Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
EPSL
Earth and Planetary Science Letters 212 (2003) 151–166
www.elsevier.com/locate/epsl

The direction–correction tilt test: an all-purpose tilt/fold test for paleomagnetic studies

Randolph J. Enkin *

Geological Survey of Canada – Pacific, P.O. Box 6000, Sidney, BC, Canada V8L 4B2

Received 6 January 2003; received in revised form 27 March 2003; accepted 31 March 2003

Abstract

The tilt or fold test in paleomagnetism is used to infer whether paleomagnetic remanence was acquired before or after tectonic tilting. While several tilt test formulations have been proposed, none fully satisfies the requirements of statistical validity, applicability to all bedding geometries, and ease of use. This paper introduces the direction–correction (DC) test, which examines the relationship between the paleomagnetic site mean directions and their corresponding bedding tilt corrections. The DC test is similar to McFadden's correlation test because both test whether or not the site directions contain information about the bedding tilts, however the DC test gives greater weight to sites with greater counter-bedding. The DC test is similar to Watson's numerical tilt tests because they determine the degree of untilting which gives optimal concentration of site directions, however the DC test uses analytical rather than numerical methods. Graphical output aids the researcher in recognizing problem sites. Using both real and simulated data, the DC test is demonstrated to be more discriminating than other tilt test formulations for all bedding geometries. The simulations show that the power of the tilt test is inversely proportional to the 95% confidence interval (α_{95}) of the overall mean. As a rule of thumb, paleomagnetists should attempt to sample sufficient sites to obtain an α_{95} less than 1/6 of the bedding attitude difference.

Keywords: paleomagnetism; spherical statistics; tilt test; fold test

1. Introduction

Paleomagnetists often collect their samples from tilted strata and apply the 'fold test' (or more generally the 'tilt test') to determine the age of magnetic remanence acquisition relative to tectonic deformation. The idea was developed by Graham [1], who recognized that if rocks were magnetized before tilting, then 'restoring the beds to the horizontal causes the magnetizations to move towards parallelism with one another'.

A number of statistical formulations have been proposed to analyze the tilt test, however none is entirely satisfactory. A good formulation is (1) statistically valid, (2) applicable to the range of geological situations paleomagnetists encounter, and (3) easy to use. This paper introduces the direction–correction (DC) tilt test, which is a

* Tel.: +1-250-363-6431; Fax: +1-250-363-6565.
E-mail address: renkin@nrcan.gc.ca (R.J. Enkin).

0012-821X/03/$ – see front matter Crown Copyright © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0012-821X(03)00238-3

powerful yet simple method which should meet the requirements of paleomagnetists.

1.1. Nomenclature

Paleomagnetic directions determined in 'geographic coordinates' are calculated relative to geographic north and the present horizontal plane. These directions are transformed to 'stratigraphic coordinates' (relative to bedding planes of strata or paleohorizontal) by rotation about the horizontal strike axis of the strata. A 'positive tilt test' concludes that paleomagnetic remanence was acquired before tilting, while a 'negative tilt test' concludes acquisition after tilting. If the tilt test renders an intermediate result between positive and negative conclusions, the remanence is often concluded to have been acquired 'syn-tilting' (noting however that other mechanisms can lead to this conclusion).

2. The DC test

The DC test involves a graph showing whether or not the geographic site mean directions are correlated to their corresponding bedding tilt corrections. It is based on the reasoning proposed by McFadden [2] for his correlation test. If the paleomagnetic remanence is acquired before tilting, then the site directions in stratigraphic coordinates are independent of the bedding attitudes, while the site directions in geographic coordinates contain information about both the initial direction distribution and the tilt rotations which bring the strata to their present attitudes. In contrast, if remanence acquisition followed tilting, the site directions in geographic coordinates contain no information on the tilt rotations.

Table 1
Example data used to construct Fig. 1

Site	D_G (°)	I_G (°)	D_S (°)	I_S (°)	k	α_{95} (°)	N	Strike (°)	Dip (°)
1	332.0	24.0	249.8	46.1	48.4	13.3	4	102.0	87.0
2	310.0	25.0	241.2	46.9	48.4	13.3	4	84.0	77.0
3	291.0	65.0	240.9	37.7	48.4	13.3	4	119.0	44.0
4	277.0	7.0	249.2	39.7	48.4	13.3	4	56.0	64.0

KR: $k_S/k_G = 30.2 > [4.26]$, KR tilt test inferred positive; AC1: $\xi_G = 1.98 < [2.34]$, $\xi_S = 0.75 < [2.34]$, AC1 tilt test indeterminate; AC2: $\xi_G = 3.81 > [2.34]$, $\xi_S = |-1.36| < [2.34]$, AC2 tilt test inferred positive; OC: max k at $108.4 \pm 16.2\% \sim 100\%$, OC tilt test inferred positive; DC: max k at $108.0 \pm 19.9\% \sim 100\%$, DC tilt test inferred positive.

Note: D, declination; I, inclination; subscript G, geographic coordinates; subscript S, stratigraphic coordinates; k, Fisher concentration; α_{95}, 95% confidence interval; N, number of specimens in site mean; strike, dip, bedding attitude using right-hand rule. In the tilt test summaries, the critical values for 95% confidence are given in square brackets.

The DC test also follows Watson and Enkin's [3] method of determining the degree of untilting required to give optimal concentration of site directions. If the best clustering occurs on either 0 or 100% untilting, then the tilt test is negative or positive, respectively. Intermediate results lead to a syn-tilting conclusion. The DC test offers an analytical rather than a numerical method of determining the optimal degree of untilting, so it is easy and efficient to use, and it is simple enough to require only ~50 lines of computer code (e.g., Appendix 1).

The DC test is best introduced with an example data set. Consider a study (Table 1) in which the paleomagnetic remanences were acquired before tilting so that site directions in stratigraphic coordinates (s_i) are tightly clustered, but not in geographic coordinates (g_i) (Fig. 1a) (the subscript i refers to the ith site). A tectonic correction brings a site direction in geographic coordinates to stratigraphic coordinates (i.e., the small circle segment which brings each g_i to its corresponding s_i) (Fig. 1b). These same rotations can be applied to the mean in geographic coordinates (G) to produce 'forward corrections of the mean' (f_i) (Fig. 1c). Inversely, backward corrections of the stratigraphic mean (S) give 'back corrections of the mean' (b_i) (Fig. 1d).

Since the same rotation which transforms g_i to s_i is used to transform G to f_i and b_i to S (Fig. 1b), the arc separations s_iS are the same as g_ib_i, and g_iG are the same as s_if_i. In this example, the s_i are tightly clustered so the g_i and b_i are close to (i.e., correlated to) each other (Fig. 1d). Conversely, the s_i and f_i are not correlated since the g_i are dispersed (i.e., not close to their mean, G) (Fig. 1c).

2.1. Construction of the DC plot

This correlation can be quantified with a DC plot (Fig. 1f) using values illustrated in Fig. 1e. On the horizontal axis is plotted c_i, the angular distance between unit vectors G and b_i:

$$c_i = \cos^{-1}(G \cdot b_i) \tag{1}$$

The vertical axis represents the projection of arc Gg_i onto Gb_i (i.e., the angular distance, d_i, from G to the intersection of the Gb_i great circle with the perpendicular great circle which goes through g_i). If d'_i is the angular distance between G and g_i, and φ_i is the angle between Gg_i and Gb_i, then spherical trigonometry on the right triangle gives:

$$d_i = \tan^{-1}(\tan(d'_i)\cos(\varphi_i)) \tag{2}$$

which is very similar to the planar projection $d = d'\cos(\varphi)$ if φ or d' is small (which is usually the case). Noting that φ is the angle between the normals to the great circles Gg_i and Gb_i, one obtains:

$$\begin{aligned} d_i &= \tan^{-1}\left(\frac{\sin(d'_i)\cos(\varphi_i)\sin(c_i)}{\cos(d'_i)\sin(c_i)}\right) \\ &= \tan^{-1}\left(\frac{(G \times g_i)\cdot(G \times b_i)}{(G \cdot g_i)\,|G \times b_i|}\right) \end{aligned} \tag{3}$$

The value d_i is taken instead of the value d'_i because if φ_i is large, d'_i could be similar in size to c_i even when g_i and b_i are not at all correlated. Care must be taken in choosing the correct quadrant for d_i. Usually the range is from $-90°$ to $+90°$ except when the bed is overturned (dip $>90°$) in which case $G \cdot g_i$ may be less than 0 and then d_i is $>90°$.

