bad" paleomagnetic data that are poorly suited for standard statistical treatment, it is a significant proportion of paleomagnetic data. What is required is a technique for estimating confidence regions that is less constrained by the few available parametric models with theoretical solutions.

The availablity of inexpensive, powerful computing resources makes it possible to estimate confidence regions empirically by means of data resampling techniques such as the bootstrap and jackknife [*Efron* 1982; *Tichelaar and Ruff*, 1989; *Wu*, 1986]. These resampling schemes replace analytical solutions with repeated simple calculations which can be used to provide confidence intervals for a given parameter. Several articles have proposed bootstrap schemes for directional data [e.g. *Davison, et al.*, 1986; *Hall*, 1987; *Ducharme et al.*, 1985; *Fisher and Hall*, 1989]. We propose a practical bootstrap approach suitable for most paleomagnetic data that provides a compact representation of confidence intervals. Our method also lends itself well to the test for common mean and the fold tests, both of fundamental importance to paleomagnetic data analysis.

STATISTICS OF PALEOMAGNETIC DIRECTIONAL DATA: PARAMETRIC APPROACHES

The Fisher Distribution

Currently available techniques for statistical analysis of paleomagnetic data have been thoroughly reviewed by *Fisher et al.* [1987] to which the reader is referred for a

Paper number 91JB00572.
0148-0227/91/91JB-00572$05.00

comprehensive discussion. We will briefly describe the methods which are of interest to the present discussion. First and foremost is the classical approach of *Fisher* [1953], discussed by many textbooks (see, for example, *McElhinny*, [1973], and *Tarling*, [1981]). The power of the approach derives from its solid statistical underpinnings and the appeal stems from its simplicity; most parameters of interest can be determined (at least approximately) using a hand calculator.

Data are assumed to be distributed according to the Fisher probability density function:

$$c(\kappa)^{-1} e^{\kappa \cos\theta}$$

where c is a constant of integration, κ is a positive concentration parameter, and θ is the angular separation of a particular direction away from the true mean direction. This distribution is circularly symmetric about a mean direction with maximum probability near the mean and minimum probability at its antipode. The parameters of interest are calculated as follows:

Denote the data set composed of unit vectors by $\mathbf{X}$. Let $\mathbf{X}$ contain n vectors, the kth vector being designated as $\mathbf{X}_k$, with $\mathbf{X}_k$ composed of the cartesian coordinates x_{ki} so that $\mathbf{X}_k = (x_{k1}, x_{k2}, x_{k3})$. The length of the resultant vector, R, is given by

$$R = \left(\sum_{k=1}^{n} \mathbf{X}_k \cdot \sum_{k=1}^{n} \mathbf{X}_k\right)^{\frac{1}{2}} \quad (1)$$

The maximum likelihood direction for a Fisher distribution is the mean and is given by

$$\bar{\mathbf{X}} = \frac{1}{R} \sum_{k=1}^{n} \mathbf{X}_k \quad (2)$$

The concentration parameter κ is estimated by $\kappa \simeq k = (n-1)/(n-R)$ in most paleomagnetic applications (but see *McFadden*, [1980], *Fisher et al.*, [1987] for a discussion of other approximations) and ranges from $0 \rightarrow \infty$; as κ goes to infinity, the dispersion goes to zero. *Fisher* [1953] shows that for $R > n - 2$ the semi-angle of a cone of 95% confidence (α_{95}) for the mean direction can be estimated by

$$\alpha_{95} = \cos^{-1}\left[1 - \frac{n-R}{R}\left(20^{\frac{1}{(n-1)}} - 1\right)\right] \quad (3)$$

For many paleomagnetic applications, α_{95} is further approximated by

$$\alpha'_{95} = \frac{140}{\sqrt{kn}} \quad (4)$$

which is usually considered to be reliable for k larger than about 7 (see, for example, *McElhinny*, [1973]).

The Bingham Distribution

The Bingham distribution [*Bingham*, 1964; 1974] was popularized by *Onstott* [1980] for use in paleomagnetic studies because it allows bimodal and elliptically distributed data. Here we call on the matrix of sums of squares and products first introduced by *Watson* [1960]. This matrix, when normalized by n, is usually referred to as the "orientation matrix" [e.g., *Scheidegger*, 1965]. The orientation matrix $\mathbf{T}$ of $\mathbf{X}$ is a second-order symmetric matrix where

$$\mathbf{T} = \frac{1}{n}\mathbf{X}^T\mathbf{X} \quad (5)$$

The orientation matrix is similar to a moment of inertia matrix with unit mass assigned to each data point on the sphere. When diagonalized, both yield the same eigenvectors, γ_i, where $i = 1, 2, 3$. However, the largest eigenvalue of the orientation matrix corresponds to the direction of the smallest moment of inertia, or the direction most heavily weighted by the data. Additionally, the eigenvalues of the orientation matrix conveniently sum to 1. The eigenvectors associated with the largest, the intermediate, and the smallest eigenvalues will be referred to as the principal (γ_1), the major (γ_2), and minor (γ_3) eigenvectors with the corresponding eigenvalues, τ_1, τ_2, and τ_3. Bingham showed that the maximum likelihood estimate for the true mean direction of a Bingham distributed sample is the principal eigenvector of $\mathbf{T}$ (not the vector mean as calculated in equation (2)). The Bingham probability density function is given by

$$c(\kappa_2, \kappa_3)^{-1} e^{(\kappa_2 \cos^2\phi + \kappa_3 \sin^2\phi)\sin^2\theta}$$

Here, κ_2 and κ_3 are negative concentration parameters. Again, θ is the angle between a given direction and the true mean direction (here estimated by the principal eigenvector of $\mathbf{T}$), and ϕ is an angle in the plane perpendicular to the true direction with $\phi = 0$ parallel to the major eigenvector in that plane. The parameters κ_2 and κ_3 are unfortunately onerous to estimate. However, *Mardia and Zemroch* [1977] have computed values for many cases of interest and look-up tables are available in terms of calculated eigenvalues (τ_i) of the orientation matrix $\mathbf{T}$. Bingham provides no explicit expression for the probability density of the mean, hence computing confidence regions about the mean is somewhat problematical. However, when the variances about the principal direction of $\mathbf{T}$ are small (when κ_2 and $\kappa_3 \ll 0$), the confidence regions can be estimated. As discussed by *Fisher et al.* [1987], the variances about the major and minor eigenvectors, respectively, may be approximated by

$$\hat{\sigma}_2^2 = \frac{1}{2(\hat{\kappa}_2)(\tau_2 - \tau_1)} \qquad \hat{\sigma}_3^2 = \frac{1}{2(\hat{\kappa}_3)(\tau_3 - \tau_1)}$$

The lengths of the major and minor semiaxes of the 95% confidence ellipse are then $s_2 = \hat{\sigma}_2\sqrt{g}$ and $s_3 = \hat{\sigma}_3\sqrt{g}$ where $g = -2\ln(0.05)$. The semiangles subtended by these axes are $\zeta_{95} = \sin^{-1}(s_2)$ and $\eta_{95} = \sin^{-1}(s_3)$. These semiaxes are parallel to the major and minor eigenvectors respectively of the orientation matrix with declinations and inclinations of Dec_ζ, Inc_ζ and Dec_η, Inc_η respectively.

The Kent Distribution

There are many cases where the appropriate Bingham concentration parameters cannot be calculated. The probability distribution function is also different from the Fisher distribution, even for the spherically symmetric case. *Kent* [1982] proposed an alternative probability distribution function which is the elliptical analogue of the Fisher distribution.

$$c(\kappa, \beta)^{-1} e^{\kappa\cos\theta + \beta\sin^2\theta\cos 2\phi}$$

Here, θ and ϕ are as before, κ is a concentration parameter similar to the Fisher κ, and β is the "ovalness" parameter. When β is zero, the Kent distribution reduces to a Fisher distribution. Calculation of approximate confidence ellipses is more straightforward than for the Bingham technique and does not require the numerical evaluation of messy integrals. The parameters of interest are calculated as follows: The

data are first rotated into the sample coordinates by the transformation

$$\mathbf{X}' = \Gamma^T \mathbf{X}$$

where $\Gamma = (\gamma_1, \gamma_2, \gamma_3)$, and the columns of Γ are called the constrained eigenvectors of $(\bar{\mathbf{X}}, \mathbf{T})$. The vector γ_1 is proportional to $\bar{\mathbf{X}}$, whereas γ_2 and γ_3 (the major and minor axes) diagonalize $\mathbf{T}$ as much as possible subject to being constrained by γ_1 (see *Kent* [1982] but note that his x_1 corresponds to x_3 in conventional paleomagnetic notation). Then the following parameters may be computed

$$\hat{\mu} = n^{-1}\sum_i x_{k1}', \quad \hat{\sigma}_2^2 = n^{-1}\sum_k (x_{k2}')^2,$$

$$\hat{\sigma}_3^2 = n^{-1}\sum_k (x_{k3}')^2. \tag{6}$$

As defined here, $\hat{\mu} = R/n$ to a very good approximation. Also to good approximation, $\hat{\sigma}_2^2 = \bar{\tau}_2$, and $\hat{\sigma}_3^2 = \bar{\tau}_3$. The semiangles ζ_{95} and η_{95} subtended by the major and minor axes of the 95% confidence ellipse are then

$$\zeta_{95} = \sin^{-1}(\sigma_2\sqrt{g}), \quad \eta_{95} = \sin^{-1}(\sigma_3\sqrt{g}) \tag{7}$$

here $g = -2\ln(0.05)/(n\hat{\mu}^2)$.

As discussed in the introduction, each of these techniques has some advantages as well as some drawbacks when applied to real (i.e. nonideal) data. We now describe a technique for estimating confidence regions which postpones the need for restrictive parametric assumptions to the distribution of the mean which is more likely to be "well behaved" than the data themselves.

A BOOTSTRAP APPROACH

For purposes of illustration, we have chosen what may be regarded as a "typical" paleomagnetic data set. The data are those from the KUL section published by *Tauxe and Opdyke* [1982] and are plotted on an equal area projection in Figure 1. Although these were treated using Fisher statistics by the authors, the data have several features which make them less than ideal candidates for such an analysis. Fisher statistics are only appropriate for data sets drawn from unimodal spherically symmetric distributions. The data shown in Figure 1 are not. As with most paleomagnetic data sets, this one is bimodal and requires separation of the data into normal or reversed modes. Whereas this poses no great problem for many directions, it is not always straightforward for intermediate directions. The data belonging to the normal mode in Figure 1 are manifestly asymmetric in distribution as will be shown. We plot the data from Figure 1 on a quantile-quantile (QQ) plot in Figure 2 as described by *Fisher et al.* [1987]. Data associated with each mode are rotated to a coordinate system with the mean direction as the pole and are then sorted by magnitude; the inclination values are converted to "coinclinations" (θ) or degrees away from the mean. Then, declination and coinclination values are plotted against expected values based on a Fisher distribution. The spherical symmetry of the Fisher distribution about the pole means that declinations should be uniformly distributed on $[0, 2\pi]$, and it can be shown that $1 - \cos\theta$ should be approximately exponentially distributed. QQ plots should be approximately linear if the designated distribution is a good model for the

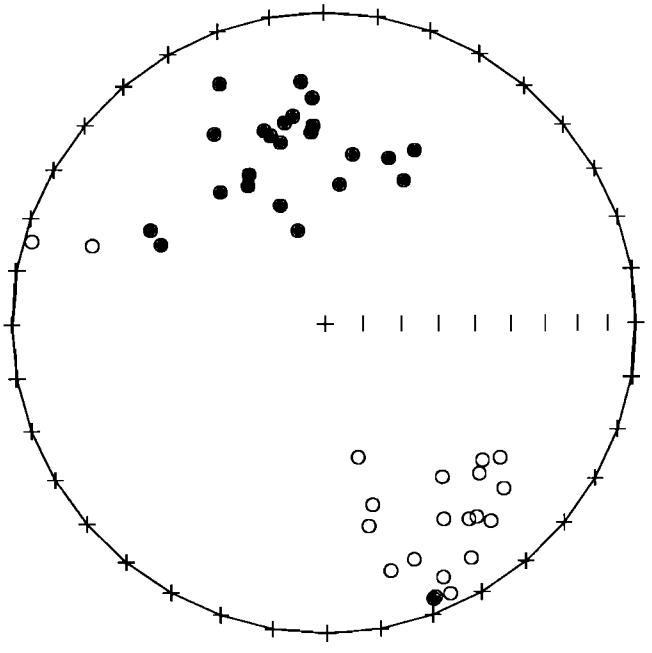

Fig. 1. Equal-area projection of directional data from the KUL section of *Tauxe and Opdyke* [1982] (in stratigraphic coordinates). North is at the top and vertical is at the center of the projection. Solid symbols are lower hemisphere, and open symbols are upper hemisphere.

data. A more quantitative measure of the goodness of fit to the Fisher model is given by the parameters $M_u(V_n) = V_n(\sqrt{n} - 0.567 + 1.623/\sqrt{n})$, for the uniform distribution on the circle and $M_e(D_n) = (D_n - 0.2/n)(\sqrt{n} = 0.26 + 0.5/\sqrt{n})$ for the exponential distribution (here n is the number of data, D_n is the Kolmogorov Smirnov statistic, and V_n is the Kuiper statistic; both statistics are measures of the departure of the empirical cumulative distribution function of the data from the expected theoretical function). Critical values are $M_u = 1.207$ and $M_e = 1.094$ at the 5% probability level (see *Fisher et al.*, [1987]) and are used here to determine whether the data are Fisherian in distribution. The declinations of the normal mode fail the test with $M_u = 1.234$. Such bimodal, asymmetrically distributed data could perhaps be treated with Bingham statistics; however, the modes are not antipodal. In fact, the angle between the modes is approximately 156°, and the dispersion and ellipticity of the two modes appear to be different. Thus the principal assumptions of the Bingham distribution may also have been violated.

