The fold test in palaeomagnetism

P. L. McFadden[*] and D. L. Jones *Department of Physics, University of Zimbabwe, PO Box MP 167, Mount Pleasant, Salisbury, Zimbabwe*

Summary. The fold test suggested by McElhinny is shown to be inappropriate. More correctly one should determine whether the mean direction of a group of sites from one limb of a fold may be distinguished statistically from the mean direction of a group of sites from another limb. Details of the application of this test are given with examples. It is shown that the test indicates significance under far less severe folding than does the test suggested by McElhinny.

1 Introduction

Graham (1949) recognized that if the magnetic remanence directions in beds of different attitudes are brought into agreement after correction for the relative rotations of the beds then the magnetization had been acquired prior to the bedding deformation. Subsequently this was recognized as a powerful test for stability of the remanence.

It is possible of course that the 'unfolding' of small deformations could produce an overall improvement in the grouping of site mean directions (since they constitute a random sample) even though the magnetization was acquired *subsequent* to the bedding deformation. Thus it is not enough to note an improvement in the overall grouping since this 'improvement' may only be a consequence of the particular random sample chosen. In order to determine whether the improvement is 'real' a statistical test is required and to this end McElhinny (1964) suggested a test using a comparison of estimates for the population precision parameter before and after unfolding the deformation. This test has been fully accepted and used wherever possible; unfortunately the test is invalid.

It is shown why the test suggested by McElhinny (1964) is invalid. The correct test is developed and it is shown that the fold test could be significant in instances which would be judged as not significant using the earlier test.

---

[*] Present address: Research School of Earth Sciences, Australian National University, PO Box 4, Canberra, ACT 2600, Australia.

2 Test suggested by McElhinny (1964)

The test is based on the assumption that the directions are Fisher distributed (Fisher 1953). Watson (1956) has shown that if the directions are Fisher distributed then

$$2\kappa (N - R) \sim \chi^2_{2(N-1)} \tag{1}$$

where κ is the precision parameter of the population, N is the number of observations and R is the vector resultant of the individual observations considered as unit vectors. Here the symbol '$\sim$' is to be read as 'is distributed as' and χ^2_m is the chi-square distribution with m degrees of freedom. Defining the statistic k as

$$k = \frac{N - 1}{N - R} \tag{2}$$

the distribution of k is given by

$$2(N-1) \frac{\kappa}{k} \sim \chi^2_{2(N-1)}. \tag{3}$$

If two independent samples, of size N_1 and N_2, are drawn from populations with precision parameters κ_1 and κ_2, giving the statistics k_1 and k_2 then from (3)

$$\frac{k_2}{k_1} \frac{\kappa_1}{\kappa_2} \sim F[2(N_1-1), 2(N_2-1)] \tag{4}$$

where $F[a, b]$ is the F distribution with a and b degrees of freedom. Further, if $\kappa_1 = \kappa_2$ then

$$\frac{k_2}{k_1} \sim F[2(N_1-1), 2(N_2-1)]. \tag{5}$$

McElhinny (1964) assumed that equation (5) is valid (with $N_1 = N_2$) as a comparison for the grouping of site mean directions before and after unfolding a bedding deformation. This assumption will now be examined in the context of the derivation of (5).

For purposes of notational simplicity let the subscript 2 refer to the population and observed statistics with the bedding in the attitude at the time of acquisition of the magnetization. Thus this subscript might refer to either the *in situ* or unfolded attitudes. Subscript 1 then refers to the population and observed statistics with the beds in the 'incorrect' attitudes.

First, because of the distortion, with the beds in their 'incorrect' attitude, the overall population *cannot* be Fisher distributed and so the distribution of equation (1) is incorrect. However, even if the overall population were Fisher distributed with the beds in their 'incorrect' attitudes there are two further problems which would also invalidate the test.

First, if R_a is the resultant length of the site mean unit vectors from one limb of the fold and R_b that from the other limb then

$$R_1^2 = R_2^2 + 2R_aR_b(\cos\theta_1 - \cos\theta_2) \tag{6}$$

where θ is the angle between the overall means from the two limbs. Thus R_1 is dependent on R_2 and so k_1 is not independent of k_2. Consequently the step (3) to (4) is invalid.

Secondly, k_1 will typically be less than k_2. However, it must be recognized that this is so essentially because κ_1 is less than κ_2. Since it is known that under these conditions $\kappa_1 \neq \kappa_2$ the step from (4) to (5) is invalid. Furthermore, this shows why the test suggested by McElhinny

(1964) requires a greater degree of distortion than that which should be required for statistical significance.

4 Correct formulation of the test

The situation produced by deformation of beds is effectively that a single population has been split into two populations (or more if several folds are being considered), each having the same precision parameter as the original population but with different true mean directions. Hence the test must be formed in terms of directional comparison, i.e. the test must be one to determine whether the two true mean directions may be distinguished statistically on the basis of the available sample.