2.2. Determining the DC slope

When the remanence acquisition predates tilting, the arc lengths c_i and d_i are similar for each site and the slope of the line passing through the origin is approximately 1 (i.e., c_i and d_i are correlated). However, if one does the equivalent calculation comparing arc lengths Ss_i and Sf_i (Fig. 1c) the c_i values are identical but the d_i values would all be close to zero leading to a slope of approximately 0 (i.e., no correlation).

The slope, s (with its associated standard error, σ), of the line passing through the origin can be determined by least squares (e.g., [4], ch. 12). Appendix 2 outlines the assumptions implicit in using least squares linear regression and their justification for the DC test.

$$s = \sum d_i c_i / \sum c_i^2, \tag{4}$$

$$\sigma^2 = \left(\frac{\sum d_i^2}{\sum c_i^2} - \frac{(\sum d_i c_i)^2}{(\sum c_i^2)^2}\right) / (N-2) \tag{5}$$

where the summations are from 1 to N (the number of sites). There are $N-2$ rather than the $N-1$ degrees of freedom usually used in linear regression since two degrees of freedom are taken by estimating the means in geographic and stratigraphic coordinates (see Appendix 2).

Using the approximation that the data are Gaussian-distributed about their true values (see Appendix 2), the estimate of the slope is distributed as Student's t with $(N-2)$ degrees of freedom. Thus the uncertainty of the slope is given by:

$$\Delta s = t_{(N-2)}^{(1-\alpha/2)} \sigma, \tag{6}$$

where the α is set to some acceptable small quantity, usually 5%. The value of $t_{(N-2)}^{0.975}$ (found in standard statistical tables) is usually between 2 and 3.

2.3. The DC slope is the optimal degree of untilting

The concentration or clustering of site directions can be calculated as a function of the degree of untilting, assuming the same proportion of tilt correction is applied to each site. This value is greatest at the 'optimal degree of untilting', γ_{max}. If the best clustering occurs in stratigraphic coordinates, the γ_{max} is 100%. Until now, γ_{max} could only be determined numerically by calculating concentration at a series of untilting proportions. However, the analytically determined DC slope is equivalent to γ_{max}. This equivalence can be shown to be exactly true for the analogous one-dimensional geometry (see Appendix 3). On a sphere, it is also exactly true for the special case of zero dispersion at a given degree of untilting. There are slight variations due to spherical distortion when there is some scatter to the directions at optimal untilting, but the differences are almost always below 1%. Larger differences are observed when dispersion is high (e.g., Fisher concentration $k<10$) or the structural differences small (e.g., bedding differences $<10°$).

It is useful to multiply the DC slope by 100% when reporting the results of the DC test. In the example, instead of reporting that the DC slope is 1.080 ± 0.199, it is more physically intuitive to state that the optimal clustering is at $108.0\pm19.9\%$ untilting. Numerical determination of $\gamma_{\max}$ gives 108.4%.

2.4. Hypothesis testing with the DC plot

The statistical significance tests for the DC test are similar to those used for other formulations of the tilt test. If s is significantly different from 0 but not significantly different from 1, that is, if we have at the same time:

$$|s|>\Delta s \text{ [i.e., } s=0 \text{ rejected]}, \tag{7}$$

and

$$|s-1|\leq\Delta s \text{ [i.e., } s=1 \text{ not rejected]}, \tag{8}$$

then the optimal degree of untilting is about 100% and the remanence is most likely pre-tilting. In Fig. 1f we see that the shaded 95% confidence region of the best fitting line brackets slope 1 but not slope 0, so the directions are correlated to the corrections and the tilt test is positive.

If the contrary unequalities hold:

$$|s-1|>\Delta s \text{ [i.e., } s=1 \text{ rejected]}, \tag{9}$$

and

$$|s|\leq\Delta s \text{ [i.e., } s=0 \text{ not rejected]}, \tag{10}$$

then the remanence directions in geographic coordinates are not correlated to the bedding corrections, optimal untilting is about 0%, and we can infer that the remanence was acquired after tilting.

If neither of these combinations is fully satisfied, then the tilt test renders an ambiguous result. If the optimal degree of untilting lies between 0 and 100% then the remanence is often inferred to have been acquired syn-tilting, but it can also mean that the remanence is an unresolved combi-

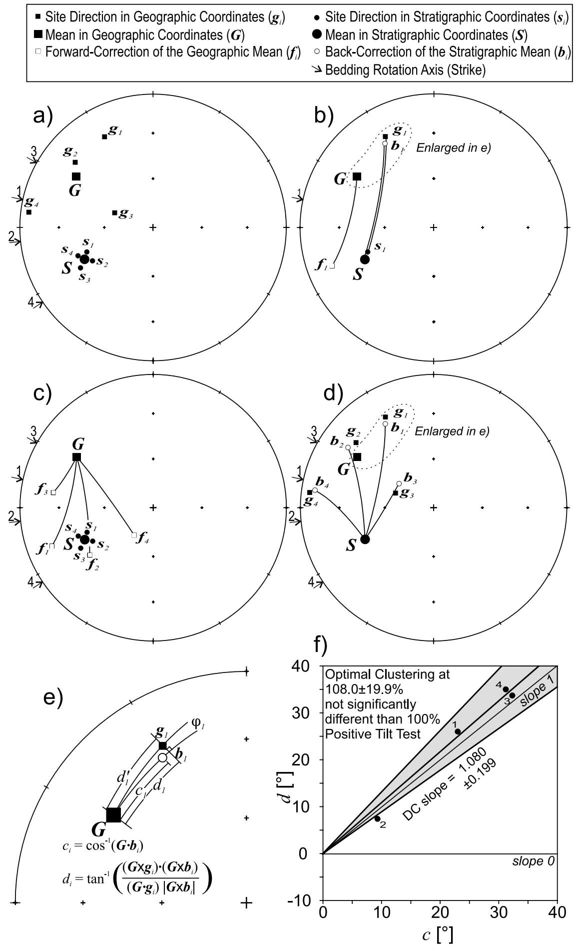


Fig. 1. Example of the DC tilt test using pre-tilting remanence (data from Table 1). See text for explanation. (a) The site directions and their means in geographic and stratigraphic coordinates. (b) The definition of the forward correction of the geographic mean and the back correction of the stratigraphic mean for site 1. (c) Forward corrections of the mean are separate from their corresponding site directions in stratigraphic coordinates. (d) Back corrections of the mean are close to their corresponding site directions in geographic coordinates. (e) Definitions of d and c for site 1. (f) The DC plot, showing how the DC slope is not significantly different from 1, but it is significantly different from 0, indicating a pre-tilting remanence.

nation of pre- and post-tilting components or that pre-tilting remanence has been strain-deformed. If s is not significantly distinguishable from both 0 and 1, then the bedding attitudes are not sufficiently different for a successful tilt test.