The data set shown in Figure 1 is by no means atypical and suggests that the search for a more flexible means of estimating confidence intervals is not unwarranted. Indeed, several articles have appeared recently which suggest various "bootstrap" approaches for the analysis of directional data (see, for example, *Fisher and Hall*, [1989] and references therein) and anisotropy data (*Constable and Tauxe*, [1990]). Bootstrap methods to estimate the confidence in a particular parameter generally proceed as follows [see *Efron*, 1982]. A "pseudosample" is created by randomly drawing data (with replacement) from the original data set. Each pseudosample is the same size as the original data set, and each data point may occur a number of times in a given pseudosample. The parameter of interest is calculated for each pseudosample, and the variability of the parameter for all the bootstrap pseudosamples is used to estimate the confidence interval of

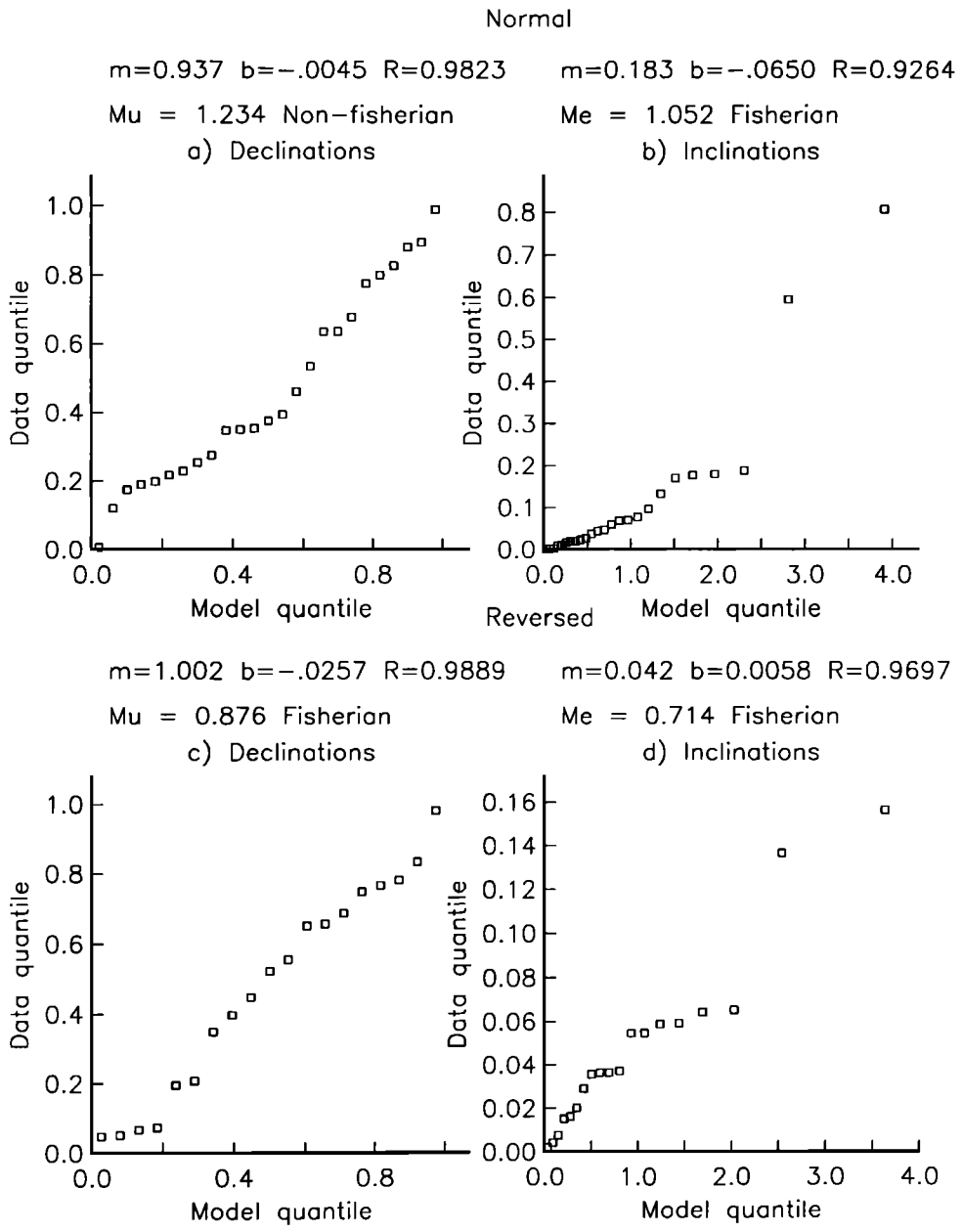

Fig. 2. Quantile-quantile plots of data from Figure 1. Data are divided into two modes and transformed to the mean. Declinations are plotted against a uniform model quantile and inclinations against an exponential model quantile. The slope of the two quantiles is given by m, the intercept by b and the correlation coefficient from a linear regression by R. M_u and M_e are parameters described in the text for discrimination of Fisher distributions. Large values indicate non-Fisherian behavior. The normal mode is plotted on top and the reversed mode below. The declinations of the normal mode are not uniformly distributed and the distribution is therefore not Fisherian.

the parameter. *Davison et al.* [1986] point out that bias can occur in the bootstrap estimates when particular data elements occur more frequently than others when averaged over the entire set of pseudosamples. This problem can be remedied by employing a "balanced" bootstrap whereby a number of replicates of the original data are concatenated into one long "source array". The data are then shuffled, and subsets of the same size as the original data are taken off sequentially to serve as pseudosamples. This ensures that each data point is used the same number of times in the set of bootstrap samples (i.e., it eliminates any potential bias in the random sampling procedure used). In practice we found that "balancing" had no effect on our results.

The advantage of the bootstrap is that it circumvents (or at least postpones) the need for parametric assumptions and substitutes repetitive calculations for analytical solutions. *Watson* [1983] discusses a bootstrap method for calculating confidence intervals for spherically symmetric, unimodal distributions. A slightly different approach for circular and spherical data was taken by *Ducharme et al.* [1985]. *Fisher and Hall* [1989] provide an excellent discussion of the various approaches and advocate a method analogous to "Studentizing" samples. In their implementation, the mean is calculated for each of 2000 "balanced" bootstrap samples and the resulting distribution is contoured and presented graphically as a grey-scale image in equal-area projection.

The nonparametric approach of *Fisher and Hall* [1989] has great appeal and certainly should be followed. However, in practice, paleomagnetists desire a more compact way of expressing confidence regions which necessitates parameterization at some level. Also, tests for discrimination of directions and dispersion (e.g. the test for common mean and fold tests) are essential tools for paleomagnetic analysis. Thus paleomagnetists have been reluctant to venture away from the rigor of tests provided for Fisher distributions such as the fold tests of *Watson and Irving* [1957] and *McFadden and Jones* [1981] and the tests for common mean of *Watson* [1956] and *McFadden and Lowes* [1981]. The approach we outline here is quite similar to that of *Fisher and Hall* [1989] in that we use a balanced bootstrap resampling scheme to "map out" the variability of the mean. The variability in the mean is much less than that in the original data. We can then imagine that a kind of spherical analogue of the Central Limit Theorem is operating so that the population of bootstrap means is asymptotically distributed according to a Kent distribution (the spherical analog of a two-dimensional normal) and estimate the associated confidence ellipses. Although *Hall* [1987] shows that elliptical confidence regions estimated from general bootstrap routines fail to capture deviations from normality, we have examined distributions of bootstrap means from a large number of paleomagnetic data sets and have found them all reasonably well approximated by an ellipse.

We will illustrate our approach using the data set shown in Figure 1. We will use the same notation as in the previous section where the number of data points is n, the data set is denoted $\mathbf{X}$ and the kth unit vector is $\mathbf{X}_k$. We construct pseudosamples using a balanced bootstrap, whereby N_b copies of $\mathbf{X}$ are concatenated together into a source array which is then shuffled by randomly selecting two elements and swapping their positions. This is repeated N_b times. Then, sequential slices of length n are taken as pseudosamples. We denote the lth pseudosample, or "bootstrapped" sample, by $\mathbf{X}_l^*$ and its kth vector by $\mathbf{X}_{lk}^*$. For each bootstrapped sample we wish to calculate an estimate of the mean direction for each of the modes. In order to do this, the data must be separated in some objective fashion into normal and reversed modes. To this end, we compute the eigenparameters of the orientation matrix of the pseudosample. The plane normal to the principal eigenvector separates the data into two hemispheres, each containing a mode of the pseudosample. This plane is used as the criterion for modal separation. Since this equatorial plane is calculated for each bootstrapped sample, intermediate directions may change allegiance from one mode to the other in different bootstrapped samples reflecting a genuine ambiguity with regard to their "proper" place. A mean direction, $\bar{\mathbf{X}}_l^*$, is estimated for each of the modes using equations (1) and (2) as described in the section on Fisher statistics.

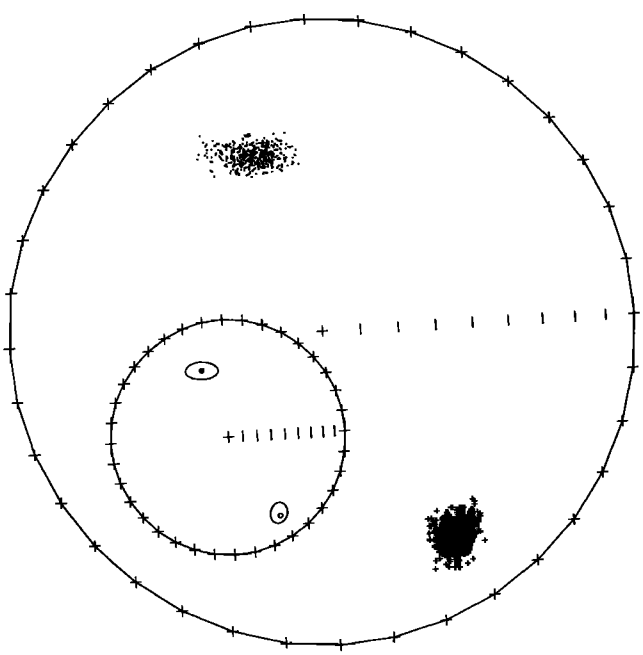

Fig. 3. Equal-area projections of mean directions for 500 bootstrap samples generated from the data of Figure 1. Dots are lower hemisphere projections, and plus signs are upper hemisphere projections. The inset shows the bootstrap confidence intervals for the mean. The projection is in stratigraphic coordinates.

Five hundred bootstrap means generated in this way for the data shown in Figure 1 are plotted in Figure 3. As is most often the case, the population of bootstrap means is approximately elliptical and is even more so when more pseudosamples are taken. In general, the number of bootstrap samples, N_b, should be sufficient to ensure that the distribution of possible means is adequately represented (something like n^2 is desirable according to *Hall* [1988]).

