Geophysical Journal International

Geophys. J. Int. (2010) **182,** 753–771 doi: 10.1111/j.1365-246X.2010.04671.x

Maximum likelihood solution for inclination-only data in paleomagnetism

Þ. Arason[1] and S. Levi[2]

[1] *The Icelandic Meteorological Office – Veðurstofa Íslands, Reykjavík, Iceland. E-mail: arason@vedur.is*
[2] *College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA*

Accepted 2010 May 20. Received 2010 January 5; in original form 2009 September 16

SUMMARY

We have developed a new robust maximum likelihood method for estimating the unbiased mean inclination from inclination-only data. In paleomagnetic analysis, the arithmetic mean of inclination-only data is known to introduce a shallowing bias. Several methods have been introduced to estimate the unbiased mean inclination of inclination-only data together with measures of the dispersion. Some inclination-only methods were designed to maximize the likelihood function of the marginal Fisher distribution. However, the exact analytical form of the maximum likelihood function is fairly complicated, and all the methods require various assumptions and approximations that are often inappropriate. For some steep and dispersed data sets, these methods provide estimates that are significantly displaced from the peak of the likelihood function to systematically shallower inclination. The problem locating the maximum of the likelihood function is partly due to difficulties in accurately evaluating the function for all values of interest, because some elements of the likelihood function increase exponentially as precision parameters increase, leading to numerical instabilities. In this study, we succeeded in analytically cancelling exponential elements from the log-likelihood function, and we are now able to calculate its value anywhere in the parameter space and for any inclination-only data set. Furthermore, we can now calculate the partial derivatives of the log-likelihood function with desired accuracy, and locate the maximum likelihood without the assumptions required by previous methods. To assess the reliability and accuracy of our method, we generated large numbers of random Fisher-distributed data sets, for which we calculated mean inclinations and precision parameters. The comparisons show that our new robust Arason–Levi maximum likelihood method is the most reliable, and the mean inclination estimates are the least biased towards shallow values.

Key words: Numerical approximations and analysis; Probability distributions.

1 INTRODUCTION

The Fisher distribution is the analogy to the normal distribution for a population of spherically distributed unit vectors. It was first derived by Langevin (1905) as the natural distribution for the alignment of an assemblage of identical magnetic moments of a paramagnetic gas in a uniform external magnetic field. Using Fisher statistics, one can obtain an unbiased estimate of the true mean direction of a sample drawn from such a distribution (Fisher 1953). The directions are often recorded as declinations (horizontal azimuth) and inclinations (the angle from the horizontal). In some studies only the inclinations are available. For example, paleomagnetic directions from borecores usually lack declinations, although the measured inclinations can reliably record past geomagnetic fields.

Briden & Ward (1966) showed that for inclination-only data, the arithmetic mean is biased towards shallow inclinations. Fig. 1 shows that the geometry of the sphere dictates that any isotropic distribution will be represented by more shallow inclinations than by steep inclinations when compared to the true mean.

In paleomagnetic applications the inclination-shallowing caused by using the arithmetic mean is usually less than a few degrees. For individual studies such a discrepancy is not very important and is usually well within the confidence limits of the study. However, because this is a one-sided bias, attempts to combine results of many studies may lead to more significant, systematic errors. For example, studies of long-term non-dipole components of the geomagnetic field require averaging the mean inclinations of many studies which might lead to a long-term effect of 1–2°. Improper procedures for estimating the mean inclinations in individual studies can seriously affect such estimates. In Fig. 2, we show contours of the approximate inclination-bias for a Fisher-distributed sample of inclination-only data.

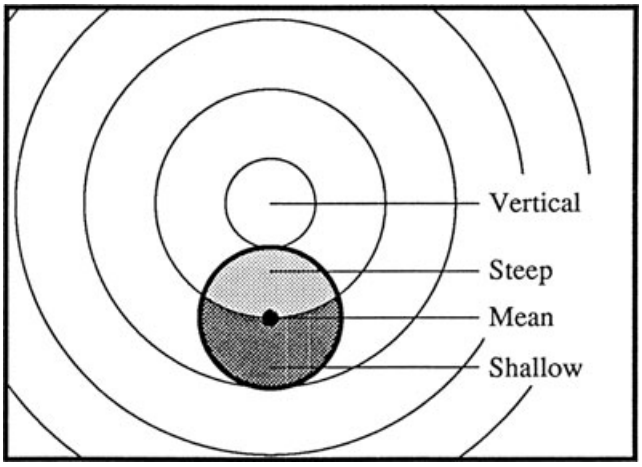

Figure 1. The geometry of the sphere dictates that any isotropic distribution about a true mean will be represented by more shallow than steep inclinations as compared to the mean. This illustration shows that the area of shallow inclinations (dark shading) is greater than the area represented by the steep inclinations (light shading). Arithmetic mean of inclination-only data will therefore result in a too shallow estimate of the mean. From Arason (1991, fig. 5.1, p. 207).