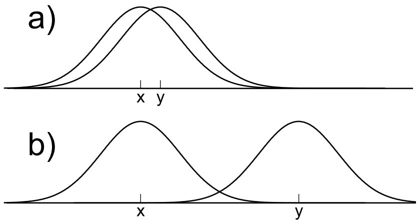
3. Comparison with other tilt test formulations

3.1. Hypothesis testing: Type I and Type II errors

Each tilt test is based on statistical inference using hypothesis testing. A null hypothesis is proposed for which the distribution of some test statistic derived from the data can be estimated either analytically or numerically. A range of values, or confidence interval, is established based on assuming that the null hypothesis is true. If the observed test statistic lies outside that confidence interval, the probability that the null hypothesis is true is small and thus the null hypothesis is inferred to be false at whatever arbitrary confidence level was chosen. The paleomagnetic community usually considers 95% as an appropriate level of confidence.

When testing a null hypothesis such as $x=y$, the probability of rejecting the hypothesis when it is true (Type I error, probability α) can be set to any value (given a model probability distribution), but the probability of accepting the hypothesis when it is in fact incorrect (Type II error, probability β) depends on the true values of x and y and is therefore unknown (Fig. 2). β decreases as the difference between the true x and y increases. The probability that the test infers that $x=y$ (i.e., x not significantly different from y) is low if their difference is large. For the tilt test this means that a significant result will be inferred only if the differences in bedding attitudes are large enough. The larger the bedding tilt differences, the lower the probability of Type II errors and thus the more predictive the tilt test will be.

While one cannot know β without knowing the true amount by which the null hypothesis is wrong, there is a tradeoff between α and β; β is large when a small value of α is chosen. Thus when one wants to test if two things are equal, the significance level should not be set too high.



c)

Truth \ Inference	x not significantly different than y	x significantly different than y
x = y	GOOD	Type I error. Probability α, arbitrarily set
x ≠ y	Type II error. Probability β, dependent on α and separation	GOOD

Fig. 2. Simple case of testing the hypothesis $x=y$ and the probability of Type II errors. (a) When the true separation between x and y is small, there is a high probability of the test failing to discriminate the difference. (b) The probability of failing to discriminate this difference decreases as the difference increases. (c) The four possibilities for a hypothesis test.

One might recommend, for example, using a 99% significance level ($\alpha=0.01$) when testing whether the paleomagnetic directions are distinct in geographic coordinates, but only a 90% significance level ($\alpha=0.1$) when testing whether the directions are the same in stratigraphic coordinates. There are several instances in the paleomagnetic literature of tilt tests passing at the 99% level, which is actually a weaker test than passing at the 95% confidence level because the directions are more likely to be within 3σ (99% significance) of each other than within 2σ (95% significance) of each other. In practice, a compromise of 95% significance for both tests is usually reasonable, but one must be aware of the possibility and implications of a Type II error.

3.2. The kappa ratio (KR) test

McElhinny [5] introduced a formulation (designated the KR test) which checks whether the concentration of site directions is significantly better in stratigraphic than in geographic coordinates. Specifically, it tests the null hypothesis that the dispersion of directions as measured by the Fisher concentration parameter, k, is the same in both coordinate systems. If the ratio of concentrations in stratigraphic and geographic coordinates, k_S/k_G, is significantly greater than 1, then the null hypothesis is rejected and one interprets that the magnetization is pre-tilting.

The KR test is easy to implement, so the test has been widely used in paleomagnetism. Nevertheless, McFadden and Jones [6] showed that the statistical inference is meaningless. While the ratio k_S/k_G is useful for qualitative analysis, it is incorrect to claim that any value is statistically significantly greater than 1, because k_S and k_G are not independent statistical quantities. Rather they are geometrically related through the measured bedding attitudes. It is incorrect to claim 95% confidence when a pre-tilting remanence passes the test 100% of the time, given sufficiently large differences in bedding dips. When bedding variations are small relative to the paleomagnetic dispersion, the KR test always fails. In the language of hypothesis testing, the KR test statistic has the characteristics of Type II errors rather than Type I errors, proving that it is statistically ill-posed. Furthermore, the KR test often infers a significantly positive or negative conclusion when a syn-tilting conclusion is more appropriate [3].

3.3. The multiple sites per limb (MSL) test

McFadden and Jones [6] proposed a tilt test for the special case of multiple sites available from each limb of a fold. Refinements of this formulation were presented by McFadden [2,7] and Fisher and Hall [8,9]. First one tests whether the Fisher distributions of site directions from each limb are compatible. If so, one then tests whether the average directions from each limb are different in both geographic and stratigraphic coordinates. If the directions are significantly different in geographic coordinates but not in stratigraphic coordinates, then the remanence is interpreted to predate tilting. The MSL formulation is statistically valid, however it requires several sites from each homoclinal limb, which is rare in paleomagnetic studies.

3.4. The isolated observation (IO) test

The IO test [2] is applicable when all sites have the same bedding attitude except one, the isolated observation. If the magnetization is pre-tilting, then the direction of the isolated observation in geographic coordinates should be further away from the other remanence directions than can be reasonably expected if they follow a Fisher distribution. However, the isolated observation should be compatible with the other sites in stratigraphic coordinates.

3.5. The azimuth correlation (AC1 and AC2) tests

A substantially different way of regarding the tilt test problem is considered in McFadden's [2] two correlation tests (AC1 and AC2, abbreviating azimuth correlation). The AC tests were the first to consider the orientations of the bedding planes as well as the remanence directions. If the remanence is pre-tilting, the site directions in stratigraphic coordinates contain no information concerning the subsequent tectonic deformation. Thus there should be no correlation between the bedding dip directions and the site directions in stratigraphic coordinates. On the other hand, in geographic coordinates one expects to see the site directions correlated to the bedding directions since both the magnetic vectors and the bedding normals rotated the same way during tilting.

In these tests, the azimuthal distribution of the remanence and dip directions around the mean direction are considered with no reference to the magnitude of the angular differences from the mean. All sites, regardless of their bedding dip, are given equal weight.

For both AC tests, the structural correction is applied to the mean direction to produce what McFadden calls a 'shifted-mean' (in this paper called a 'correction of the mean') corresponding to each site. In the first definition (AC1), the correlation of the azimuths of the site directions and the corrections of the mean around the mean direction is tested for each coordinate system separately. For AC2, the same correlation is tested but the corrected means come from the alternate coordinate system. The two definitions are almost identical if the mean directions are not very different in the two coordinate systems, but only the second is appropriate when tectonic corrections pull the mean significantly in one direction.

3.6. Relation of the DC test to the AC test

The test statistic for the AC tests is the sum of $\cos(\varphi_i)$ while the DC test statistic is related to the sum of $d_i \approx d'_i \cos(\varphi_i)$. This means the AC tests give equal weight to all sites when testing the correlation between the site mean directions and the bedding attitudes, while the DC test gives higher weight to the sites which have more divergent bedding attitudes and thus higher d'. When the tectonic structure is symmetrical (for example, a two-limb geometry where each limb has the same number of sites), the c_i and d'_i values are nearly identical, in which case the DC slope (Eq. 4) is proportional to the sum of $\cos(\varphi_i)$, meaning that there is no practical difference between the AC and DC tests. However when the number of sites with divergent beddings are unbalanced, the AC tests lose their ability to identify the correlation. It was this observation which originally led to the development of the DC test.

The IO test was introduced in [2] to avoid this weakness in the AC tests. A single site with counter-bedding (the isolated observation) is swamped in the noise of the other sites when only azimuths are considered, but that site will have a large d giving it great weight in the DC test. To avoid this problem McFadden [2] recommended averaging the site directions of a limb together before applying the AC tilt test when one limb has many more sites than the others, otherwise the 'site means from the one limb will completely swamp the others, and it is unlikely that any correlation will be discernible'.