ESTIMATION OF BOOTSTRAP CONFIDENCE REGIONS

Once N_b mean directions have been calculated for each mode based on the bootstrapped samples, we calculate the mean of the bootstrapped means $\bar{\mathbf{X}}^*$, which should very nearly equal $\bar{\mathbf{X}}$, the mean of the original data set. Disagreement may be an indication that N_b is too small. We then compute the orientation matrix $\bar{\mathbf{T}}^*$ and the constrained eigenvectors of Γ for the set of bootstrapped mean directions. For sufficiently large N_b the principal direction of the orientation matrix coincides very nearly with the Fisher mean. Inspection of plots of bootstrap means from a wide variety of paleomagnetic data sets (like those shown in Figure 3) has led us to the conclusion that the distribution of the bootstrap means is nearly always elliptical, suggesting that a Kent model could be applied to each mode of bootstrap means to parameterize the confidence region. The approximate 95% confidence regions can thus be estimated as follows. For the present case we are interested in the distribution of bootstrapped means, so N_b is used instead of n in equation (6) to calculate $\hat{\sigma}_2^2$ and $\hat{\sigma}_3^2$; since we are interested in the confidence in the mean of the original data and not the confidence in the mean of the means, the appropriate value for $g = -2\ln(0.05)/\hat{\mu}^2$. These values in equation (7) give the approximate 95% confidence ellipse for the sample mean direction. The confidence ellipses calculated in this way are shown in the insert to Figure 3.

Testing of Bootstrap Confidence Regions on Simulated Data

Before accepting the bootstrap method as an improvement over the conventional approach to estimating confidence regions, it is desirable to understand how it performs. We tested it on a variety of simulated data sets drawn from Fisher distributions with assorted samples sizes, n, and concentration parameters, κ. A sample simulated from a Fisher distribution is shown in Figure 4*a*. The true mean direction is the vertical. In Figure 4*b*, we plot 100 means calculated from

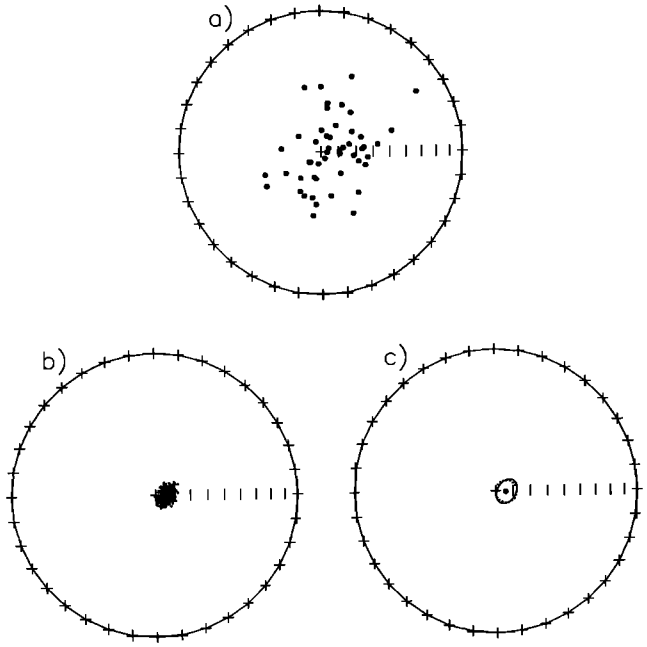

Fig. 4. Equal-area projections. (*a*) A typical Fisher distribution with $\kappa = 10$ and $n = 50$. (*b*) Means of 500 bootstrap samples. (*c*) Mean of the data shown in Figure 4*a* with bootstrap (solid line) and α_{95} (dotted line).

the bootstrap pseudosamples, and in Figure 4*c* we show the elliptical bootstrap confidence region, here called b_{95} (solid line), and the circular region (dotted line) commonly used by paleomagnetists, corresponding to α'_{95} in equation (4).

If the calculated regions are true 95% confidence regions, we would expect that if we performed 100 such simulations then on the average 95 of the calculated confidence regions should include the true mean. (So the sample shown in Figure 4 happens to be one of the approxiamtely 5% of data sets whose confidence regions do not include the true mean.) If on average fewer than 95 enclose the true mean, then the confidence regions are biased too small; if more, then they are too large. One might expect that the elliptical confidence regions will be too small for our simulated data sets, because we are assuming an elliptical form for the underlying distribution, when in fact it is spherically symmetric. We would also expect the bias to decrease as the number of data contributing to the mean increase and we obtain a better sampling of the underlying distribution. We find that this is indeed the case. Figure 5*a* shows the average (and one standard deviation in the mean) number of points lying outside the confidence region as a function of n (for $\kappa = 50$), and 5*b* shows them as a function of κ (for $n = 10$). The confidence regions are computed by four different methods. The solid line at 5 is what one would expect for an unbiased result. The squares represent the number of means lying outside the α_{95} cone of confidence for a Fisher distribution computed using equation (3), the triangles those computed using the approximation in equation (4), while the diamonds are those obtained from the bootstrap method and the assumption that the bootstrap means are drawn from a Kent distribution. The octagons were also obtained by a bootstrap technique, but one in which it was assumed that the bootstrapped means were distributed according to a Fisher distribution rather than the elliptical Kent. It is clear that all the methods except the more accurate means of computing Fisherian α_{95} produce confidence regions that are biased too small. There is no excuse for using α'_{95}; it is not difficult to compute the vastly more accurate approximation found in equation (3), and one can then obtain extremely reliable confidence regions if the distribution underlying the data is Fisherian. Comparing the results for the two bootstrap methods allows us to determine that for an underlying Fisherian distribution the bias is substantially greater for the Kent elliptical regions than for the Fisher circles when the sample size is small (this is to be expected since small samples will show greater departures from spherical symmetry). The bias in both bootstrap estimates is essentially independent of κ for fixed n, but for fixed κ it decreases fairly rapidly with increasing n, so that by $n = 25$ it has reached what we would regard as an acceptable level. It should be noted that one would not normally use the bootstrap method on data unless they exhibit significant departures from an underlying Fisher model. The Fisher simulations shown here thus represent a worse than average case in terms of the expected bias.

Application to Paleomagnetic Data

We are now in a position to compare the confidence regions calculated by the bootstrap method outlined here with

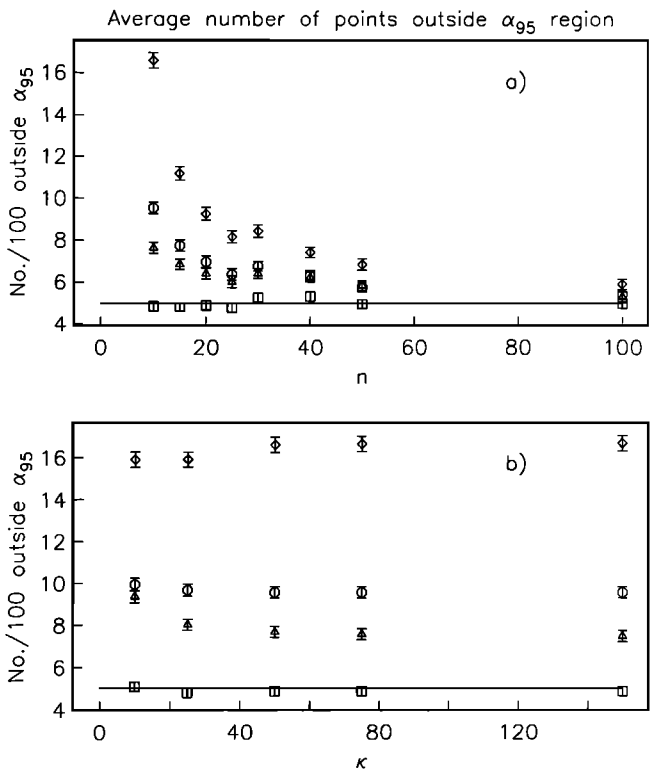

Fig. 5. (*a*) Average number of points lying outside 95% confidence regions as a function of n, the number of data contributing to the Fisher mean based on simulations from Fisher distributions. Solid line represents the unbiased result; squares, give Fisherian confidence cones calculated using equation (3); triangles, approximation to Fisher result given by equation (4); octagons, bootstrap means with an assumed Fisher distribution; diamonds, bootstrap means with an assumed Kent distribution. (*b*) Same as Figure 5*a* but as a function of κ. Error bars are one standard deviation in the mean of the confidence confidence regions calculated for 100 sets of 100 simulations each.

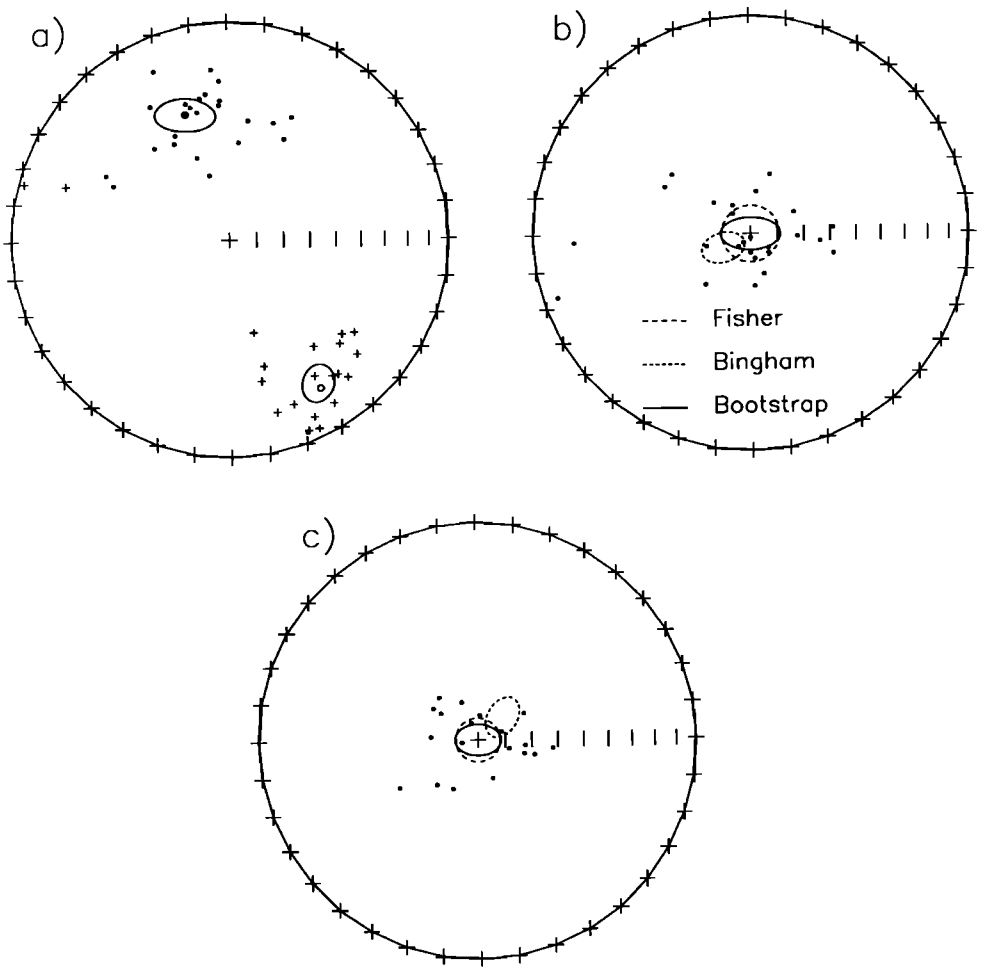

Fig. 6. (*a*) Data from Figure 1 plotted with estimated 95% confidence regions. Projection same as in Figure 1 with conventions of Figure 3. (*b*) Comparison of 95% confidence regions calculated by the method proposed in the text (bootstrap) and by Fisher and Bingham methods. Data are normal mode, rotated such that the eigenvectors of the orientation matrix (Γ) of the set of bootstrap means serve as axes. The center of the diagram is the principal axis and major and minor axes point left-right and up-down respectively. The principal eigenvector for the set of bootstrap means also coincides with the mean of the original data (see text). (*c*) Same as for Figure 6*b* but reversed mode.

confidence regions calculated by standard parametric methods on some "real" data. The data from Figure 1 are shown again in Figure 6*a* with the bootstrap confidence regions. Now, we rotate the normal and reversed data so that the major and minor eigenvectors for each mode lie in the plane of the projection; the principal direction (here parallel to the mean) is therefore at the center of the diagram. Figure 6*b* shows the normal mode with the corresponding confidence regions and Figure 6*c* shows the reversed mode.