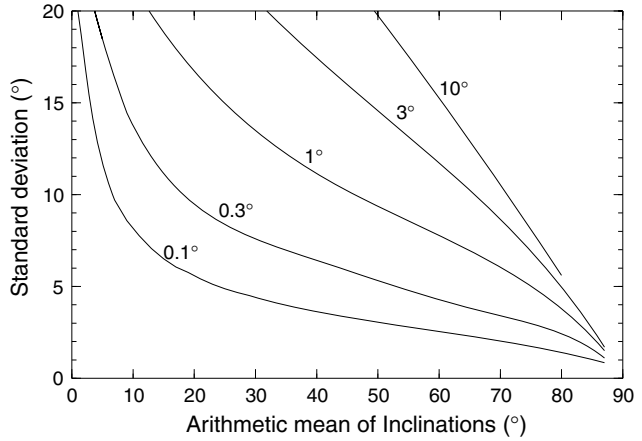

Figure 2. Contours of the approximate inclination shallowing of the mean for the arithmetic mean and ordinary standard deviation of a Fisher-distributed sample of inclination-only data.

Briden & Ward (1966) derived the marginal distribution for inclination-only data and its likelihood function. For a known dispersion of the distribution one can, in principle, make a simple correction for the inclination bias. The task is to simultaneously identify the mean inclination and the precision parameter, which is a measure of the spread of the distribution. Briden & Ward (1966) presented a graphical method to estimate the maximum likelihood mean inclination and precision parameter. Although the method is quite simple to use, it would be useful to be able to calculate the values directly from the observed data on a computer.

Harrison (1974) made a correction for the inclination-bias by comparing the standard deviation of the inclinations to those of randomly generated Fisher-distributed directions with a known mean and dispersion. Matching the arithmetic mean inclination and standard deviation of the randomly generated data set to the observed data, he estimated the mean inclination and precision parameter. Another early method assumes that the dispersion of the data is related to secular variation of the geomagnetic field, and knowledge of the field's dispersion is used to constrain the true precision parameter; a correction term is then applied to the arithmetic mean inclination (Peirce 1976; Cox & Gordon 1984). This method is only applicable to paleomagnetic results from lava flows, which represent spot recordings of the geomagnetic field, and where each inclination represents a single flow (also, care must be taken that the data adequately represent the spectrum of geomagnetic secular variation.). Clark (1983, 1988) studied the properties of the marginal Fisher-distribution and biases of various statistics. Westphal *et al.* (1998) suggested a graphical method to estimate the mean inclination and verify that the data are Fisherian.

In the early 1980s, two different numerical methods (Kono 1980; McFadden & Reid 1982) claimed to solve the inclination-only problem outlined by Briden & Ward (1966). Kono (1980) equated the expectation values of $\cos\theta$ and $\cos^2\theta$ of the distribution to the data, where θ is the co-inclination ($\theta = 90^\circ - I$) and I is the inclination. In principle, this is a method of moment estimation and is asymptotically unbiased as the sample size increases. For some data sets of steep and dispersed inclinations, the Kono method gives no solution. McFadden & Reid (1982) criticized the Kono method, both on theoretical grounds and because it resulted in obviously biased estimates for steep and dispersed inclinations. They suggested instead to solve the maximum likelihood problem. As they were unable to solve the exact formulas, they made some simplifications. Unfortunately, there was an error in one of their key equations (McFadden & Reid 1982, eq. 40, p. 317). Although this error was modified by several workers, some paleomagnetists are still using the erroneous original method, and we are not aware of a published correction to the method. An outline of the McFadden–Reid method can be found in a book where the authors corrected the equation without mentioning that it is different from the original work (McElhinny & McFadden 2000, eq. 3.2.32, p. 97). Even with the correction, the McFadden–Reid method is based on approximations that are inappropriate for dispersed steep inclinations, leading to inclination biases.