3.7. The optimal concentration (OC) test

Watson [3,10] reasoned that the tilt test in paleomagnetism should be implemented as a parameter estimation problem. The important aspect of the tilt test is not the degree of parallelism of the site directions, but rather how the rocks were oriented at the time of magnetization. Site directions are most concentrated when the beds are in (or not significantly different from) the attitudes they had at the time of remanence acquisition. Therefore, paleomagnetists should determine the degree of untilting, γ_{max}, which leads to optimal concentration. For example, if γ_{max} is not significantly different than 100%, then one can conclude that remanence was acquired while the beds were flat-lying. Since analytical estimates of γ_{max} and its confidence interval were not available, a parametric resampling strategy was proposed.

McFadden [7] questioned 'the validity of the basic assumption that the magnetization was acquired where the magnetic directions cluster most tightly'. He was correct in that maximum concentration of a primary remanence will not be found on exactly 100% untilting because of ordinary statistical variations. However, when bedding corrections are larger than the site direction dispersion, the bedding corrections always disrupt the distribution leading to lower concentration. (If the difference in bedding attitudes is smaller than the dispersion of magnetization directions, then no tilt test formulation can be effective.)

3.8. Relation of the DC test to the OC test

Since the DC slope is an analytical calculation of γ_{max}, the interpretation of the DC and OC tests is identical. Note that the DC method treats poorly defined sites the same as well-defined sites, whereas the OC method recognizes that sites have different confidence intervals. It is possible to include site weights in the least squares estimates, however that goes counter to general paleomagnetism practice of using all sites which pass certain minimum acceptability criteria, such as site $\alpha_{95} < 15°$.

It is quite common for the OC confidence interval for γ_{max} to be smaller than the DC confidence interval. Implicit in the OC tests is the questionable assumption that all sites share a common true direction and that all sources of dispersion are included in the within-site scatter. If, for example, each site is a lava flow, then the within-site scatter is typically much smaller than the between-site scatter because of secular variation. The OC numerical simulations estimate a too-small confidence interval for the optimal degree of untilting, and the tilt test often concludes that optimal untilting is 'significantly different' than 100% even though the remanence is primary. In contrast, the DC test uses the between-site dispersion to estimate the confidence interval.

Numerical simulations show that the OC and DC tests render the same confidence interval for γ_{max} when the individual sites have $k_w N_w = k_B$ (where k_B is the between-site precision, k_w is the within-site precision, and N_w is the number of specimens within a site mean). When $k_w N_w > k_B$, there must be a between-site contribution to dispersion which is not sampled within-site, and the OC method can not be expected to provide correct inference.

4. Comparison of tilt test formulations with real and simulated data

To compare the DC tilt test formulation against the others, real data and simulations of random data were subjected to all applicable tests. The advantage of simulated data is that we are privy to the 'truth'. The remanence is known to predate tilting, since it is constructed to be so. By choosing thousands of random samples, one can determine the validity of the hypothesis tests and the critical values they use. The frequency of Type I errors should equal the α probability. The frequency of Type II errors should decrease as the differences in bedding dips increase, and their measured probability β gives an indication of the power of the tilt test formulation.

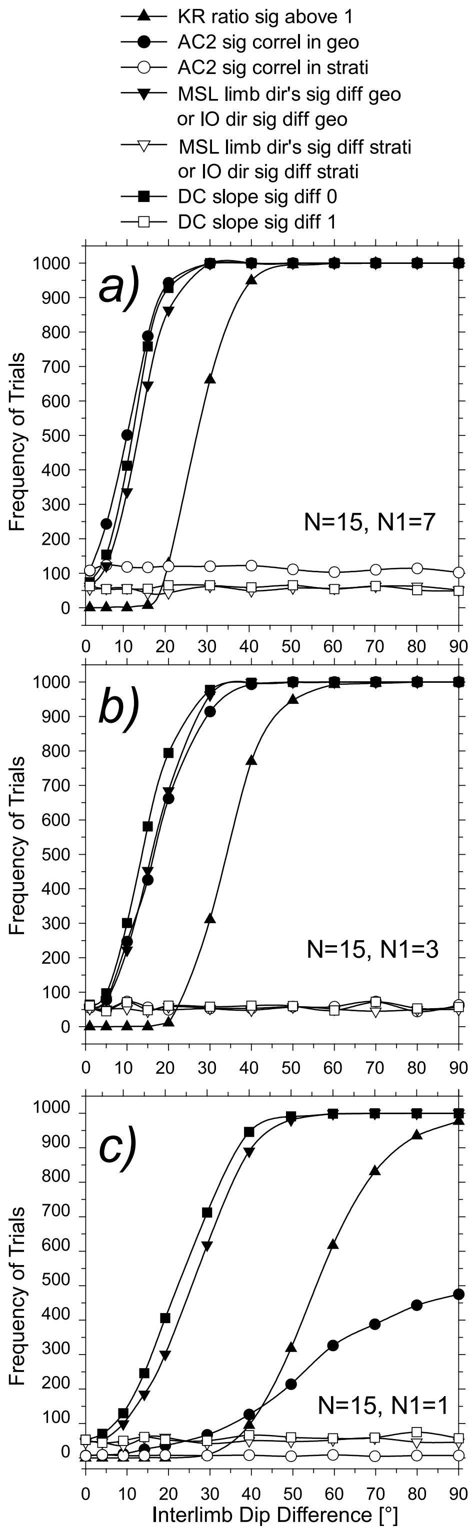
The simplest situation is the two-limb geometry, and most of the characteristics and behaviors of the tilt test can be explored with it. For each trial, a sample of N site directions is chosen from a Fisher distribution with vertical mean and precision k. (With the mean 90° from the horizontal strike axis, each degree of dip correction leads to a degree of remanence direction rotation. A different mean would lead to a smaller rotation but the characteristics of the tilt test would be the same.) One limb with $N1$ sites is tilted, arbitrarily to the west, by a chosen dip. For each set of these four variables N, $N1$, k, and dip, 1000 trials were performed. The random number generator seed was also noted so that the data could be reconstructed. Hypothesis tests were applied using a 95% confidence interval.

The KR, AC1, AC2, and DC tests can be applied to all situations. If $N1 = 1$, then the IO test can be applied, otherwise the MSL test was used. The OC test, being equivalent to but slower than the DC test, was not included.

Typical results are illustrated in Fig. 3, for $N = 15$ sites, $k = 30$, where the dip was varied from 1 to 90°. Open symbols indicate the number of times per 1000 trials that Type I errors occurred (i.e., MSL: a significant difference in the two limb directions was incorrectly inferred in stratigraphic coordinates; AC2: a significant correlation between beddings and directions was incorrectly inferred in stratigraphic coordinates; DC: γ_{max} was incorrectly found to be significantly different from 100%). Type I errors for the KR test, where k_G was incorrectly found to be significantly greater than k_S, never occurred. Closed symbols indicate the frequency of correct inferences, or equivalently the lack of Type II errors (i.e., KR: k_S/k_G was significantly greater than 1; MSL: the two limb directions were significantly different in geographic coordinates; AC2: the directions and beddings were significant correlated in geographic coordinates; DC: γ_{max} was significantly different than 0%).

After 1000 trials, a well-behaved test will have $\sim$50 Type I errors, regardless of the simulation parameters. When the dip is only 1°, Type II errors occur almost every trial. With a 90° dip, most tests have no Type II errors. A measure of the power of a test is the dip beyond which Type II errors are rare. The more powerful the test, the lower this dip will be.