For the comparison of directions from one limb of the fold with another limb to be valid it is necessary that all the directions originally belonged to a single population. Thus, given that the samples from each limb were drawn from beds which acquired their magnetization contemporaneously, there is causal reason to believe that the precision parameter is the same for the populations on each limb (note that these precisions should be freed from within-site dispersion (McFadden 1981)). This hypothesis may be tested using equation (5) (Watson 1956) since the relation is valid for comparing the precisions of the populations on each limb. If this hypothesis may be rejected statistically then the validity of the fold test (in that particular instance) must be questioned because of the following possibilities. First, the precisions of the populations from the different limbs may be different because the original populations were actually different. In such a case there would be no causal reason for expecting the original true mean directions to be the same and so the test would be invalid. Secondly, the precisions may appear to be different owing to the bedding distortion being more complex than recognized by the observer. This may arise as a consequence of sites having been placed too close to a fold axis, the resultant rotation of such sites being different from that of other sites assigned to the same limb. Alternatively some of the sites assigned to a particular limb may have been derived from beds with a different, but unrecognized rotation. In either of these cases the data are incorrect and this would naturally invalidate any test.

Having confirmed that the precisions from the different limbs are the same the fold test may be performed simply by testing whether the overall mean directions from the different limbs may be distinguished statistically (McFadden & Lowes 1981). For the simple case of two limbs the hypothesis of a common true mean direction may be rejected if

$$\frac{[R_a + R_b - R^2/(R_a + R_b)]}{2(N - R_a - R_b)} > \left(\frac{1}{p}\right)^{1/(N-2)} - 1 \tag{7}$$

where R is the length of the resultant vector of all the site mean directions (i.e. of the vectors $\vec{R}_a$ and $\vec{R}_b$), p is the level of significance and N is the number of sites.

If m limbs ($m > 2$) are being considered simultaneously then the ks may be tested as shown by McFadden & Lowes (1981) and the relevant distribution for the fold test is

$$\left(\frac{N - m}{m - 1}\right) \frac{\Sigma R_i - R^2/\Sigma R_i}{2(N - \Sigma R_i)} = f \sim F[2(m - 1), 2(N - m)] \tag{8}$$

where the summations are for $i = 1$ to m and the R_i are the lengths of the vector resultants (from each limb) of the site mean unit vectors. If the observed value of f exceeds the critical value of the F distribution at the required level of significance then the hypothesis of a common true mean direction may be rejected. It should be noted that both of the tests

Table 1. Values of $(1/p)^{1/(N-2)} - 1$ for the application of the fold test with $p = 0.05$.

N	$(1/p)^{1/(N-2)} - 1$	N	$(1/p)^{1/(N-2)} - 1$
3	19.00	13	0.3130
4	3.472	14	0.2836
5	1.714	15	0.2592
6	1.115	16	0.2386
7	0.8206	17	0.2211
8	0.6475	18	0.2059
9	0.5341	19	0.1927
10	0.4542	20	0.1811
11	0.3950	25	0.1391
12	0.3493	30	0.1129

(equations 7 and 8) are conditional on a common precision but independent of the actual value of the precision parameter. Also, the tests are conditional upon the observed values of the R_i.

If the hypothesis of a common true mean direction may be rejected using the *in situ* observations but not when using the unfolded observations then it may be accepted that the observed magnetization was acquired before distortion of the beds. If the situation is reversed then it may be be accepted that the magnetization was acquired after the bedding distortion. If the hypothesis of a common true mean direction cannot be rejected either before or after unfolding then the distortion of the beds was too small to show a significance with the available sample.

For convenience numerical values of the right side of equation (7) are for $N = 3$ to 30 and $p = 0.05$ are given in Table 1.

4 Examples

Consider the following data, which were drawn at random from a Fisher population with $\kappa = 10$ and arbitrarily split into two groups as if they were drawn from different limbs of a fold.

$$N_a = 4, \qquad R_a = 3.770, \qquad k_a = 13.04;$$

$$N_b = 3, \qquad R_b = 2.823, \qquad k_b = 11.30.$$

The angle between the means of the two groups was $18.1°$ giving $R = 6.513$. If this sample had actually been obtained from field sampling the value of κ would not be known and equation (7) would therefore have to be used. From this equation the hypothesis of a common true mean direction would be rejected at the 95 per cent level of confidence if

$$\frac{6.593 - R^2/6.593}{0.814} > 0.8026. \tag{9}$$

Using $R = R_2 = 6.513$, the left side equals 0.1954 and thus the hypothesis of a common true mean direction would not be rejected (as expected in this instance). Using (9) it may be seen that for rejection of the hypothesis R must be less than 6.250. This means that the angle between the means of the two groups must exceed $37.5°$ for rejection of the hypothesis with the observed values of R_a and R_b. Hence, in this instance, for the fold test to be significant the distortion of the beds would have to produce a relative rotation of at least $19.4°$.