Both the Kono and McFadden–Reid methods gained popularity in paleomagnetism. The McFadden–Reid method gives results which are almost identical to the arithmetic mean, and it should not be used in its original form. Both the modified McFadden–Reid and the Kono methods give reasonable results for true inclinations below 60°, provided the precision parameter is greater than 10 (Arason 1991; Arason & Levi 1995; Levi & Arason 2006).

Enkin & Watson (1996) presented a new approach. They weighted the likelihood function with a Bayesian factor. This weighting gives better constraints for solutions of very dispersed and steep inclinations. They presented three approaches depending on the dispersion and steepness of the data: arithmetic mean, Gaussian estimate and marginal likelihood. Enkin & Watson (1996) used better approximations to the Bessel functions than McFadden & Reid (1982).

These methods use various assumptions and approximations that turn out to be inappropriate and lead to biases for many data sets. In this study, we describe a new method for directly evaluating the maximum likelihood estimates of the mean inclination and the precision parameter. Similar to previous studies for estimating the mean inclination of inclination-only data, we assume that the directions are Fisher-distributed. We also discuss the effects when the directions are not Fisherian. Evaluation of the likelihood function and its derivatives is problematic. For ordinary paleomagnetic data, the direct evaluation of these functions includes exponential terms that can potentially lead to overflow in any ordinary programming language. We were successful in analytically cancelling these exponential terms from the likelihood functions, and, using more detailed

computation of the Bessel and other functions, we are now able to accurately calculate the location of the maximum of the likelihood function. The new Arason–Levi method is robust, and it gives accurate maximum likelihood estimates of the mean inclination and precision parameter for any data set. Extensive comparisons between the various methods using randomly generated, Fisher-distributed data sets favour our method.

2 THE LIKELIHOOD FUNCTION

The joint density function of the Fisher distribution is

$$g(\theta, \varphi) = \left(\frac{\kappa \sin\theta}{4\pi \sinh\kappa}\right) \exp[\kappa \cos\theta_0 \cos\theta + \kappa \sin\theta_0 \sin\theta \cos(\varphi - \varphi_0)], \quad (1)$$

where κ is the precision parameter describing the concentration of the distribution, θ is the angle from the vertical down ($\theta = 90^\circ - I$, where I is inclination) and φ is the azimuthal angle, and (θ_0, φ_0) represents the true mean direction. The derivation of the likelihood function and its derivatives, presented here, is similar to previous analyses of inclination-only data (Briden & Ward 1966; Kono 1980; McFadden & Reid 1982; Clark 1983; Enkin & Watson 1996).

For inclination-only data of sample size N, we only have the observed co-inclinations θ_i: $\theta_1, \ldots, \theta_N$, and not the azimuths, and the data must be regarded as a random sample from the marginal distribution of eq. (1)

$$f(\theta) = \int_0^{2\pi} g(\theta, \varphi)\, \mathrm{d}\varphi = \left(\frac{\kappa \sin\theta}{2 \sinh\kappa}\right) \exp[\kappa \cos\theta_0 \cos\theta] I_0(\kappa \sin\theta_0 \sin\theta), \quad (2)$$

where $I_0(x)$ is the hyperbolic Bessel function of order zero, also called the modified Bessel function of first kind and order zero.

It is worth noting that the bias in the arithmetic mean inclination can be calculated by

$$\Delta I = \int_0^{\pi} \theta f(\theta)\, \mathrm{d}\theta, \quad (3)$$

where ΔI is the inclination bias and $f(\theta)$ is from eq. (2). This bias was calculated numerically for selected values of κ, shown in Fig. 3.

The likelihood function is

$$H(\theta, \kappa) = \prod_{i=1}^{N} f(\theta_i) = \left(\frac{\kappa}{2 \sinh\kappa}\right)^N \prod_{i=1}^{N} \sin\theta_i \exp(\kappa \cos\theta \cos\theta_i) I_0(\kappa \sin\theta \sin\theta_i) \quad (4)$$

and the log-likelihood function is

$$h(\theta, \kappa) = \ln[H(\theta, \kappa)] = N \ln\left[\frac{\kappa}{2 \sinh\kappa}\right] + \sum_{i=1}^{N} (\kappa \cos\theta \cos\theta_i + \ln[I_0(\kappa \sin\theta \sin\theta_i)]) + \sum_{i=1}^{N} \ln[\sin\theta_i]. \quad (5)$$