Results for the AC1 test are not illustrated in Fig. 3 because it requires the mean direction to be the same before and after tectonic correction, which is not generally the case. The simulations show that when the two limbs have the same number of sites and are given equal and opposite dips, the AC1 and AC2 tests give identical results. But with unbalanced structures, the mean in geographic coordinates is different from that in stratigraphic coordinates, which makes a proportion



(about 3%) of trials fail the AC1 test regardless of the dip (Type II errors, which should go to zero frequency).

4.1. Two limbs with equal numbers of sites

When $N = 15$, $N1 = 7$, the two limbs have about equal weight (Fig. 3a). The AC2, DC and MSL tests all perform well when the dip is greater than $\sim 20°$, in that there are few Type II errors. The KR test requires steeper dip of about 40°.

Note how all the tilt tests except AC2 have Type I errors about 50 times or 5% as expected. The AC2 test fails about 10% of the time, because the tabulated critical values [2] mistakenly correspond to 90% rather than 95% confidence. The AC test statistic, $\xi = \Sigma\cos(\varphi_i)$, is a measure of the correlation between bedding corrections and site directions. Recognizing that anticorrelation (i.e., ξ significantly less than 0) is as important as correlation (i.e., ξ significantly greater than 0), $|\xi|$ rather than ξ is compared to the critical value, thus doubling the null hypothesis rejection region. Numerical simulations confirm the anomalous rejection rate. If a true 95% confidence limit was used, the number of Type II errors would go up, and the AC2 test would be seen to be less powerful than the DC test.

4.2. Two limbs with unbalanced sites; isolated observation

Simulations were also done for unbalanced two-limb geometries, specifically $N1 = 0.2N$ (e.g., Fig. 3b) and $N1 = 1$ (e.g., Fig. 3c). In the first case, the MSL test can still be applied, however for the latter case the IO test is substituted. The substantial difference in comparison to the balanced limbs trials is that a greater dip is necessary to

←

Fig. 3. Comparison of tilt tests using simulations of two-limb geometry ($N = 15$, $k = 30$) as described in the text. For each dip, 1000 trials were performed. Open symbols shows the frequency of Type I errors (expected to occur 50/1000 times), and closed symbols shows 1000 minus the number of Type II errors. (a) Balanced limbs geometry ($N1 = 7$). (b) Unbalanced limbs geometry ($N1 = 3$). (c) Isolated observation geometry ($N1 = 1$).

avoid Type II errors. This effect is most pronounced with the AC2 test.

Note that the frequency of Type I errors with the AC2 test goes down the more unbalanced the structure becomes. In principal, Type I errors should always occur with the arbitrarily chosen probability, and should not be dependent on the bedding geometry. With unbalanced bedding geometries, the mean direction is dominated by the site directions from the main limb, and thus most of the φ_i azimuths remain similar in geographic and stratigraphic coordinates. By not giving added weight to the sites with counter-bedding, the AC tests obtain too few Type I errors and too many Type II errors.

4.3. Relationship between bedding dip difference and tilt test power

Many simulations with different combinations of N, $N1$, k, and dip were done. To first order, the results always look similar to those illustrated in Fig. 3. When the bedding dip difference, dip, is small, tilt tests almost always fail because Type II errors are very common. With large dip, there are no Type II errors, so the tilt test only fails because of Type I errors. One measure of the power of a tilt test is the dip_0 at which the test usually renders the correct result, say 900 correct inferences from 1000 trials (i.e., $\beta = 10\%$); the smaller the dip_0, the more powerful the test.

The DC test always performs well, followed closely by the MSL or IO test, with the KR test worst of all. The AC2 test apparently is the most powerful when the limbs are balanced, however the confidence level is mistakenly set at 90% instead of the stated 95%. Larger dip_0 is needed when the structure is less balanced, with the AC2 test becoming the worst with decreasing $N1/N$.

For a given tilt test and bedding structure (i.e., the proportion $N1/N$), dip_0 is found to be proportional to $1/\surd(kN)$. Since the confidence interval of the mean of site directions, α_{95}, is approximately $140°/\surd(kN)$, this relationship implies that there is a minimum bedding dip difference which should be sampled proportional to the anticipated α_{95}. For a balanced structure, the DC test is useful ($\beta < 10\%$) with dip differences $3.4\alpha_{95}$; for unbalanced limbs ($N1/N = 0.2$), $\mathrm{dip}_0 = 3.7\alpha_{95}$; and for isolated observation geometry, $\mathrm{dip}_0 = 5.8\alpha_{95}$.

This analysis leads to a simple rule for paleomagnetists in the field. Remembering that the simulations were done with the mean direction 90° from the bedding strike, a typical real study will require even greater dip difference to render a positive tilt test. I recommend looking for beds with differences of greater than six times the anticipated α_{95} value. If the rock types being studied typically render a final α_{95} of $\sim 5°$, then one should search for sites with counter-bedding of at least 30°. If the tectonic structures are not steep enough, then the paleomagnetist must take more sites to try to reduce α_{95}.

Table 2
Example data for MSL: Lupata Volcanics [11]

Site	D_G (°)	I_G (°)	D_S (°)	I_S (°)	k	α_{95} (°)	N	Strike (°)	Dip (°)
1	346.7	−60.9	337.6	−53.5	667	2.0	9	35	10
2	346.2	−60.3	337.0	−52.9	941	1.6	9	35	10
3	338.2	−64.6	329.5	−56.1	544	2.2	9	35	10
4	332.2	−59.5	326.8	−50.3	137	5.2	7	35	10
5	356.3	−64.7	343.9	−58.0	1800	1.2	9	35	10
6	344.8	−54.9	339.0	−56.7	222	3.7	8	320	5
7	344.8	−50.2	339.0	−52.6	49.3	6.9	10	320	5

KR: $k_S/k_G = 2.09 < [2.69]$, KR tilt test indeterminate; MSL: $\mathrm{t.s.}_G = 1.24 > [0.82]$, $\mathrm{t.s.}_S = 0.072 < [0.82]$, MSL tilt test inferred positive; AC1: $\xi_G = |-3.99| > [3.09]$, $\xi_S = 0.17 < [3.09]$, AC1 tilt test inferred positive; AC2: $\xi_G = 5.12 > [3.09]$, $\xi_S = 1.20 < [3.09]$, AC2 tilt test inferred positive; OC: max k at $94.0 \pm 30.3\% \sim 100\%$, OC tilt test inferred positive; DC: max k at $95.0 \pm 63.9\% \sim 100\%$, DC tilt test inferred positive.
Note: see Table 1.

Table 3
Example data for isolated observation: paleozoic carbonates, Crowsnest Transect, Canadian Rockies front ranges [12]

Site	D_G (°)	I_G (°)	D_S (°)	I_S (°)	k	α_{95} (°)	N	Strike (°)	Dip (°)
DKK01	74.5	47.8	278.2	74.2	40.3	9.6	7	171.3	57.1
DKK73	46.4	50.1	330.7	73.9	141.6	4.1	10	161.5	40.2
DKK74	44.4	43.8	319.1	72.7	220.3	5.2	5	157.4	49.4
DKK77	48.6	57.0	0.4	64.5	26.5	9.6	10	189.8	26.6
OSF63	45.7	56.8	313.8	77.0	49.2	13.2	4	157.8	36.6
OSF64	257.3	47.2	3.3	80.6	77.1	9.0	5	335.1	46.6
OSG30	44.1	56.3	302.0	71.6	65.0	4.6	8	160.5	42.6
OSG31	45.0	46.7	313.0	65.0	147.7	7.9	11	166.1	53.2

KR: $k_S/k_G = 7.95 > [2.48]$, KR tilt test inferred positive; IO: $\gamma_G = 78.4° > [18.8°]$, $\gamma_S = 12.3° < [20.0°]$, IO tilt test inferred positive; AC1: $\xi_G = |-7.22| > [3.30]$, $\xi_S = |-2.80| < [3.30]$, AC1 tilt test inferred positive; AC2: $\xi_G = 6.04 > [3.30]$, $\xi_S = 4.40 > [3.30]$, AC2 tilt test indeterminate; OC: max k at $83.8 \pm 7.6\%$ different from 100%, OC tilt test indeterminate; DC: max k at $85.0 \pm 19.3\% \sim 100\%$, DC tilt test inferred positive.
Note: see Table 1. The critical value for the IO test is the angular distance from the mean beyond which the isolated observation has a < 5% chance of being found if the directions are Fisher-distributed.