The asymmetrically distributed normal data are not well modeled by the Fisher distribution as already noted (see Figure 2); hence the α_{95} cannot reflect the asymmetry (see Table 1 for parameters). The reversed mode, however, has fewer intermediate directions and is compatible with a Fisher distribution (Figure 2); in this case the Fisher and bootstrap confidence regions are quite similar.

The difficulty with the Bingham approach for data sets whose modes are not antipodal and identically distributed is evident from the fact that the mean directions for each mode (plotted at the center of the diagram) are not the same as the direction of the principal eigenvector of the original data set (located at the center of the Bingham confidence regions). In Table 1 we list the confidence regions calculated by both the Bingham and Kent methods for the normal mode data taken alone. Please note that even when each mode is treated separately by the Bingham method, the principal eigenvector of the normal mode does not coincide with the Fisher or vector mean.

The reason for this is straightforward; the mean is computed from the vector sum of the contributing directions and gives equal weight to each datum. In contrast, the principal eigenvector is the axis about which the data would have minimum moment of inertia if they were regarded as point masses. The moment of inertia associated with a point mass about a given axis is given by the product of the mass with the squared perpendicular distance from the axis; thus the principal axis is the one that minimizes the squared distances of the directions from it. In fact, *Watson* [1986] has shown that in some cases the principal direction is a more robust estimate of the true direction than is the mean. We illustrate this here with a simple example. In Figure 7 we show a random sample drawn from a Fisher distribution

TABLE 1. Statistical Parameters for the Normal Mode of Data Shown in Figure 1

Method	$\bar{Dec}$	$\bar{Inc}$	η_{95}	Dec_η	Inc_η	ζ_{95}	Dec_ζ	Inc_ζ
Bootstrap	341.5	39.0	11.9	90.0	20.4	6.3	120.0	44.0
Fisher	341.5	39.1	(α_{95} = 11.0,	∂p = 19.1,	∂m = 12.8)			
Bingham	345.8	40.9	11.9	89.5	15.3	6.5	195.5	45.1
Kent	345.8	40.9	11.6	89.5	15.3	6.3	195.5	45.1

$\bar{Dec}$ and $\bar{Inc}$ are the estimated values for the true mean directions by each technique. Dec_η and Inc_η are the directions of the major semi-axes and Dec_ζ and Inc_ζ are for the minor semiaxes.

with κ = 25 and n = 20. Two extra data points are contributed with directions 85° away from the true direction (vertical). The mean for the data set has a declination of 330.8° and an inclination of 82.0°, while the principal direction is oriented with a declination and inclination of 293.5° and 85.0° respectively. The principal direction is closer to the true direction, hence is less affected by the presence of outliers.

The principal direction may therefore be a more robust estimate (i.e., not as sensitive to outliers) of the paleomagnetic direction of interest than is the mean. It should be noted that when the data distribution is symmetric the mean and principal axis will yield the same direction; however, even the most "ideal" paleomagnetic data sets will exhibit small departures from symmetry, so there will inevitably be differences between the mean and principal directions. Bootstrap confidence regions can be calculated for the principal direction in the same manner as described here for the mean. However, if the data distribution is not well modeled by an ellipsoid (e.g., if it exhibits strong asymmetry in the form of streaking) then it is likely that neither the mean nor the principal axis will provide a good estimate for the direction of interest. *Van Alstine* [1980] suggested the use of the mode as preferable; the mode is the maximum likelihood or most probable direction and thus remains uninfluenced by strong asymmetries in the data. Bootstrapping of modal estimates in the manner described by *Fisher and Hall* [1989] might well be appropriate in such cases.

The normal mode shown in Figure 6*b* is shown again in Figure 8*a* with the horizontal (north) and vertical directions shown as a triangle and a square, respectively, for reference. The Fisher cone of confidence (α_{95}) and the bootstrap ellipse

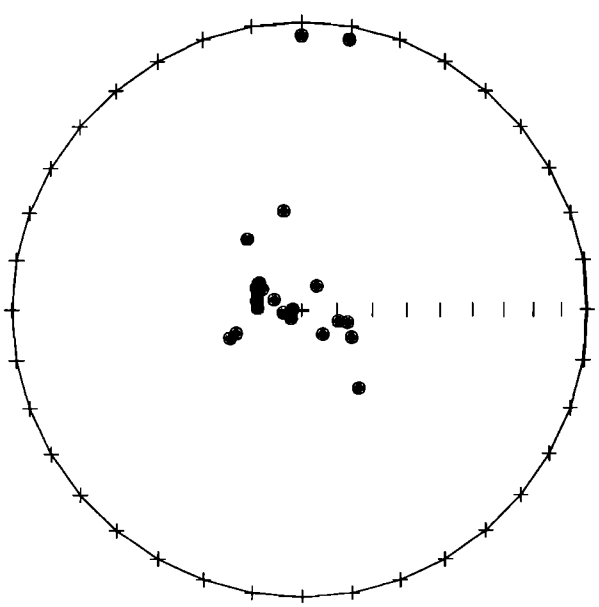

Fig. 7. A simple example illustrating the more robust nature of the principal direction as opposed to the mean (see text).

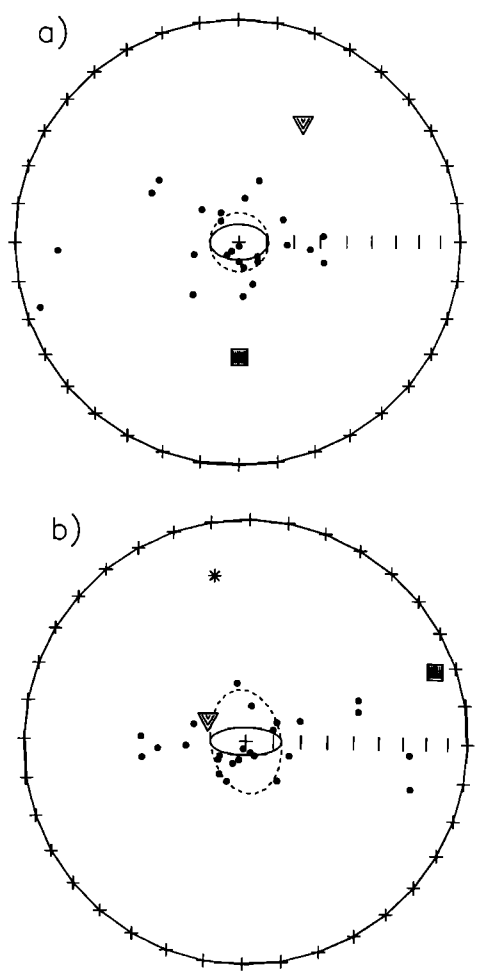

Fig. 8. (*a*) Data from Figure 6*b*, with bootstrap and Fisher confidence ellipses. The triangle is the horizontal north direction, and the square is the vertical. (*b*) Same data as in Figure 8*a* but transformed to virtual geomagnetic poles. The asterisk represents the sampling site. The dashed ellipse is the 95% confidence ellipse calculated from the α_{95} (see text). The major semiaxis is ∂m, and the minor semiaxis is ∂p. The solid ellipse is the bootstrap confidence ellipse calculated by the proposed method.

of confidence (β_{95}) are also shown. The directional data are transformed into virtual geomagnetic poles (VGPs). Since no strict assumptions are made about the underlying distribution of the data, we may compute the confidence regions for the VGPs using the same bootstrap algorithm as for the directions. This liberates us from the uncomfortable necessity of choosing which data set is Fisherian, the directions or the VGPs, when it is often evident that neither meets the necessary requirements. The VGPs associated with the normal mode are plotted in Figure 8*b*, along with the position of the geographic north pole (triangle) and the Greenwich meridian at the equator (square). The sampling site is shown as an asterisk. We plot the bootstrap confidence regions together with the Fisher confidence region distorted in the usual way [e.g. *McElhinny*, 1973)] i.e.

$$\partial m = \alpha_{95}\frac{\cos\theta}{\cos\bar{I}}$$

$$\partial p = \frac{1}{2}\alpha_{95}(1+3\cos^2\theta)$$

where $\bar{I}$ is the inclination of the mean, θ is given by $\tan\bar{I} = 2\tan\theta$, and α_{95} is the angle giving the 95% circle of confidence of *Fisher* [1953] (equation (3) and Table 1). The parameter ∂m is the uncertainty in the paleomeridian (the minor semiaxis), and ∂p is the uncertainty in the paleoparallel (the major semiaxis) of the paleopole. Because of the nature of the asymmetry of the original data distribution, the ellipticity of the distorted Fisher confidence region is perpendicular to that of the actual distribution of VGPs.

COMPARING PALEOMAGNETIC DIRECTIONS

The calculation of confidence regions for paleomagnetic data is largely motivated by a need to compare estimated directions with either a known direction (for example, the present field) or another estimated direction (for example, that expected from a particular paleopole). The first case is straightforward. If the known test direction lies outside the confidence interval computed for the estimated direction, then the estimated and known directions are different at the

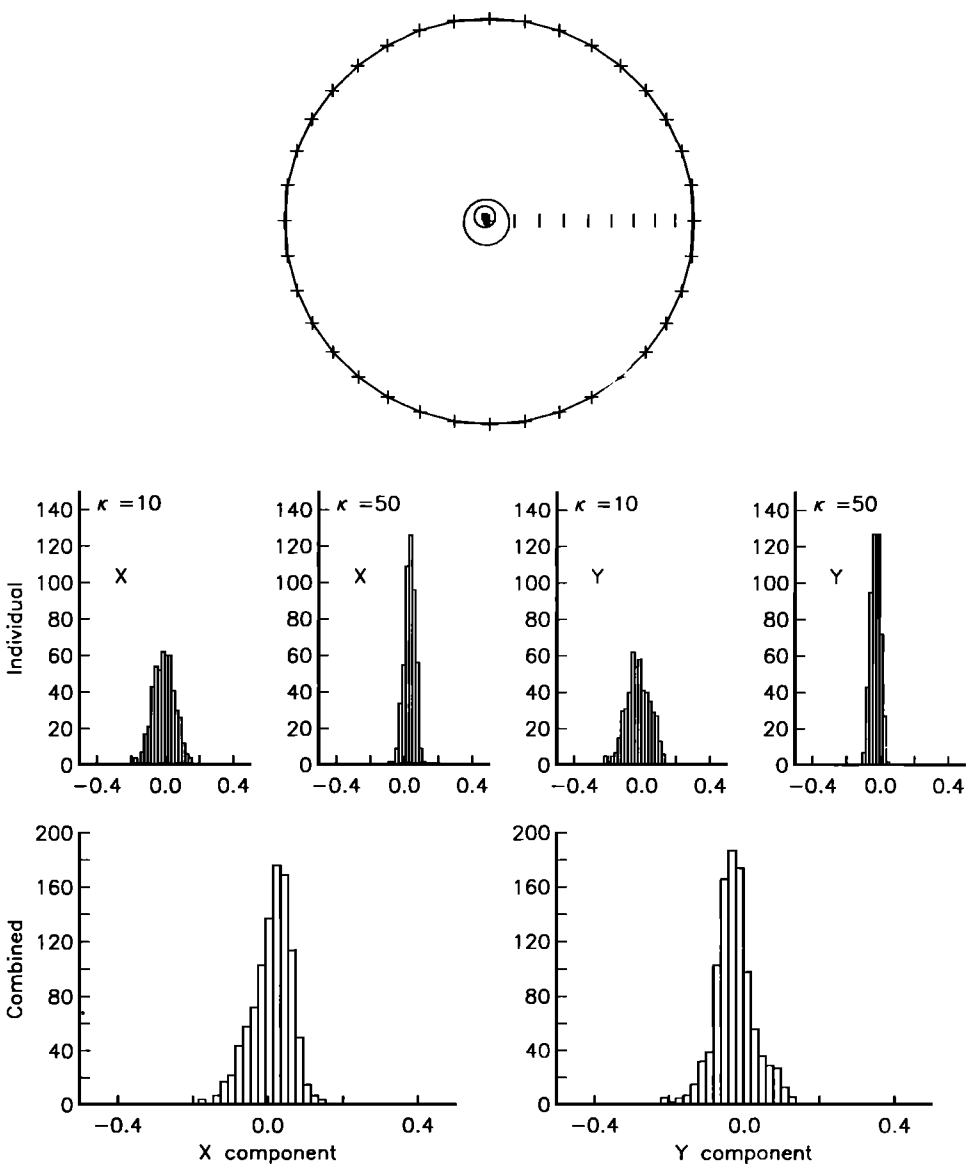

Fig. 9. (Top) equal-area projection of mean and α'_{95} of two simulated data sets with κ of 10 and 50. (Middle) histograms of Cartesian coordinates of mean directions of 500 bootstrapped samples from the two data sets. (Bottom) Cartesian coordinates of bootstrapped means of both data sets combined.

specified confidence level. Similarly, in the second case, if the two confidence regions do not overlap, the two directions must be different at the given level of certainty. Also, when one confidence region includes the mean of the other set of directions, the difference in directions is not significant. However, when the two confidence regions overlap, but neither includes the mean of the other, determining the significance of the difference becomes more difficult. For special cases where the two data sets are Fisher distributed with the same dispersion, criteria that rely on the comparison of the R values for the two data sets separately and together have been developed to test the significance of the difference in data sets [*Watson*, 1956]. The statistic F is calculated as follows:

$$F = (n-2)\frac{(R_1 + R_2 - R)}{(n - R_1 - R_2)} \quad (8)$$

where R_1, R_2, and R are the resultant magnitudes of the first, second, and combined data sets, respectively, and n is the total number of data points in both data sets. F is then compared with the value given in an F table for 2 and $2(n\text{-}2)$ degrees of freedom.