The likelihood function H is always positive for $0 \le \kappa$, $0^\circ \le \theta \le 180^\circ$, $0^\circ < \theta_i < 180^\circ$. Because $\ln(x)$ is a monotonically increasing function, for $x > 0$, both functions H and h have maximum values at the same location (θ, κ), where

$$\frac{\partial h}{\partial \theta} = 0 \quad \text{and} \quad \frac{\partial h}{\partial \kappa} = 0, \quad (6)$$

$$\frac{\partial h}{\partial \theta} = -\kappa \sin\theta \sum_{i=1}^{N} \cos\theta_i + \kappa \cos\theta \sum_{i=1}^{N} \left(\sin\theta_i \left[\frac{I_1(\kappa \sin\theta \sin\theta_i)}{I_0(\kappa \sin\theta \sin\theta_i)}\right]\right) = 0, \quad (7)$$

$$\frac{\partial h}{\partial \kappa} = N/\kappa - N \coth\kappa + \cos\theta \sum_{i=1}^{N} \cos\theta_i + \sin\theta \sum_{i=1}^{N} \left(\sin\theta_i \left[\frac{I_1(\kappa \sin\theta \sin\theta_i)}{I_0(\kappa \sin\theta \sin\theta_i)}\right]\right) = 0, \quad (8)$$

where $I_1(x)$ is the hyperbolic Bessel function of order one.

There is a distinct possibility that the maximum occurs at the boundary ($\theta = 0^\circ$ or 180°), without satisfying eq. (6). Numerical simulations show that this is sometimes the case. Therefore, we are also interested in the κ that maximizes the likelihood function at

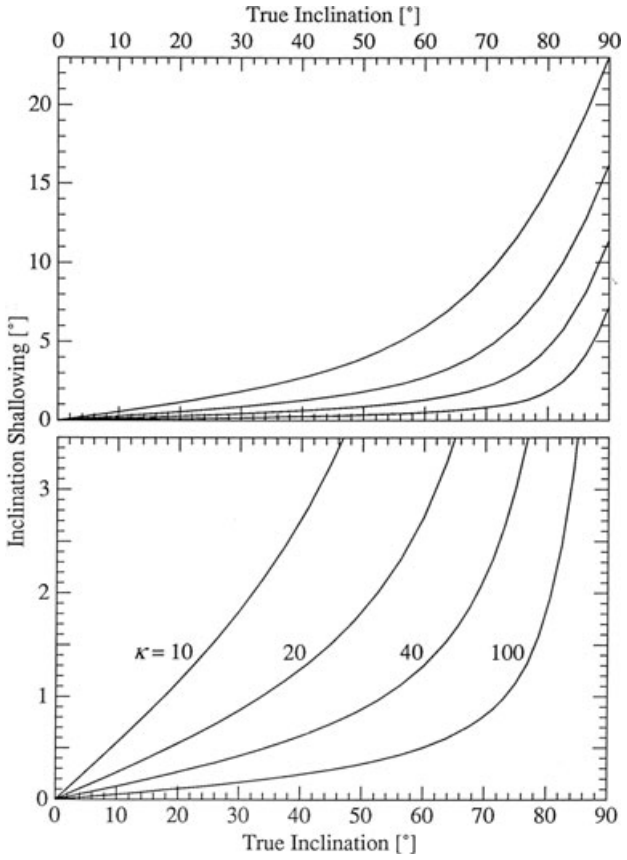

Figure 3. The inclination shallowing of the mean resulting from the arithmetic mean of inclination-only data versus the true inclination for Fisher-distributed data. Values of the precision parameter κ are 10, 20, 40 and 100. Expanded scale is shown on the lower panel. For high κ or low inclination this effect is not very serious. If κ were known, one could apply a simple correction to arithmetic means of inclination-only data. However, κ is usually not known and one needs to estimate κ and I simultaneously. From Arason (1991, fig. 5.2, p. 209).

the boundary

$$\frac{\partial h}{\partial \kappa} = N/\kappa - N\coth\kappa + \cos\theta \sum_{i=1}^{N} \cos\theta_i = 0, \tag{9}$$

$$\coth\kappa - 1/\kappa = L(\kappa) = \cos\theta \frac{1}{N} \sum_{i=1}^{N} \cos\theta_i, \tag{10}$$

where $L(\kappa)$ is the Langevin function, and $\cos\theta = \pm 1$ for $\theta = 0^\circ$ or 180°.