4.4. Application to real data

Real data are used to illustrate how the different tests behave in different settings, and the data are tabulated to help programmers verify their calculations. As an example of a two-limb study, consider the Lupata Volcanics [11] which was used by McElhinny [5] as a data set which the KR test infers does not pass the tilt test. With an interlimb angle of only ~15°, the study provides a marginal application of the fold test. However the dispersion is so low that the tilt test is able to discern between pre- and post-tilting remanence. As Table 2 shows, only the KR test fails to infer a positive tilt test with these data.

In the study of Cretaceous remagnetizations in the southern Canadian Rockies [12], the Crowsnest Transect provided an isolated observation bedding geometry. Only one site had counter-bedding from a small drag fold in an otherwise monoclinal sequence. In this case, the AC2 test and the OC tilt tests are indeterminate (Table 3). The AC2 test fails because the single NE-dipping site is swamped by the SW-dipping sites. The

Table 4
Example data for syn-tilting remanence acquisition: Cretaceous Ventura member in Manning Park, southern Canadian Cordillera [13]

Site	D_G (°)	I_G (°)	D_S (°)	I_S (°)	k	α_{95} (°)	N	Strike (°)	Dip (°)
01	48.0	54.1	354.7	63.6	85.0	5.0	11	186.0	31.4
02	41.8	59.0	341.6	62.7	68.6	5.9	10	186.0	31.4
03	38.9	51.4	350.4	61.5	56.4	6.1	11	178.1	31.7
04	38.4	52.1	358.7	64.4	140.2	3.9	12	172.9	26.2
05	24.9	67.1	13.1	52.1	70.7	4.4	16	264.4	16.4
06	43.2	64.1	37.5	53.0	135.2	3.4	14	293.0	11.6
07	23.4	63.0	9.2	49.9	71.8	6.1	10	250.6	16.0
08	39.6	63.8	20.8	52.8	106.3	4.2	12	253.8	15.8

KR: $k_G/k_S = 2.34 < [2.48]$, KR tilt test indeterminate; MSL: $\text{t.s.}_G = 2.13 > [0.65]$, $\text{t.s.}_S = 3.05 > [0.65]$, MSL tilt test indeterminate; AC1: $\xi_G = 5.11 > [3.30]$, $\xi_S = 6.52 > [3.30]$, AC1 tilt test indeterminate (AC1 $\xi_S = 0$ at 40.5% untilting); AC2: $\xi_G = |-5.57| > [3.30]$, $\xi_S = 7.57 > [3.30]$, AC2 tilt test indeterminate (AC2 $\xi_S = 0$ at 41.6% untilting); OC: max k at $35.4 \pm 8.1\%$ different from 0% and 100%, OC tilt test indeterminate (syn-tilting conclusion); DC: max k at $36.2 \pm 18.6\%$ different from 0% and 100%, DC tilt test indeterminate (syn-tilting conclusion).
Note: see Table 1.

OC test fails because the individual sites do not include all sources of dispersion.

Finally, the Ventura member red beds of the Methow Basin of British Columbia and Washington hold a syn-tilting magnetization at several localities [13]. In Manning Park, a two-limb fold was sampled (Table 4). All the tests reveal indeterminate results between positive and negative tilt tests, leading to a syn-tilting interpretation. The γ_{max} determined numerically (OC) and analytically (DC) are almost identical, but quite different from the zero-correlation point for the AC tests. The OC confidence interval is about half the DC confidence interval, but the latter is preferred because it accounts for all sources of dispersion.

5. Conclusion

The DC tilt test formulation provides an analytical method of calculating the degree of bedding correction which renders the least dispersion of site mean directions. It also tests whether or not the site directions are correlated to the bedding corrections. Thus it combines and improves aspects of both Watson's OC and McFadden's AC tests.

The arithmetic for the DC test is not difficult to program (Appendix 1). Graphical output provides the researcher with a view of which sites are the most important for the tilt test, and which sites may be outliers, making this test less of a 'black box' method. Paleomagnetism analysis programs which include the DC test can be downloaded: http://www.pgc.nrcan.gc.ca/tectonic/enkin.htm.

Simulations show that the DC test is statistically valid, and it is applicable to all bedding geometries. It is the most powerful of all tilt test formulations, in that it requires the least bedding structure to succesfully infer a positive (or negative) tilt test. Thus the DC tilt test could come into general use as an all-purpose tilt test for paleomagnetic studies.

Acknowledgements

Thanks to Vincent Courtillot for encouraging me to develop the DC test. Phil McFadden's work provided the initial inspiration, and his comments on an early version of this manuscript aided the final version greatly. Thanks to Ted Irving for sharing his experience, intuition and precision. Judith Baker, as well as shepherding through thousands of simulations, was a sounding board for these ideas. Stuart Gilder and Frank Lowes provided excellent reviews which improved the paper. This is GSC publication 2002261. ***[VC]***

Appendix 1. Example computer code

The following subroutine, written in BASIC, calculates the DC slope and tests it against slopes 1 and 0. For N sites, the unit vector remanence direction in geographic coordinates is given by GX(I), GY(I) and GZ(I), and the bedding attitude is given by Strike(I), Dip(I). The unit vector means in geographic and stratigraphic coordinates are given by GMeanX, GMeanY, GMeanZ and SMeanX, SMeanY, SMeanZ.

```
   SumCC = 0. : SumCD = 0. : SumDD = 0.
   FOR I = 1 TO N
     'determine BX, BY, BZ, the back correction of StratiMean
        CALL BedCorrection (SMeanX, SMeanY, SMeanZ, Strike(I), -Dip(I), BX, BY, BZ)
     'C = angular distance from GeoMean to back-correction of StratiMean (Equation 1)

   CosC = GMeanX*BX + GMeanY*BY + GMeanZ*BZ
   C(I) = Arccos (CosC)
 'D = projection of GeoSite-GeoMean arc onto BackCorrection-GeoMean arc (Equation 3)
   MeanDotG = GMeanX*GX(I) + GMeanY*GY(I) + GMeanZ*GZ(I)
   MeanCrossGX = GMeanY*GZ(I) - GMeanZ*GY(I)
   MeanCrossGY = GMeanZ*GX(I) - GMeanX*GZ(I)
   MeanCrossGZ = GMeanX*GY(I) - GMeanY*GX(I)
   MeanCrossBX = GMeanY*BZ - GMeanZ*BY
   MeanCrossBY = GMeanZ*BX - GMeanX*BZ
   MeanCrossBZ = GMeanX*BY - GMeanY*BX
   MeanCrossBM = SQR (MeanCrossBX^2 + MeanCrossBY^2 + MeanCrossBZ^2)
   MeanCrossGDotMeanCrossB = MeanCrossGX*MeanCrossBX + MeanCrossGY*MeanCrossBY + _
                                                 MeanCrossGZ*MeanCrossBZ
   IF MeanDotG = 0. THEN  'site direction 90° from mean
        D(I) = 90.
   ELSEIF MeanCrossBM = 0. THEN  'back-correction coincedent or antipodal with GeoMean
        D(I) = Arccos (MeanDotG)
   ELSE
        TanD = MeanCrossGDotMeanCrossB / MeanDotG / MeanCrossBM
        D(I) = Arctan (TanD)
        IF MeanDotG < 0. AND TanD < 0. THEN D(I) = D(I) + 180.   'D>90° from mean
   END IF
 'Calculate DC slope and confidence interval (Equations 4-6)
   SumCC = SumCC + C(I)*C(I)
   SumCD = SumCD + C(I)*D(I)
   SumDD = SumDD + D(I)*D(I)
NEXT I
Slope = SumCD / SumCC                                                 '(Equation 4)
Sigma = SQR( (SumDD - SumCD*SumCD/SumCC) / (SumCC*(N-2.)) )        '(Equation 5)
DeltaSlope = StudentsT(N-2) * Sigma                                   '(Equation 6)
 'Compare DC slope to 1 (for positive test) and 0 (for negative test). (Equations 7-10)
SlopeMax = Slope + DeltaSlope
SlopeMin = Slope - DeltaSlope
IF (SlopeMax-1.) * (SlopeMin-1.) < 0. THEN
     IF SlopeMin < 0. THEN
          PRINT "Indeterminate"
     ELSE
          PRINT "Positive"
     END IF
ELSEIF SlopeMax * SlopeMin < 0. THEN
     PRINT "Negative"
ELSE
     PRINT "Indeterminate"
END IF
```