The Watson approach has been extended to include data sets with different dispersions by *McFadden and Lowes* [1981]; (see also *McWilliams*, [1984] and *Demerest* [1983]). In cases where the data are not Fisher distributed, however, discrimination of data sets with similarly overlapping confidence regions can be quite difficult.

The test for common mean asks the question, "Can the means be discriminated from one another?" This question can be cast in terms of our bootstrapping scheme. If the set of bootstrap means is examined, are there two distinct groups or is there just one? We explore this idea by considering first several data sets simulated by drawing at random 20 directions from a Fisher distribution. First, we consider two data sets sharing a common true mean (the vertical) but having values of κ of 10 and 50 respectively. The sample means and associated α'_{95} are shown in Figure 9. Now we generate 500 bootstrap pseudosamples and calculate means for each. The Cartesian coordinates of the set of bootstrap means are plotted as histograms in Figure 9 for each individual population and combined. Because the directions are essentially vertical, the Z components had virtually no variation, being quite close

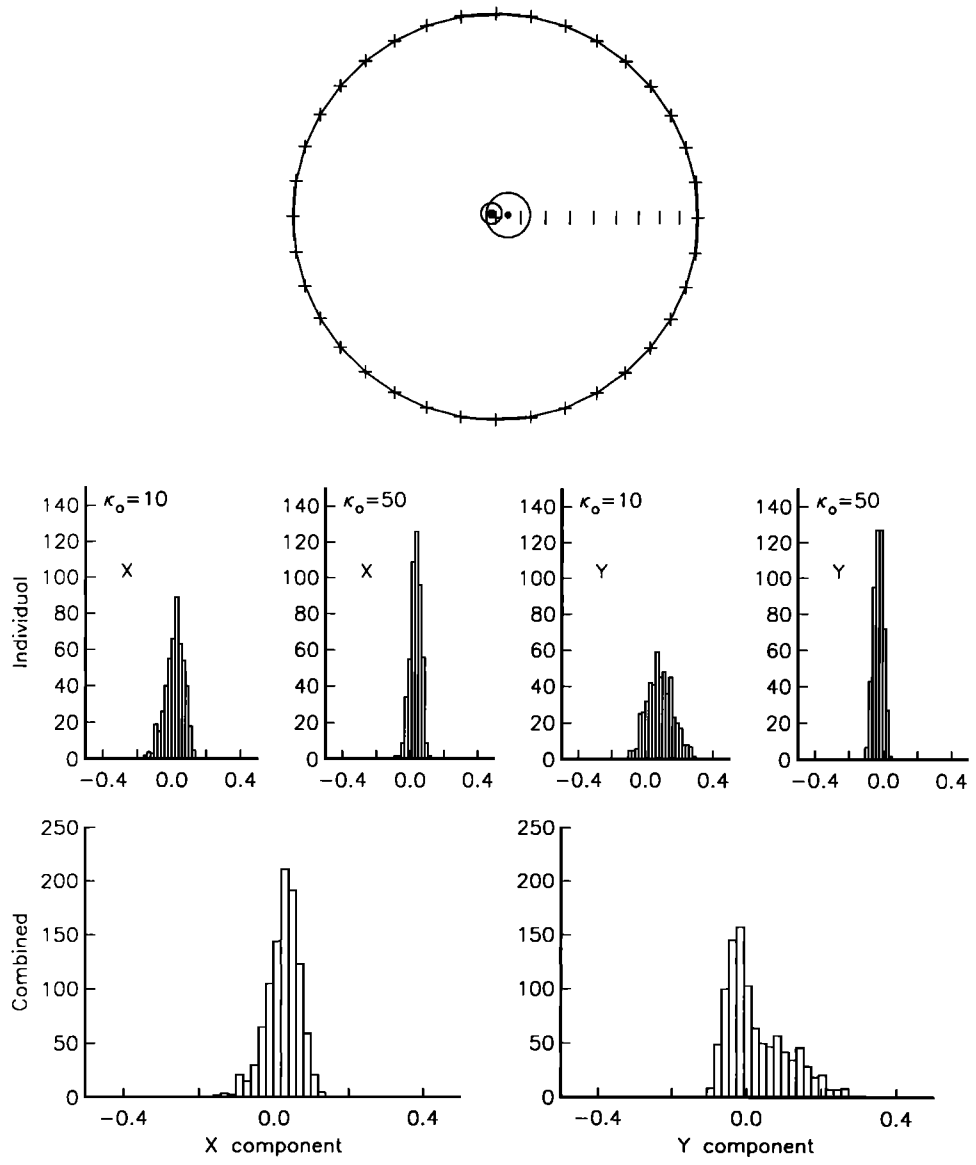

Fig. 10. Same as Figure 9 but data set with κ of 10 rotated such that the mean lies outside the confidence interval of the other data set.

to 1; hence they were left off the figure to alleviate clutter. (Also since the resultant is normalized to 1, only two of the three Cartesian coordinates are independent). The combined histograms are unimodal but have sloping shoulders reflecting the presence of the two different dispersions. In Figure 10 we show the case for two data sets having κ of 10 and 50 as before. Here, the mean of the data set with the higher dispersion lies outside the confidence region of the other, but its confidence region includes the mean of the tighter distribution. These data sets pass the test for common mean described above, and the histograms of the combined sets of bootstrap means are again unimodal.

In Figure 11 we show two data sets, each with $\kappa = 10$. Each mean lies outside the confidence region of the other and the test of *Watson* [1956] must be applied (equation (8) with $F < 2.77$). These data sets pass Watson's test for common mean ($F = 2.03$), and again the histograms of the combined bootstrap means are all unimodal. The data sets shown in Figure 12 (both $\kappa = 10$), however, fail the test for common mean ($F = 3.9$). In this case, the histograms of the combined sets of Cartesian coordinates of the bootstrap means has two peaks consistent with the presence of two directions. There will of course be cases very close to the cutoff value in which the histogram test will be equivocal. For such cases, a more sophisticated approach is necessary. It is well known that the shape and number of modes observed in a histogram can be strongly dependent on the number and position of the bins selected. What is really required is an objective method for determining the number of modes (hence the number of directions) in the probability density function underlying the distribution of bootstrap means. The histograms of Cartesian components just provide a relatively crude estimate of this. *Donoho* [1988] describes a rigorous statistical test to determine the minimum number of modes required by the underlying density function. His test has the merit of being independent of the statistical distribution underlying the data and is ideal for the application required here. However, the computational algorithm has not yet been implemented. The computations involved are not trivial and increase rapidly with the number of modes. Until the Donoho test can be implemented, there will be ambiguous cases.

We illustrate the bootstrap approach for a test for common

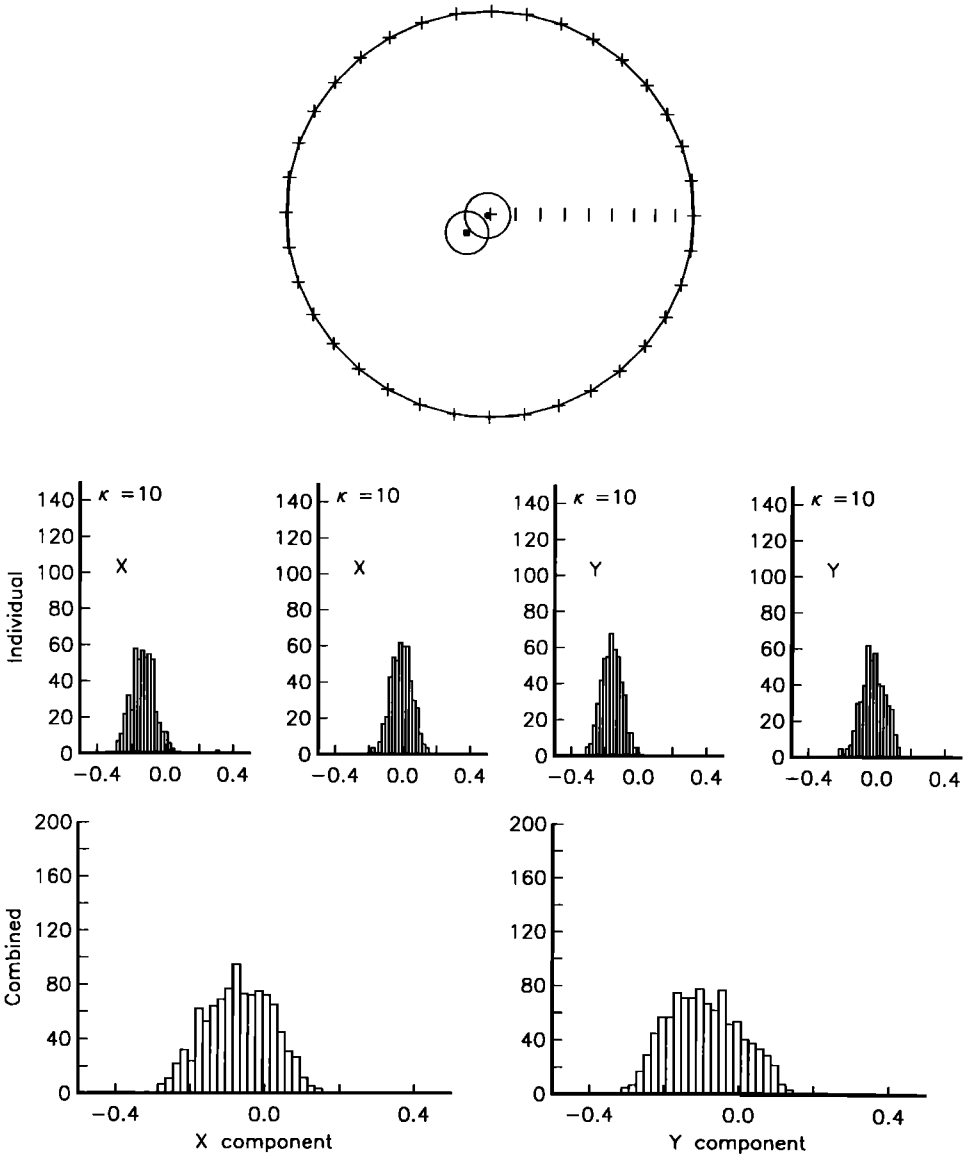

Fig. 11. Same as Figure 9 but both data sets have κ of 10. One is rotated such that the mean of each data sets lies outside the cone of confidence of the other. The data sets pass the test for common mean.