Although it is probably never the case, we finally consider the possibility that the maximum is on the other boundary, $\kappa = 0$. This represents a case of completely (and exactly) random directions. See Appendix A for calculating the log-likelihood function for this case. On this boundary the function is independent of θ and has the value

$$h(\theta, \kappa = 0) = -N\ln[2] + \sum_{i=1}^{N} \ln[\sin\theta_i]. \tag{11}$$

3 EVALUATION OF ESSENTIAL FUNCTIONS

Problems with previous methods are often caused by the rather poor approximations of some of the functions needed to calculate the likelihood function and its derivatives. Also, some of the methods require somewhat *ad hoc* approaches for different data sets. To be consistent, we use the same method for all data sets. In addition, we use the most accurate calculations available for determining the likelihood function and its derivatives to prevent introducing artificial effects into the solution. Moreover, we do not want to arbitrarily restrict the iteration process from some undesired region of the parameter space (θ, κ). Our results show that we can calculate the likelihood function and its derivatives with desired accuracy for all regions of the parameter space.

The hyperbolic Bessel functions, $I_0(x)$ and $I_1(x)$ cannot be expressed as a finite combination of elementary functions. Some previous studies of the inclination-only problem used crude approximations of the functions, leading to inaccurate estimates of the mean inclination. Direct evaluation of the Bessel functions is problematic because they increase exponentially, and we may need to evaluate the functions for very high values, for example $I_0(1000) \approx 2.5 \times 10^{432}$. Such direct evaluation of the functions will lead to inaccurate results or overflow problems in any ordinary computer programming language for some real paleomagnetic data sets. As shown in Appendix A, these functions can be accurately evaluated for any combination of the parameters $(\kappa, \theta, \theta_i)$.

In Appendix A, we calculate the Bessel function, $I_0(x)$, and the ratio $I_1(x)/I_0(x)$ using the approximations of Press *et al.* (1989) and Olver (1972, eqs 9.8.1–4). To avoid inaccuracies and overflow problems due to the exponential behaviour of the functions, we define the functions $B(x)$ and $R(x)$

$$B(x) = I_0(x)/\mathrm{e}^x, \tag{12}$$

$$R(x) = I_1(x)/I_0(x) = \frac{I_1(x)/\mathrm{e}^x}{I_0(x)/\mathrm{e}^x}. \tag{13}$$

In Appendix A, we show that the functions $B(x)$ and $R(x)$ can be accurately calculated for all values of x without ever evaluating the exponential element. For $0 \leq x$, the functions take values in the range $0 < B(x) \leq 1$ and $0 \leq R(x) < 1$. Although, Enkin & Watson (1996) use a better approximation to the Bessel functions than McFadden & Reid (1982), their approximation may still give rise to inaccuracies up to 3.7 per cent ($|\varepsilon| > 0.02$). Our approximations to the Bessel functions are many orders of magnitude more accurate, for example error of our $B(x)$ is always $|\varepsilon| < 2 \times 10^{-7}$. Fig. 4 shows the functions $B(x)$ and $R(x)$.

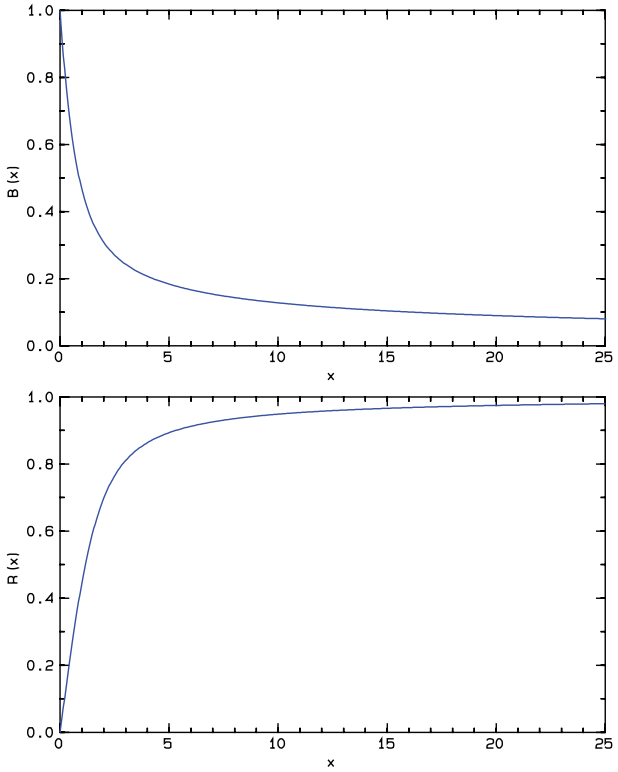

Figure 4. The functions $B(x)$ and $R(x)$ are easily calculated for any value of x. They are used to evaluate the hyperbolic Bessel functions $I_0(x)/\mathrm{e}^x = B(x)$ and $I_1(x)/I_0(x) = R(x)$.