Appendix 2. Statistical assumptions and justifications

Least squares linear regression is based on several statistical assumptions. In the DC test, some assumptions are not perfectly fulfilled, however these problems are here shown to be negligible.

A2.1. The linear model

The test assumes that the DC data follow the model of straight line passing through the origin; that is $d_i = sc_i + e_i$, where the e_i are errors in the data. Examination of Fig. 1 (especially data for site 2) shows that $c_i \approx 0$ occurs when the corresponding site direction is close to the mean directions in both geographic and stratigraphic coordinates and thus the d_i will be close to 0. If the paleomagnetic dispersion is small and the remanence predates tilting, then the d_i and c_i will be almost identical, so the DC plot will be linear through the origin. A problem arises when the paleomagnetic dispersion is greater than the differences in the bedding attitudes, in which case the the proportionality is not apparent. This occurs when there are insufficient bedding attitude differences between sites, thus making the study inappropriate for application of the tilt test.

A2.2. All dispersion is in the dependent variable

The least squares estimate of slope (Eq. 4) assumes that there is an independent variable with no uncertainty, and that all scatter is found in the dependent variable. For the DC test, d is put on the vertical axis because there is usually more scatter in the remanent directions than in the measurement of the bedding orientations.

While this is usually the case, sometimes paleohorizontal is difficult to measure. In such a case, the tilt test is compromised. In the extreme case of paleohorizontal differences being smaller than the uncertainty in the bedding measurements, a positive tilt test would appear negative. The appendix to the study of the Spences Bridge volcanics [14] demonstrates a case where uncertainty in bedding measurements in volcanic flows leads to an ambiguous tilt test in a paleomagnetic data set, while conglomerate and contact tests indicate primary (pre-tilting) remanence acquisition. The conclusion is that large bedding uncertainties tend to make a pre-tilting remanence appear to be syn-tilting in age.

A2.3. Uniformity and Gaussian distribution of residuals

Residuals are the differences between the data and the mathematical model. They are uniform if their mean is zero and their scatter is the same for

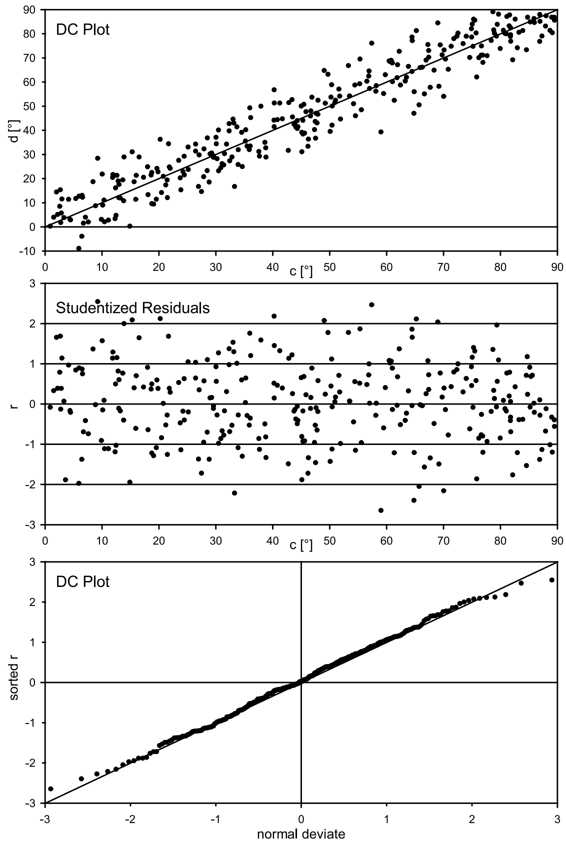


Fig. 4. Demonstration that the DC plot is linear through the origin, with Gaussian-distributed residuals. (a) DC plot for a sample of 300 sites with random bedding dips from 0 to 90°. (b) Studentized residuals about the best fitting line through the origin. (c) Quantile–quantile plot shows the goodness of fit to a Gaussian distribution.

any subsample of observations (no heteroscedascity). The principal of least squares does not require that the residuals be Gaussian-distributed, but this approximation underlies the confidence interval estimate. These characteristics are probably impossible to derive analytically, however they are very clear in simulations with large numbers of sites. Consider Fig. 4, simulated using 300 sites. A Fisher distribution ($k = 30$) was given random bedding tilts up to 90°. By inspection, the data on the DC plot (Fig. 4a) follow a band of constant width above and below the line $d = c$. Studentized residuals (Fig. 4b) measure how many standard deviations each point is away from the line when that line is fit without that point. In this example, 14 of 300 data are greater than two standard deviations from the mean, as expected for a Gaussian distribution. There is no evidence of serial correlations within the data set. The fit to the Gaussian distribution is made clear using a quantile–quantile plot (Fig. 4c) that compares the sorted Studentized residuals with Gaussian deviates (the expected position of the N data where they are perfectly Gaussian-distributed).

When dispersion is high, the assumptions do not hold. In the extreme case of site directions uniformly distributed over the sphere, no correlation will be seen with the bedding. But for $k > 3$, all large N simulations have an apparently Gaussian distribution of residuals on the DC plot. With real data, it is always a good idea to inspect the Studentized residuals and Q–Q plot to verify that the model is well-followed.

A2.4. Independence of observations and number of degrees of freedom

Independence means that the plotted data include no information about each other. The DC plot uses quantities derived from the N site directions, however it is first necessary to determine the means of site directions in geographic and stratigraphic coordinates. If a site direction is changed, these means change, and thus the points on the DC plot change. This lack of independence does not invalidate the least squares estimates, but it must be included in the calculation of the confidence interval around the estimated DC slope.

This is analogous to the calculation of variance which first requires an estimate of the average. The variance is underestimated if the sum of square deviation is divided by N rather than the correct $N-1$ because one degree of freedom is taken before variance can be calculated. The degree of freedom is the total number of independent measures of errors in the residuals.

Standard least squares regression of a line passing through the origin uses $N-1$ degrees of freedom when calculating the confidence interval of the slope. In the DC test, calculating two means leaves $N-2$ degrees of freedom, leading to slightly larger confidence intervals. The resulting confidence intervals have been confirmed numerically.