If the test suggested by McElhinny (1964) were used the minimum angle of rotation required for significance may be determined as follows. The critical value of the F distribution with 12 and 12 degrees of freedom is 2.686 for testing at the 95 per cent level of confidence. Thus for significance it would be necessary that

$$\frac{k_2}{k_1} = \frac{12.32}{k_1} > 2.686.$$ (10)

Hence R_1 would have to be less than 5.692 and the angle between the means of the two groups would have to exceed $61.3°$. Thus for significance the distortion of the beds would have to produce a relative rotation of at least $43.2°$, more than twice that required by the correct test. If the fold axis is not perpendicular to the true mean direction of the original magnetization then the bed distortion required for significance is even larger.

The second example is drawn from the Lupata Alkaline Volcanics (Gough & Opdyke 1963). This example was used by McElhinny (1964) to show that the fold test was not significant in that instance. The relevant data are given in Table 2. Now $k_2/k_1 = 2.10$ and the critical value of the F distribution at the 95 per cent level of confidence is 2.686, as in the above example. On this basis McElhinny (1964) concluded that the test was not significant. However, using the correct test, the hypothesis of a common true mean direction may be rejected at the 95 per cent level of confidence if

$$\frac{6.9837 - R^2/6.9837}{0.0326} > 0.821,$$ (11)

assuming that the precisions on the two limbs are the same. In this example $k_b/k_a = 2.18$ and the critical value of the F distribution (with 8 and 2 degrees of freedom) at the 95 per cent level of confidence is 19.4. Consequently there is no reason to suppose that the populations

Table 2. Relevant data for the Lupata Alkaline Volcanics.

Group 1

Site	*In situ* D	I	After unfolding D	I
1	346.7	− 60.9	337.4	− 52.9
2	346.2	− 60.3	337.2	− 52.3
3	338.2	− 64.6	329.7	− 55.8
4	332.2	− 59.5	326.3	− 50.4
5	356.3	− 64.7	343.4	− 57.6
Mean	343.7	− 62.2	334.6	− 54.0
	$R_a = 4.9853$		$k_a = 272.2$	

Group 2

Site	D	I	D	I
6	344.8	− 54.9	337.9	− 56.7
7	344.8	− 50.2	339.0	− 52.1
Mean	344.8	− 52.6	338.5	− 54.4
	$R_b = 1.9984$		$k_b = 594.5$	

Overall statistics
In situ: $R_1 = 6.9631$; $k_1 = 163$; $N = 7$
After unfolding: $R_2 = 6.9825$; $k_2 = 343$; $N = 7$
$R_a + R_b = 6.9837$
$2(N - R_a - R_b) = 0.0326$

on each limb do not share a common precision parameter. With $R = R_2 = 6.9825$ the left side of equation (11) is 0.074 and the hypothesis of a common true mean direction after unfolding cannot be rejected. Using $R = R_1 = 6.9631$ the left side of equation (11) is 1.26 and the hypothesis of a common true mean direction *in situ* may be rejected. Hence the fold test is in fact significant and it may be concluded that the Lupata Alkaline Volcanics acquired their observed magnetization before the folding occurred.

In this example the precisions have not been freed from within-site dispersion. The precision freed from within-site dispersion is higher than the precision including within-site dispersion and so the distortion required for significance of the fold test is reduced even further. Consequently in this example it was not necessary to perform a two-tier analysis.

5 Conclusion

It has been shown that the statistical test suggested by McElhinny (1964) for determining significance of the fold test is invalid and that to obtain (apparent) significance using this test a far greater distortion would be required than is actually necessary for significance using the correct test. The correct statistical procedure for applying the fold test is to determine whether the mean direction of a group of sites from one limb may be distinguished from the mean direction of a group of sites from another limb of the fold.

Since the test suggested by McElhinny (1964) is too stringent it is almost certain that any workers who have claimed the presence of a significant fold test in an investigation will be correct. The only possibility that these conclusions were incorrect is that the precision of the magnetization from one limb may have differed from the precision from another limb. The important consequence is that there may have been many instances in which a fold test was significant but such significance was rejected on the basis of the incorrect test.

References

Fisher, R. A., 1953. Dispersion on a sphere, *Proc. R. Soc. A*, 217, 295–305.
Gough, D. I. & Opdyke, N. D., 1963. The palaeomagnetism of the Lupata Alkaline Volcanics, *Geophys. J. R. astr. Soc.*, 7, 457–468.
Graham, J. W., 1949. The stability and significance of magnetism in sedimentary rocks, *J. geophys. Res.*, 54, 131–167.
McElhinny, M. W., 1964. Statistical significance of the fold test in palaeomagnetism, *Geophys. J. R. astr. Soc.*, 8, 338–340.
McFadden, P. L., 1981. Two-tier analysis in palaeomagnetism, *Geophys. J. R. astr. Soc.*, submitted.
McFadden, P. L. & Lowes, F. J., 1981. The discrimination of mean directions drawn from Fisher distributions, *Geophys. J. R. astr. Soc.*, 67, 19–33.
Watson, G. S., 1956. Analysis of dispersion on a sphere, *Mon. Not. R. astr. Soc. geophys. Suppl.*, 7, 153–159.