Eq. (8) includes a hyperbolic cotangent term, coth κ. Our representation of the hyperbolic cotangent function is shown in Appendix A.

We evaluate the log-likelihood function $h(\theta, \kappa)$ of eq. (5) for any value $0 \leq \kappa$ and $0^\circ \leq \theta \leq 180^\circ$ for any given data set of θ_i, $i = 1, \ldots, N$.

$$\begin{aligned} h(\theta, \kappa) &= N\ln\left[\frac{\kappa}{2\sinh\kappa}\right] \\ &\quad + \sum_{i=1}^{N} (\kappa\cos\theta\cos\theta_i + \ln[I_0(\kappa\sin\theta\sin\theta_i)]) \\ &\quad + \sum_{i=1}^{N} \ln[\sin\theta_i] \\ &= A_1 + A_2 + A_3. \end{aligned} \tag{14}$$

The first term (A_1) can be calculated for any value of κ, by a careful choice of approximations depending on κ. The exponential element of the Bessel function in the second term (A_2) can be analytically cancelled by using the function $B(x)$ of eq. (12).

$$A_2 = \sum_{i=1}^{N} (\kappa\cos\theta\cos\theta_i + \ln[\exp(\kappa\sin\theta\sin\theta_i)B(\kappa\sin\theta\sin\theta_i)]), \tag{15}$$

$$= \sum_{i=1}^{N} (\kappa\cos(\theta_i - \theta) + \ln[B(\kappa\sin\theta\sin\theta_i)]). \tag{16}$$

For any given data set, the last term (A_3) of eq. (14) is a constant and will therefore not affect the location of the maximum of the likelihood function. There is a problem if one of the observed values is vertical, that is $\theta_i = 0^\circ$ or 180° (ln 0). Otherwise, A_3 can directly be calculated. The details of accurately evaluating the log-likelihood function is shown in Appendix A.

4 OUR METHOD OF LOCATING THE MAXIMUM

Similar to the Gaussian estimates of Enkin & Watson (1996), we take the arithmetic mean of the co-inclinations as our initial guess for θ

$$\hat{\theta}_0 = \bar{\theta} = \frac{1}{N}\sum_{i=1}^{N}\theta_i \tag{17}$$

and the inverse variance of the co-inclinations as our initial guess for κ.

$$\hat{\kappa}_0 = \left(\frac{1}{N-1}\sum_{i=1}^{N}\left(\theta_i - \bar{\theta}\right)^2\right)^{-1}, \tag{18}$$

where the θs are in radians. Arranging eq. (7), we iterate $\hat{\theta}_j$ to the next $\hat{\theta}_{j+1}$, using $\hat{\kappa}_j$ as our best estimate of κ

$$\tan\hat{\theta}_{j+1} = \frac{\sum_{i=1}^{N}\left(\sin\theta_i\left[\frac{I_1(\hat{\kappa}_j\sin\hat{\theta}_j\sin\theta_i)}{I_0(\hat{\kappa}_j\sin\hat{\theta}_j\sin\theta_i)}\right]\right)}{\sum_{i=1}^{N}\cos\theta_i}. \tag{19}$$

In evaluating the Bessel function ratio $I_1(x)/I_0(x)$, we use the function $R(x)$, see eq. (13) and eqs (A20) and (A21) in Appendix A.

From eq. (8), we iterate $\hat{\kappa}_j$ to the next $\hat{\kappa}_{j+1}$, using $\hat{\theta}_{j+1}$ from eq. (19) as our best estimate of θ

$$\hat{\kappa}_{j+1} = \left(\coth\hat{\kappa}_j - \frac{1}{N}\sum_{i=1}^{N}\left\{\cos\hat{\theta}_{j+1}\cos\theta_i + \sin\hat{\theta}_{j+1}\sin\theta_i\left[\frac{I_1(\hat{\kappa}_j\sin\hat{\theta}_{j+1}\sin\theta_i)}{I_0(\hat{\kappa}_j\sin\hat{\theta}_{j+1}\sin\theta_i)}\right]\right\}\right)^{-1}. \tag{20}$$

As for eq. (19), we use $R(x)$ to calculate the Bessel function ratio, and our approximations for the hyperbolic cotangent are shown in Appendix A, eqs (A23)–(A25).