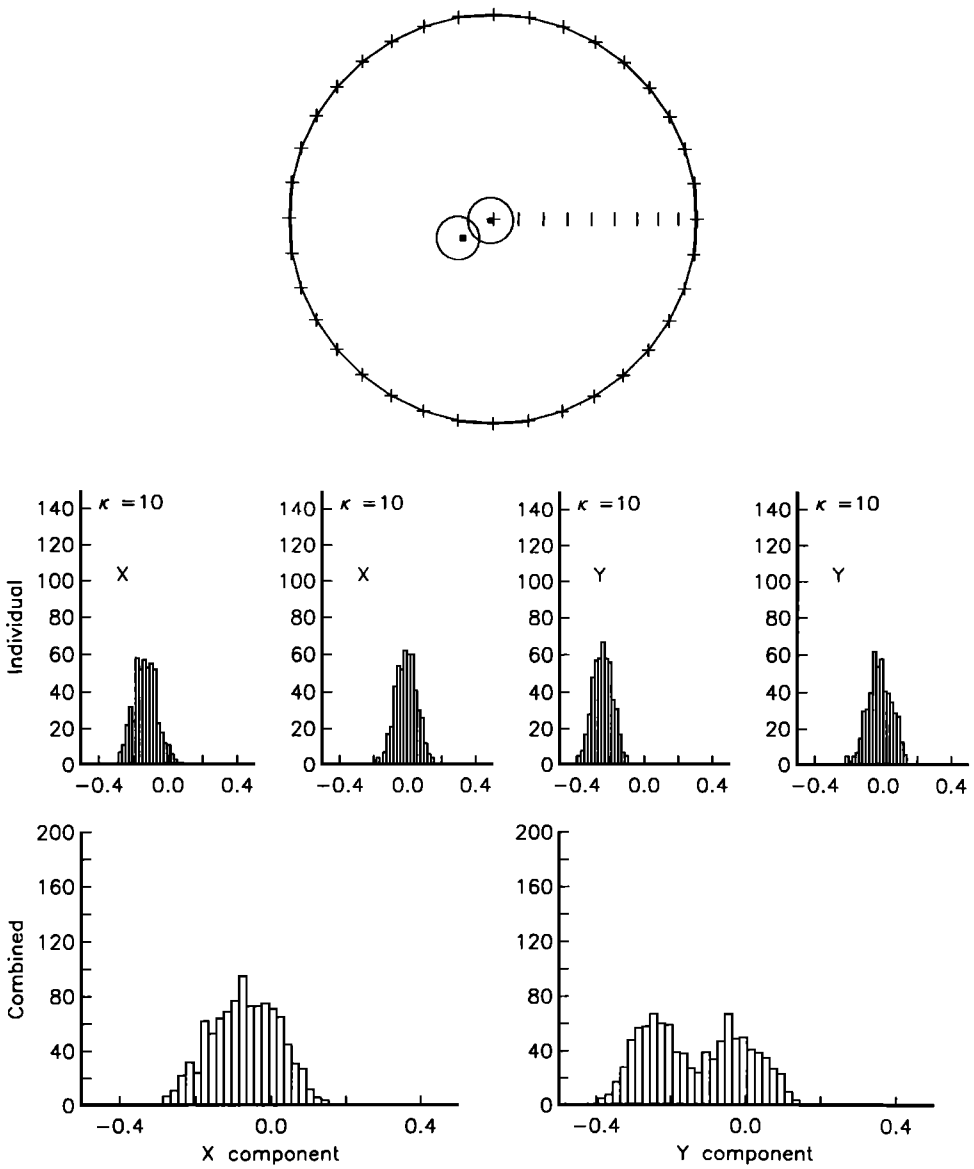

Fig. 12. Same as Figure 11 but the data sets fail the test for common mean.

mean with a data set from the Mojave block recently published by *MacFadden et al.* [1990]. These data are shown plotted on an equal-area projection in Figure 13. The first question that might reasonably be asked of the data is: Are these data Fisherian? A quantile-quantile plot of the data (Figure 14) shows that the reversed mode fails the test for Fisher distribution. A second question that springs to mind is: Are the reversed and normal modes antipodal? In other words: do the combined data sets of the normal mode and the antipodes of the reversed mode pass the test for common mean? Since the reversed mode is not Fisherian, the test is not clear cut. On the basis of the overlap of the Fisher confidence intervals for the means of the normal data and the antipodes of the reversed data (shown in Figure 15*a*), *MacFadden et al.* [1990] claimed that they were; i.e., the data were presumed to pass the test for common mean. However, as we noted earlier, simple overlap of confidence regions is inconclusive. The bootstrap confidence region of the antipode of the reverse mode just grazes that of the normal mode (see Figure 15*b*). In Figures 15*c* and 15*d* we use our bootstrap approach to examine the data and plot histograms of the Cartesian coordinates of the bootstrap mean directions as before. In Figure 15*d*, the coordinates of the reversed mode are transposed to the corresponding antipodes. If the normal and reverse directions are distinct, then we should be able to identify two distinct peaks or modes in at least one of the histograms of Figure 15*d*. It certainly looks as though this is possible for the Y component (the histogram shows twin peaks, even when the antipodes of the reversed data are used), strongly suggesting that these data fail the reversals test.

On the basis of the data shown in Figure 13, *MacFadden et al.* [1990] postulate a rotation of the Mojave Block of some 20°. We calculate an expected direction based on a mean paleomagnetic pole calculated from the eight poles of comparable age, (early Miocene (12 Ma) to latest Oligocene (24 Ma)), compiled by *Irving and Irving* [1982]. We do not use the pole based on running averages (used by MacFadden et al.), because it is contaminated by data from completely different ages; however, the expected directions calculated for the Mojave Block are similar to those used by MacFadden et al. Since the data fail the reversals test and the reversed data

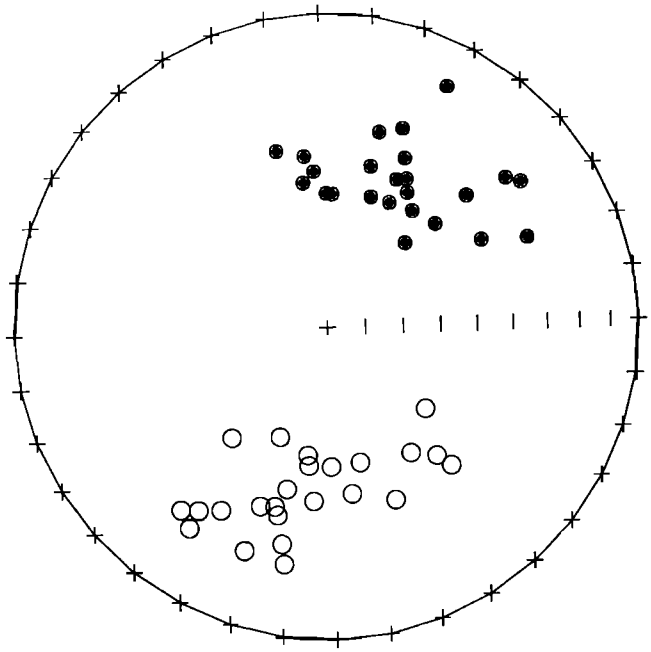

Fig. 13. Data of *MacFadden et al.* [1990] plotted on equal-area projections (see caption for Figure 1).

are actually quite close to the expected direction, we feel that the significance of a 20° rotation has been overestimated.

In fact, one could use the bootstrap approach illustrated here to address the problem of rotation. One would use the directions predicted by poles chosen to represent the "expected direction" as one data set and the (unimodal) study data set as another. One would then bootstrap each data set separately and put the Cartesian coordinates of the bootstrap means together on a single diagram. If several modes are evident, then two directions can be discriminated.

APPLICATION TO THE PALEOMAGNETIC FOLD TEST

In the classic fold tests of *Watson and Irving* [1957], and *McElhinny* [1964], the estimated Fisher precision parameter k is calculated for the data both before (k) and after (k') correction for bedding tilt. The question of whether or not the directions are better grouped after tilt correction is addressed by an F test comparing the ratio k'/k to the value from the F tables for a given number of degrees of freedom. *McFadden and Jones* [1981] show that this approach is invalid because, generally speaking, if one of the data sets (say before tilt correction) is Fisher distributed then the rotation would distort the second data set out of a Fisher distribution; hence the assumptions for calculating k are violated. However, the proposed modification assumes that a fold test involves a single population which has been split into two populations, each Fisher distributed with the same k as the original population. Unfortunately, folds are often curved surfaces, and there are several structural corrections even for sites of the same "limb". Thus the number of "limbs" approaches the number of sites, and the McFadden and Jones test becomes cumbersome.

The difficulties inherent in the fold test are illustrated in Figure 16 (data from *McCabe et al.*, [1983]). Two limbs of a broad fold were sampled. Each site has its own tilt correction because of the curved nature of the fold. The directions from the two fold limbs are corrected for bedding attitude in steps, as was done by *McCabe et al.*, [1983]. We have used the steps of 0%, 30%, 60%, and 100% unfolding for the present illustration. The only group which is Fisherian is that for 60% unfolding (see Table 2), and it seems that no parametric fold tests are applicable. We attempt here to develop tests using bootstrapped parameters that may have broader applicability than those based on Fisher distributions.

We suggest two approaches to the problem; the first is analogous to the classic fold test in that it relies on the degree of dispersion present in the data set as a function of the degree of unfolding, while the second is similar to the technique used by McFadden and Jones in that it looks for differences in the mean directions.

The complete data set at various degrees of unfolding was subjected to the bootstrap analysis described herein. The data are plotted on equal-area projections in Figure 16 in a coordinate system given by Γ, the principal axes of the bootstrap means. Below each stereonet we plot histograms of the eigenvalues of the orientation matrix obtained for each pseudosample (the τ_{1i}^*). The eigenvalues for the set of bootstrap means are listed in Table 3. As the data become better grouped, the histogram of the principal eigenvalues becomes narrower and better defined and it moves closer to 1. The shape of the intermediate and minimum histograms also change; the distinction between the major and minor eigenvalues (τ_2 and τ_3 respectively) becomes blurred as the data distribution becomes more symmetric, and the intermediate eigenvalues move closer to zero.

The eigenvalues $\bar{\tau}_2^*$ and $\bar{\tau}_3^*$ can be taken as estimates of the variances about the mean in the directions of the major and minor axes. If we assume that the components of the mean in the direction of $\bar{\gamma}_2^*$ and $\bar{\gamma}_3^*$ are normally distributed (this is asymptotically true according to *Kent* [1982]), then the ratio of the estimated variances in each direction during different degrees of unfolding can be reasonably expected to be F distributed. Hence the fold test as formulated here is also an F test with the ratio of the eigenvalues and the appropriate degrees of freedom. If, for instance, the ratio of the eigenvalues $\bar{\tau}_2^*$ before and after unfolding is greater than the F value for the appropriate number of degrees of freedom at some confidence level, then the hypothesis that the components of the means in the direction of γ_2^* are drawn from the same distribution is rejected at that confidence level. Thus we could compare the ratios of the $\bar{\tau}_2^*$ and $\bar{\tau}_3^*$ for the various degrees of unfolding (shown in Figure 17) in order to test for significance of the difference in clustering of the data sets. Also shown is the traditional fold test for comparison. The $\bar{\tau}_3^*$ parameters do not change much with unfolding in this case. The $\bar{\tau}_2^*$, however, behave very much like the k in the traditional approach. The slightly lower k values result from the fact that k "averages" $\bar{\tau}_2^*$ and $\bar{\tau}_3^*$ to some extent.

In a different approach to the fold test we could investigate whether the mean directions from each limb of the fold can be discriminated at the various stages of unfolding. This would again be done as described for the test for common mean. The bootstrap means from each limb of the fold would be combined and plotted in the form of a single histogram for each Cartesian component. If the fold correction improves the grouping we should be able to find the degree of unfolding at which each histogram has only a single distinguishable mode, and the dispersion about that mode is a minimum.