Appendix 3. A one-dimensional analog of the DC test

The relation between the DC test and Watson's OC test is best explained using a one-dimensional analog. Consider two collections of points on an axis, analogous to the site directions on a sphere in geographic coordinates (0% untilting) and stratigraphic coordinates (100% untilting). In Fig. 5, the 0% axis is plotted vertically on the left and the 100% axis on the right. By connecting the corresponding points, each point traces out a line analogous to the small circle on the sphere traced out by a site direction during incremental bedding correction. The slopes of these lines are analogous to the bedding dips in the spherical case.

The distance from the mean to the ith point on the 0% axis corresponds to d_i defined above. In order to construct c_i, we take the mean on the 100% axis and track back to the 0% axis using a line with the ith slope. The distance from the 0% mean to the back correction of the 100% mean forms c_i (Fig. 5).

If the degree of untilting is x and the position along the line is y, then for each site we have:

$$y_i = m_i x + g_i, \tag{A1}$$

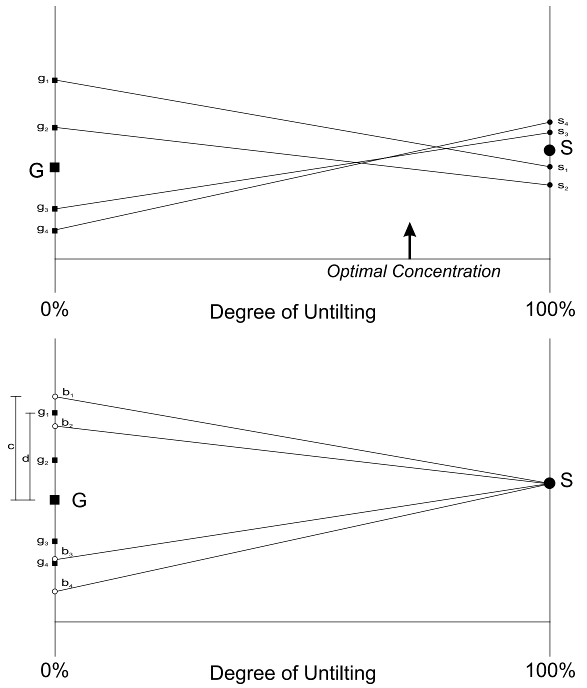


Fig. 5. Example of a one-dimensional analog to the DC test: The vertical axis on the left (right) corresponds to site directions in geographic (stratigraphic) coordinates and the slope of the line connecting points corresponds to the bedding dip in spherical geometry. Symbols are the same as in Fig. 1.

where g_i is the value of y_i at $x=0$ (i.e., the site direction in geographic coordinates) and m_i is the slope of the line corresponding to the bedding dip. The distance from the mean in geographic coordinates, $G=\Sigma g_j/N$, to the individual site directions b_i is:

$$d_i = g_i - \left(\sum g_j/N\right). \tag{A2}$$

Since the mean in stratigraphic coordinates is $S=\Sigma(m_j+g_j)/N$, the distance from G to the back correction of the mean is:

$$c_i = S - m_i - G = \left(\sum m_j/N\right) - m_i. \tag{A3}$$

On the DC plot, the slope of the least squares line passing through the origin is:

$$s = \sum d_i c_i / \sum c_i^2$$

$$= \frac{-\left(\sum m_i g_i - \sum m_i \sum g_i/N\right)}{\left(\sum m_i^2 - \left(\sum m_i\right)^2/N\right)} \tag{A4}$$

For the OC test, one must find the degree of untilting which provides maximum concentration, or equivalently minimum variance. For a degree of untilting x, the variance is given by:

$$\begin{aligned}\mathrm{var}(y) &= \sum (y_i \bar{y})^2/(N-1)\\ &= \{x^2\left(\sum m_i^2 - \left(\sum m_i\right)^2/N\right) +\\ &\quad 2x\left(\sum m_i g_i - \sum m_i \sum g_i/N\right) +\\ &\quad \left(\sum g_i^2 - \left(\sum g_i\right)^2/N\right)\}/(N-1),\end{aligned} \tag{A5}$$

where $\bar{y}(x) = (\Sigma m_i/N)x + \Sigma g_i/N$. The minimum variance is found at $\hat{x}_0$ when:

$$\begin{aligned}\mathrm{d}(\mathrm{var}(y))/\mathrm{d}x &= \{2x\left(\sum m_i^2 - \left(\sum m_i\right)^2/N\right) +\\ &\quad 2\left(\sum m_i g_i - \sum m_i \sum g_i/N\right)\}/(N-1) = 0.\end{aligned} \tag{A6}$$

Therefore:

$$\hat{x}_0 = \frac{-\left(\sum m_i g_i - \sum m_i \sum g_i/N\right)}{\left(\sum m_i^2 - \left(\sum m_i\right)^2/N\right)} \tag{A7}$$

which is identical to s (Eq. A4). Thus the degree of untilting which renders optimal concentration is equivalent to the slope of the DC plot.

References

[1] J.W. Graham, The stability and significance of magnetism in sedimentary rocks, J. Geophys. Res. 54 (1949) 131–167.
[2] P.L. McFadden, A new fold test for palaeomagnetic studies, Geophys. J. Int. 103 (1990) 163–169.
[3] G.S. Watson, R.J. Enkin, The fold test in paleomagnetism as a parameter estimation problem, Geophys. Res. Lett. 20 (1993) 2135–2137.
[4] J.R. Green, D. Margerison, Statistical Treatment of Experimental Data, Elsevier, Amsterdam, 1978, 382 pp.
[5] M.W. McElhinny, Statistical significance of the fold test in palaeomagnetism, Geophys. J. R. Astron. Soc. 8 (1964) 338–340.
[6] P.L. McFadden, D.L. Jones, The fold test in palaeomagnetism, Geophys. J. R. Astron. Soc. 67 (1981) 53–58.
[7] P.L. McFadden, The fold test as an analytical tool, Geophys. J. Int. 135 (1998) 329–338.
[8] N.I. Fisher, P. Hall, New statistical methods for directional data; I, Bootstrap comparison of mean directions and the fold test in palaeomagnetism, Geophys. J. Int. 101 (1990) 305–313.
[9] N.I. Fisher, P. Hall, A general statistical test for the effect of folding, Geophys. J. Int. 105 (1991) 419–427.
[10] L. Tauxe, G.S. Watson, The fold test; an eigen analysis approach, Earth Planet. Sci. Lett. 122 (1994) 331–341.

[11] D.I. Gough, N.D. Opdyke, The palaeomagnetism of the Lupata Alkaline Volcanics, Geophys. J. R. Astron. Soc. 7 (1963) 457–468.
[12] R.J. Enkin, K.G. Osadetz, J. Baker, D. Kisilevsky, Orogenic remagnetizations in the front ranges and inner foothills of the southern Canadian Cordillera: chemical harbinger and thermal handmaiden of Cordilleran deformation, Geol. Soc. Am. Bull. 112 (2000) 929–942.
[13] R.J. Enkin, J.B. Mahoney, J. Baker, M. Keissling, R.A. Haugerud, Syntectonic remagnetization in the southern Methow block: Resolving large displacements in the southern Canadian Cordillera, Tectonics 21 (2002) 10.1029/2001TC001294.
[14] M.L. Haskin, R.J. Enkin, J.B. Mahoney, P.S. Mustard, J. Baker, Deciphering shallow paleomagnetic inclinations – 1: Implications from correlation of Albian volcanic rocks along the Insular/Intermontane Superterrane boundary in the southern Canadian Cordillera, J. Geophys. Res. 108 (2003) 10.1029/2002JB001982.