Maximum values on the boundary ($\theta = 0^\circ$ or 180°) are found by iterating from an initial guess of eq. (18) using eq. (21).

$$\hat{\kappa}_{j+1} = \left(\coth\hat{\kappa}_j - \cos\theta\frac{1}{N}\sum_{i=1}^{N}\cos\theta_i\right)^{-1}. \tag{21}$$

As in eq. (20), we use the approximations for the hyperbolic cotangent shown in Appendix A.

Alternating between eqs (19) and (20), a stable pair $(\hat{\theta}, \hat{\kappa})$ is typically located at the maximum of the likelihood function, usually in fewer than 10 iterations. When the maximum is close to the boundary ($\theta \approx 0^\circ$ or 180°), convergence is much slower, and more than 100 iterations might be required. In such cases, the value of the log-likelihood function is often maximum at the boundary, and it is calculated using eq. (21); the result is compared to the value in the interior, and the highest value is chosen.

Estimates of the angular standard deviation, θ_{63}, are calculated using

$$\cos\theta_{63} = 1 + \frac{\ln(1 - 0.63(1 - \mathrm{e}^{-2\kappa}))}{\kappa}, \tag{22}$$

and we estimate the 95 per cent confidence limits of the mean inclination, α_{95}, using (Fisher 1953; Kono 1980, eq. 12):

$$\cos\alpha_{95} = 1 - \frac{N-1}{N(\kappa-1)+1}(20^{1/(N-1)} - 1). \tag{23}$$

5 OUTLINE OF OUR METHOD

Our method of locating the maximum includes the following five steps:

1. We select an initial value using eqs (17) and (18). Then we iterate using alternatively eqs (19) and (20) until the solution of successive iterations changes very little. Both inequalities 24 and 25 must be simultaneously satisfied to terminate the iteration process

$$\Delta\theta = |\hat{\theta}_{j+1} - \hat{\theta}_j| < 0.000\,001^\circ \tag{24}$$

and

$$\Delta\kappa/\kappa = \left|\frac{\hat{\kappa}_{j+1} - \hat{\kappa}_j}{\hat{\kappa}_{j+1}}\right| < 0.000\,001. \tag{25}$$

Once the iteration terminates at $(\hat{\theta}, \hat{\kappa})$, we calculate the value of the log-likelihood function $h_1 = h(\hat{\theta}, \hat{\kappa})$ by the method described in Appendix A. If the calculations do not converge (i.e. inequalities 24 and 25 are not satisfied) in 10 000 iterations, a convergence problem is identified. The extremely small changes defined by inequalities 24 and 25 are not important for most data sets, but may be significant for near vertical solutions.

2. It is possible that there is no local maximum in the interior of the parameter space, and that the likelihood function is maximum at one of its boundaries, that is: $\theta = 0^\circ$, $\theta = 180^\circ$ or $\kappa = 0$. For the first two boundaries, the initial guess of the precision parameter comes from eq. (18). Then we iterate through eq. (21) until the condition of inequality 25 is satisfied. Once the iteration terminates at $\hat{\kappa}$, we calculate the value of the log-likelihood function $h_2 = h(0^\circ, \hat{\kappa})$, and $h_3 = h(180^\circ, \hat{\kappa})$ by the method described in Appendix A. Finally, we set $\hat{\kappa} = 0$ (and $\hat{\theta} = 90^\circ$) and calculate the value of the log-likelihood function on this boundary by eq. (11), $h_4 = h(90^\circ, 0)$. Although the maximum is probably never on this boundary ($\kappa = 0$)—in our numerical comparison it was never the case—we calculate this value for completeness.

3. We compare the four values of the log-likelihood function from steps 1 and 2: h_1, h_2, h_3 and h_4. The solution $(\hat{\theta}, \hat{\kappa})$ that gives the highest value is chosen as our maximum likelihood solution.

4. To check the robustness of the chosen solution, we evaluate the log-likelihood function in a 16-point oval array around our solution, where $\hat{\theta} - 0.01^\circ \le \theta \le \hat{\theta} + 0.01^\circ$ and $\hat{\kappa} - 0.1$ per cent $\le \kappa \le \hat{\kappa} + 0.1$ per cent (Fig. 5). If the chosen solution does not have the highest value compared to the 16 surrounding points, a robustness flag is raised.