The best approach to determining whether there is a

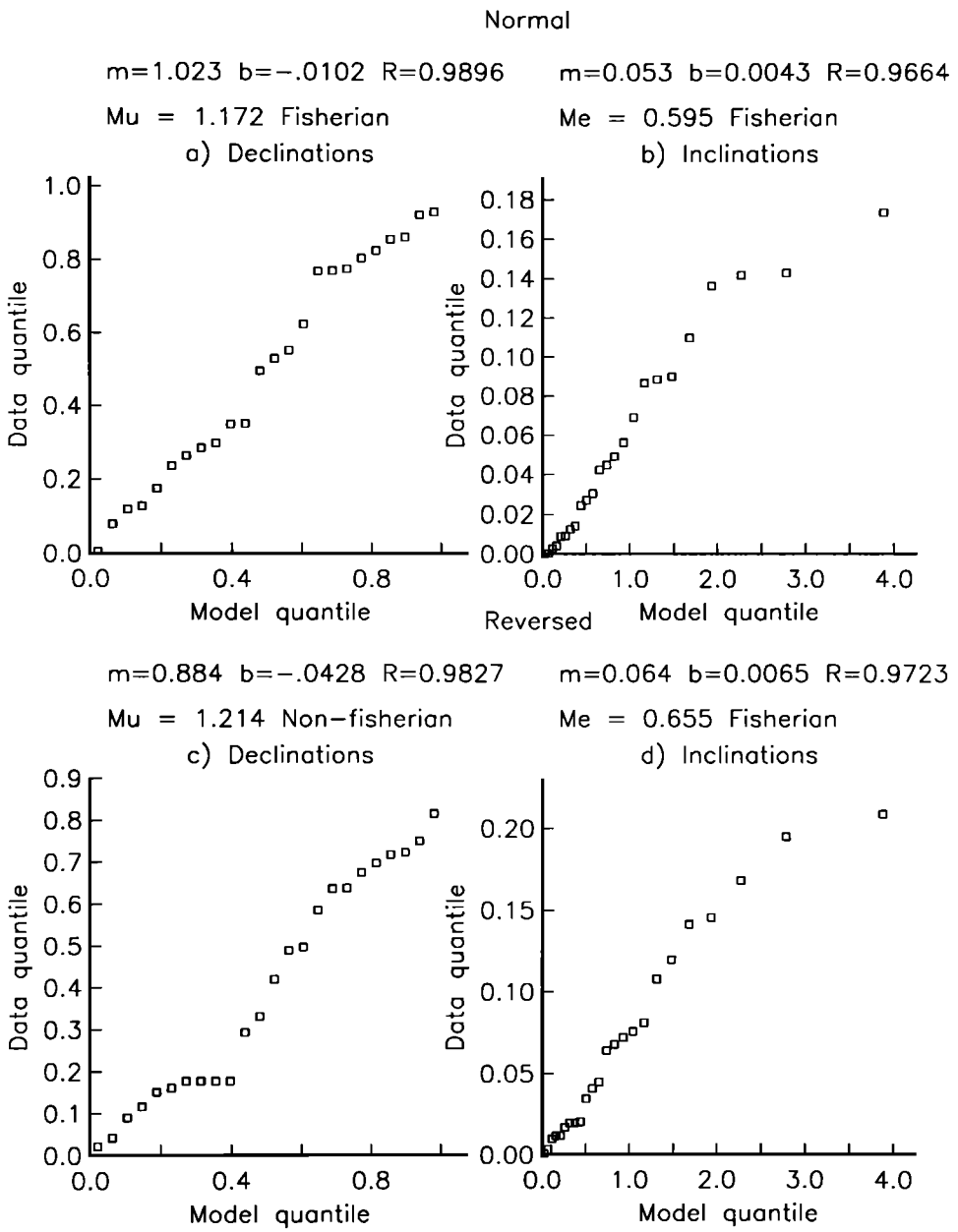

Fig. 14. Quantile-quantile plot of data in Figure 13. See explanation of Figure 2.

positive result to the fold test will almost certainly involve a combination of these two approaches. The method just described may be used to determine the optimum degree of unfolding that gives the most tightly grouped data as determined from the histograms of bootstrap means. Then the first method may be used to test whether there is a significant improvement in the clustering of the data after unfolding. To pass this fold test, the optimum degree of unfolding must be 100% and there must be a significant F value for at least one of the two quantities $\bar{\tau}_2^*/\bar{\tau}_2^{*\prime}$ and $\bar{\tau}_3^*/\bar{\tau}_3^{*\prime}$.

Discussion and Conclusions

Naturally, there are costs and benefits in adopting or rejecting various parametric assumptions. The advantage of making parametric assumptions, in particular, Fisherian ones, is that a data set may be described in terms of a few parameters (the direction of the mean and the concentration parameter) and this description is very informative. Also, there have been several extremely useful tests developed which allow discrimination of directions and dispersions for pairs of Fisher distributed data sets. The disadvantage of parametric assumptions is that many paleomagnetically interesting data sets are not Fisherian and the Fisher inferences can be misleading; tests based on flawed assumptions may lead to flawed interpretations. On the other hand, the advantage of a completely nonparametric bootstrap (such as that of *Fisher and Hall* [1989]) is that statistical inferences can be made without the concern for the distortions possible from parametric assumptions. The disadvantage is that a fairly sophisticated computer is required and that compact tabulation of parameters describing the data set is impossible. Indeed, each data set must be used in its entirety in order to make

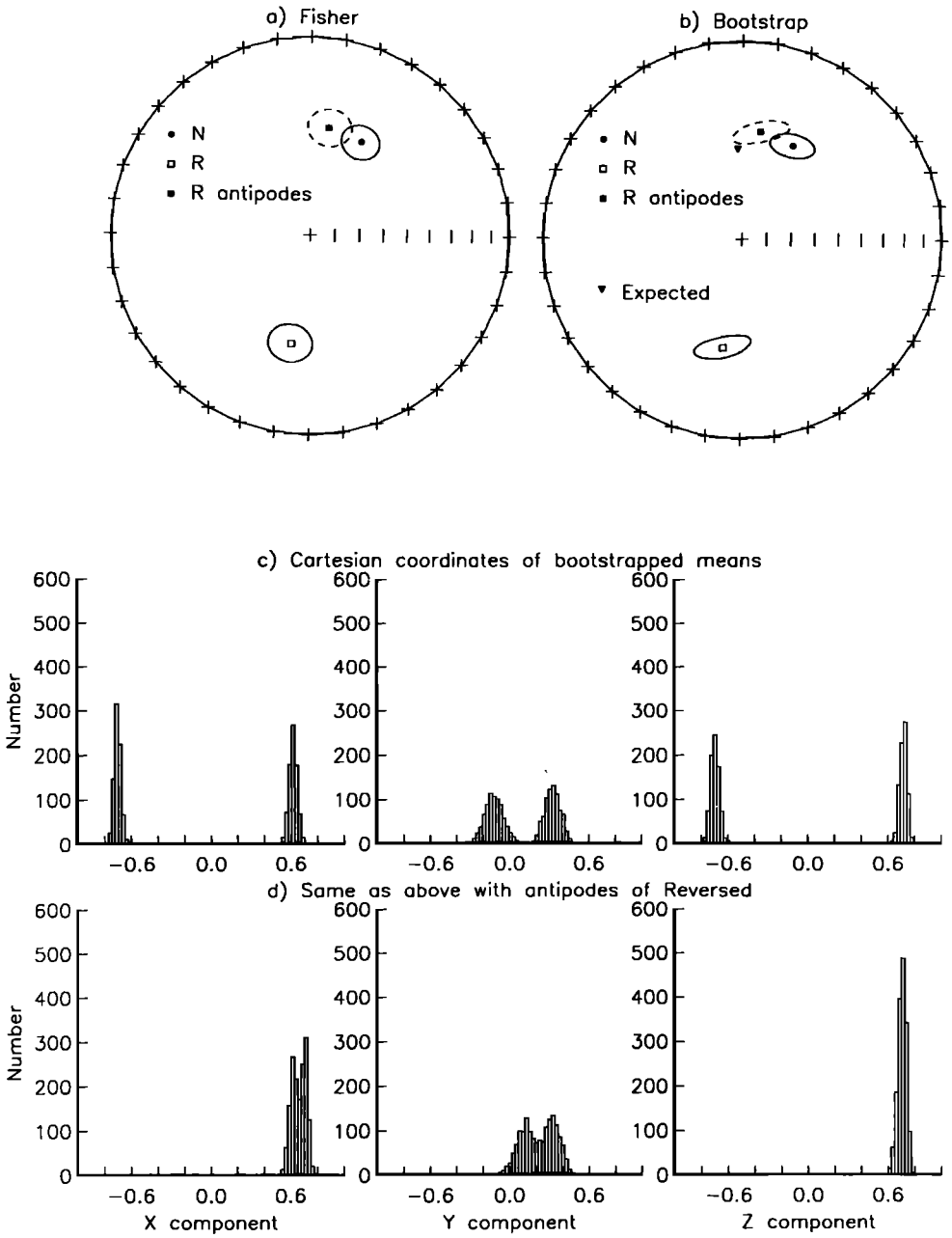

Fig. 15. (*a*) Means and Fisher α'_{95} for data shown in Figure 13. (*b*) Means and bootstrapped confidence regions as described in text. (*c*) Cartesian coordinates of directions of the bootstrapped means of the data shown in Figure 13, (N_b = 800). (*d*) Same as in Figure 15*c* but with the antipodes of the reversed mode.

comparisons, making some sort of international archive for all published paleomagnetic data essential. The question then is, when, if ever, do the benefits of bootstrapping outweigh the costs?

Before such a question can be seriously considered, the assumptions, benefits, and penalties must be clearly understood. To this end we review briefly what our bootstrap procedures are. Then we discuss the pros and cons of the method and conclude with recommendations on when this approach may be advantageous. The procedure for calculating approximate 95% confidence intervals is as follows:

1. Draw a random sample (with replacement) from the original data of the same size (n) to serve as a pseudosample.
2. Repeat step one a large number of times (say n^2).
3. Calculate the eigenparameters of the orientation matrix for each pseudosample and separate the data into two modes using the plane perpedicular to the principal eigenvector as the criterion for separation.
4. Calculate a mean for each mode for all pseudosamples.
5. Calculate approximate ellipses of 95% confidence for the set of bootstrapped means for each mode by assuming that they are Kent distributed.

This is not a balanced bootstrap. We found in our simulations that balancing was much slower, required far more memory, and produced identical results to the procedure outlined above, provided the number of pseudosamples was large. Given the comparatively small memories on most personal computers, we favor a larger number of

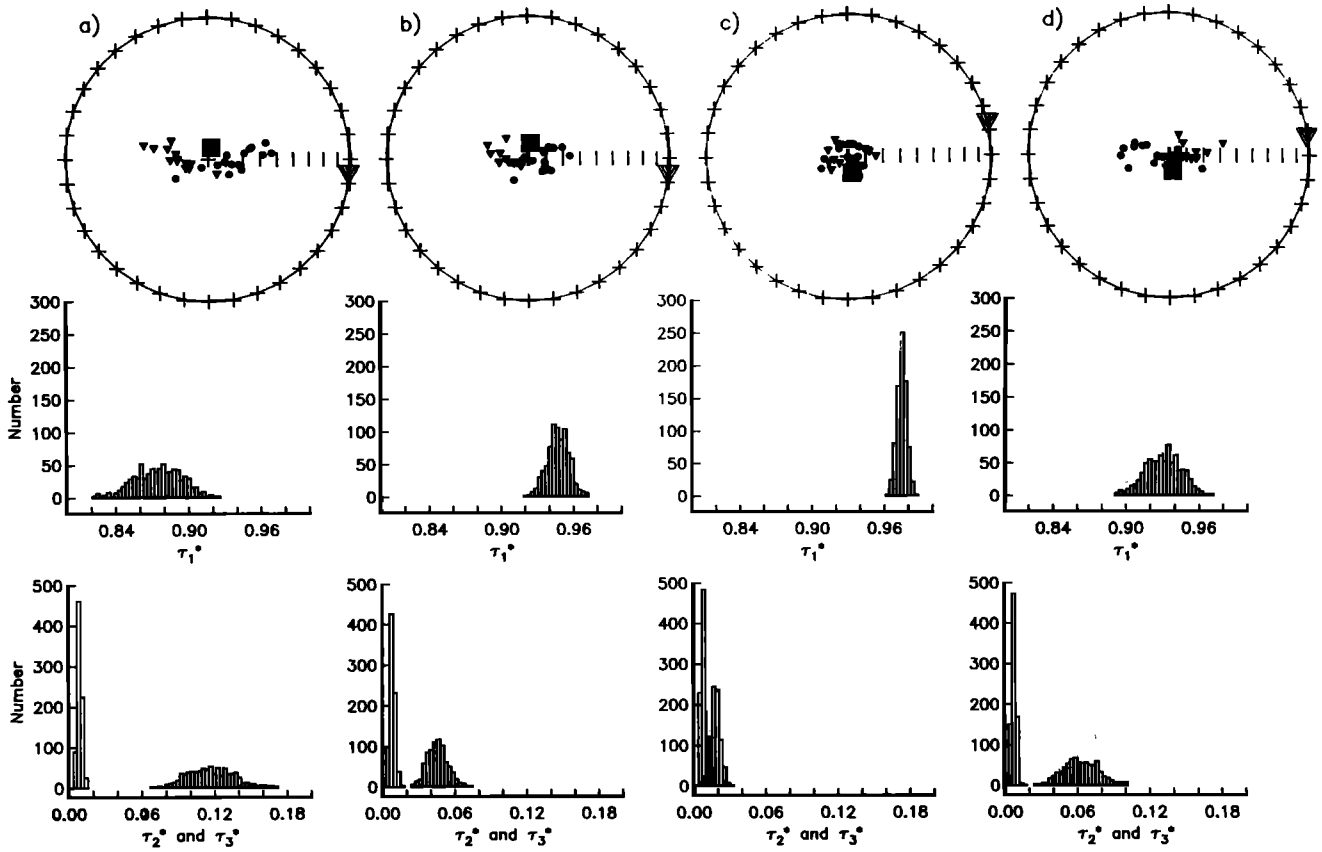

Fig. 16. Fold test for data of *McCabe et al.* [1983]. Circles are equal-area projections in sample coordinates (see Figure 6*b* and 6*c*). Circles are from the northern limb and triangles are from the southern limb. For each step of unfolding, data are treated by bootstrap technique described in the text with 800 bootstrap samples. Counted are the eigenvalues for each bootstrap sample (τ_i^* in our notation).

pseudosamples over balancing, based on the experience that our random number generator is not biased and the resulting bootstrap is therefore already well balanced.

The final step above necessitates a parametric assumption about the distribution of the means. The nonparametric bootstrap of *Fisher and Hall* [1989] expresses the statistical behavior of a given data set by means of a grey-scale contoured image of the bootstrap means. However, it is often desirable to have a more compact representation of this behavior which can be shown in tabular form. Some parametric assumption must be made in order to provide such a representation. On the basis of examination of a wide assortment of paleomagnetic data sets and bootstrap means generated therefrom, we feel that the assumption of approximately elliptical contours of the confidence intervals is a reasonable compromise. Thus we have not avoided parametric assumptions entirely, we have just postponed their application to the set of means, which are more likely to be "well behaved" than the directions.

TABLE 2. M_u and M_e values for the Data in Figure 1

Percent Unfolding	M_u	M_e	Fisherian?
0	1.915	0.507	no
30	1.575	0.468	no
60	1.088	0.468	yes
100	1.475	0.851	no

The advantages of our method are several. The procedure automatically divides data sets into modes and the data need not be edited "by hand". No a priori decisions must be made as to the allegiance of a particular intermediate point to a particular mode as this may change from pseudosample to pseudosample. Furthermore, the technique applies equally well to directions or pole positions and generates confidence intervals which more accurately reflect elliptical data distributions than do Fisher confidence intervals. When applied to the principal directions of the pseudosamples, as opposed to the means, the resulting estimate of the true direction is more robust in the presence of outliers than is the mean. In addition, by examining histograms of the Cartesian coordinates of the bootstrapped means (or principal directions) of two data sets, the presence of a common mean can in most cases be unambiguously determined without resorting to parametric assumptions which often do not apply. Finally, using the eigenvalues of the average orientation matrix instead of κ in the fold test eliminates the discomfort of calculating concentration parameters intended for spherically symmetric distributions on grossly asymmetrical data sets.

There are also several disadvantages of our bootstrap technique. First of all, based on the results of our simulation experiments, bootstrapped confidence regions are in general too small, and for data sets smaller than about 25 (in each mode) the effect is probably unacceptably large. Of course

TABLE 3. Average Eigenvalues for the Data Shown in Figure 16

Percent Unfolding	$\bar{\tau}_1^*$	σ^2	$\bar{\tau}_2^*$	σ^2	$\bar{\tau}_3^*$	σ^2
0	0.875	0.020	0.117	0.019	0.009	0.002
30	0.947	0.009	0.045	0.008	0.008	0.002
60	0.9745	0.004	0.018	0.004	0.007	0.002
100	0.930	0.015	0.063	0.015	0.007	0.002

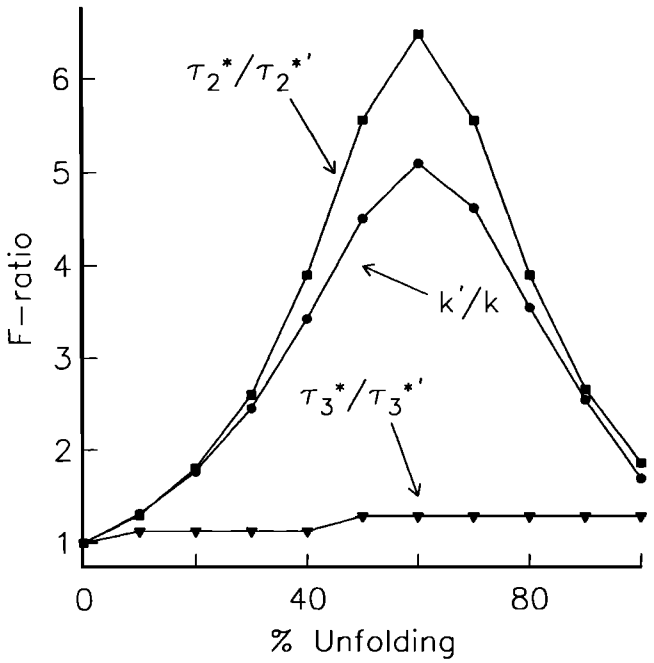

Fig. 17. Plots of various parameters controlled by dispersion of the data. Primed values are for data corrected for tilt.

the α'_{95} calculated with the usual approximation (equation (4)) are also too small, even if the data are drawn from a Fisher distribution. Second, each determination of a bootstrap confidence ellipse (using thousands of pseudosamples) takes about a minute on our laboratory computer (a Compaq 386/25 running under Xenix). A version which runs on a Macintosh II (MacBoot) takes a few minutes. We anticipate that a PC-AT version (DosBoot) will take somewhat longer, and so bootstrapping is certainly slower than calculating Fisher or Kent confidence parameters. Also, a computer with a fairly hefty memory is required (at least 4Mb), but memory is getting cheaper every month and this should not be a major concern. A third disadvantage is that the procedure for separating normal and reverse mode data assumes that the modes are more than 90° apart and are fairly discrete. Girdle distributions will fail miserably under this method.

Although this technique is much more automated and robust than the standard approaches, the user is encouraged (nay, emplored) to examine the data graphically, perhaps using the technique of *Fisher and Hall* [1989]. Quantile-quantile plots can also be useful, and if the Fisher shoe fits (as it often does), by all means wear it. Our technique performs nearly as well as the Fisher approach on Fisherian data but is much slower and is not necessary. Finally, data sets smaller than a few dozen are not suitable for bootstrapping.

Acknowledgments. It is our pleasure to acknowledge useful conversations with R. Parker, J. Tarduno and G. Watson. We are particularly grateful for the careful and extremely helpful reviews by M. Debiche and P. McFadden. This work was bootstrapped onto NSF grant EAR 90-17157 to L.T. and C.C.

REFERENCES

Bingham, C., Distributions on the Sphere and Projective Plane, Ph. D. thesis, Yale University, New Haven, Conn., 1964.

Bingham, C., An antipodally symmetric distribution on the sphere, *Ann. Stat.*, *2*, 1201–1225, 1974.

Constable, C., and L. Tauxe, The bootstrap for magnetic susceptibility tensors, *J. Geophys. Res.*, *95*, 8383–8395, 1990.

Davison, A. C., D. V. Hinkley and E. Schechtman, Efficient bootstrap simulation, *Biometrika*, *73*, 555–566, 1986.

Demarest, H. H., Error analysis for the determination of tectonic rotation from paleomagnetic data, *J. Geophys. Res.*, *88*, 4321–4328, 1983.

Donoho, D. L., One-sided inference about functionals of a density, *Ann. Stat.*, *16*, 1390–1420, 1988.

Ducharme, G. R., M. Jhun, J. P. Romano, and K. N. Truong, Bootstrap confidence cones for directional data, *Biometrika*, *72*, 637–645, 1985.

Efron, B., *The Jackknife, the Bootstrap and Other Resampling Plans*, SIAM Reg. Conf. Ser. App. Math, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1982.

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

Fisher, N. I., and P. Hall, Bootstrap confidence regions for directional data, *J. Am. Stat. Assoc.*, *84*, 996–1002, 1989.

Fisher, R. A., Dispersion on a sphere, *Proc. R. Soc., London Ser. A*, *217*, 295–305, 1953.

Hall, P., On the bootstrap and likelihood based confidence regions, *Biometrika*, *74*, 481–493, 1987.

Hall, P., Theoretical comparison of bootstrap confidence intervals, *Ann. Stat.*, *16*, 927–953, 1988.

Irving, E., and G. A. Irving, Apparent polar wander paths Carboniferous through Cenozoic and the assembly of Gondwana, *Geophys. Surv.*, *5*, 141–188, 1982.

Kent, J. T., The Fisher-Bingham distribution on the sphere, *J. R. Stat. Soc., Ser. B*, *44*, 71–80, 1982.

MacFadden, B. J., M. O. Woodburne, and N. D. Opdyke , Paleomagnetism and Neogene clockwise rotation of the Northern Cady Mountains, Mojave Desert of Southern California, *J. Geophys. Res.*, *95*, 4597–4608, 1990.

Mardia, K. V., and P. J. Zemroch, Table of maximum likelihood estimates for the Bingham distribution, *J. Statist. Comput. Simul.*, *6*, 29–34, 1977.

McCabe, C., R. Van der Voo, D. R. Peacor, C. R. Scotese, and R. Freeman, Diagenetic magnetite carries ancient yet secondary remanence in some Paleozoic carbonates, *Geology*, *11*, 221–223, 1983.

McElhinny, M. W., Statistical significance of the fold test in paleomagnetism, *Geophys. J. R. Astron. Soc*, *8*, 338–340, 1964.

McElhinny, M. W., *Paleomagnetism and Plate Tectonics*, Cambridge University Press, New York, 1973.

McFadden, P. L., The best estimate of Fisher's precision parameter κ, *Geophys. J. R. Astron. Soc.*, *60*, 397–407, 1980.

McFadden, P. L. and D. L. Jones, The fold test in paleomagnetism, *Geophys. J.R. Astron. Soc.*, *67*, 53–58, 1981.

McFadden, P. L. and F. J. Lowes, The discrimination of mean directions drawn from Fisher distributions, *Geophys. J. R. Astron. Soc.*, *67*, 19–33, 1981.

McWilliams, M. O., Confidence limits on net tectonic rotation, *Geophys. Res. Lett.*, *1*, 825–827, 1984.
Onstott, T. C., Application of the Bingham distribution function in paleomagnetic studies, *J. Geophys. Res.*, *85*, 1500-1510, 1980.
Scheidegger, A. E., On the statistics of the orientation of bedding planes, grain axes and similar sedimentological data, *U.S. Geol. Surv. Prof. Pap.*, *525-C*, C-164-C-167, 1965.
Schneider, D. A., and D. V. Kent, The paleomagnetic field from equatorial deep-sea sediments: axial symmetry and polarity symmetry, *Science*, *242*, 252–256, 1988.
Tarling, D. H., *Paleomagnetism*, Chapman and Hall, London, 1981.
Tauxe, L., and N. D. Opdyke, A time framework based on magnetostratigraphy for the Siwalik sediments of the Khaur area, Northern Pakistan, *Paleogeogr. Paleoclimatol., Paleoecol.*, *37*, 43–61, 1982.
Tichelaar, B. W. and L. J. Ruff, How good are our best models?, *EOS Trans. AGU*, *68*, 593–606, 1989.
Van Alstine, D. R., Analysis of the modes of directional data with particular reference to paleomagnetisim, *Geophys. J. R. Astron. Soc.*, *61*, 101–113, 1980.
Watson, G. S., More significance tests on the sphere, *Biometrika*, *47*, 87–91, 1960.
Watson, G. S., *Statistics on a Sphere*, John Wiley, New York, 1983.
Watson, G. S., Some estimation theory on the sphere, *Ann. Inst. Statist. Math.*, *38*, 263–275, 1986.
Watson, G. S., A test for randomness, *Mon. Not. Royal Astr. Soc., Geophys. Supp.*, *7*, 160-161, 1956.
Watson, G. S. and E. Irving, Statistical methods in rock magnetism , *Mon. Not. R. Astron. Soc.*, *7*, 289–300, 1957.
Wu, C. F. J., Jackknife, bootstrap and other resampling methods in regression analysis, *Ann. Stat.*, *14*, 1261–1295, 1986.

C. Constable, Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, La Jolla, CA 92093-0225,
N. Kylstra and L. Tauxe, Geological Research Division, Scripps Institution of Oceanography, La Jolla, CA 92093-0220.

(Received May 31, 1990;
revised February 4, 1991;
accepted February 26, 1991.)