5. Finally, we calculate estimates of the angular standard deviation, θ_{63}, and 95 per cent confidence limits of the mean inclination, α_{95}, using eqs (22) and (23).

6 CONVERGENCE AND ROBUSTNESS OF OUR METHOD

To assess the robustness of our method, we generated 368 000 random Fisher-distributed data sets. The data sets had 23 true inclinations, $I =: -90^\circ, -85^\circ, -80^\circ, -75^\circ, -70^\circ, -60^\circ, -50^\circ, -40^\circ, -30^\circ, -20^\circ, -10^\circ, 0^\circ, 10^\circ, 20^\circ, 30^\circ, 40^\circ, 50^\circ, 60^\circ, 70^\circ, 75^\circ, 80^\circ, 85^\circ$ and 90°. Precision parameters, κ, were assigned values of 10, 20, 40 and 100. The number of samples in each data set, N, was 5,

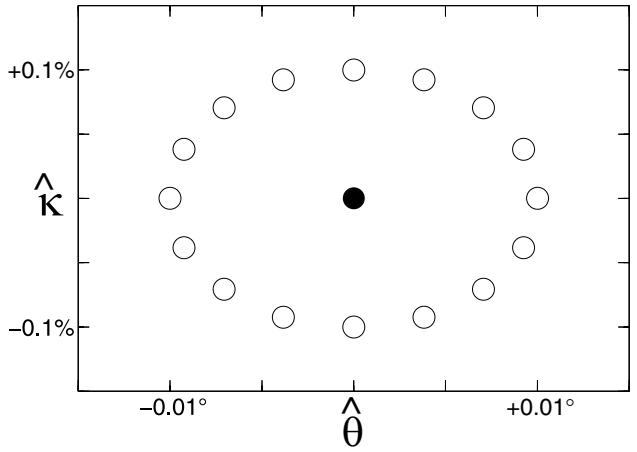

Figure 5. The scheme of verifying the robustness of our maximum likelihood method. For our maximum likelihood estimate $(\hat{\theta}, \hat{\kappa})$ (filled circle), we evaluated the log-likelihood function at 16 surrounding points (open circles) to verify that our estimate is at least a local maximum at this level of accuracy ($\hat{\theta} \pm 0.01^\circ$, $\hat{\kappa} \pm 0.1$ per cent).

10, 20 and 100. For each of these $23 \times 4 \times 4 = 368$ combinations, we generated 1000 random Fisher-distributed data sets. We use a method very similar to that of Fisher *et al.* (1987, p. 59). Our method is described in detail by Arason (1991).

For each of the 368 000 inclination-only data sets, we calculated the maximum likelihood solution pair $(\hat{\theta}, \hat{\kappa})$ by the method outlined in the previous section. We encountered 857 cases of convergence problems or robustness failure. This compares to 28 203 and 52 843 cases of a no-solution by the Kono and McFadden–Reid methods, respectively. Convergence problems in step 1 of our method occurred in fewer than 0.2 per cent of the data sets. In all these cases, the solution was approaching the vertical; usually $\hat{\theta} < 2^\circ$ (or $>178^\circ$), the convergence was very slow; and the conditions of inequalities 24 and 25 were not satisfied. In the majority of these cases, the solution still fulfilled our robustness check. One can employ a more sophisticated algorithm to handle these exceptional cases, but for practical reasons we chose to keep our algorithm simple.

For steep and dispersed inclinations, the maximum value of the log-likelihood function is frequently on the boundary of the parameter space. For data sets where the true inclination was $\pm 70^\circ$, $\pm 80^\circ$ and $\pm 90^\circ$, this happened in 5 , 17 and 28 per cent of the cases, respectively.

For each solution, we evaluated the log-likelihood function in an array around the solution pair shown in Fig. 5. Although this check does not guarantee that the iteration process has converged to the maximum, it should identify cases where the process went astray. Compared with the surrounding array, our solution was at the maximum value of the log-likelihood function in over 99.95 per cent of the 368 000 data sets. This check indicates that the solution is probably accurate to $\theta \pm 0.01^\circ$ and $\kappa \pm 0.1$ per cent.

Therefore, we are confident that, at least for $\kappa > 10$ and $N > 5$, our method of locating the maximum likelihood solution is extremely robust.

7 UNIQUENESS OF NEAR VERTICAL SOLUTIONS

For dispersed and steep data, the likelihood function sometimes has its maximum value on the boundary of the parameter space, that is a vertical solution ($I = \pm 90^\circ$).

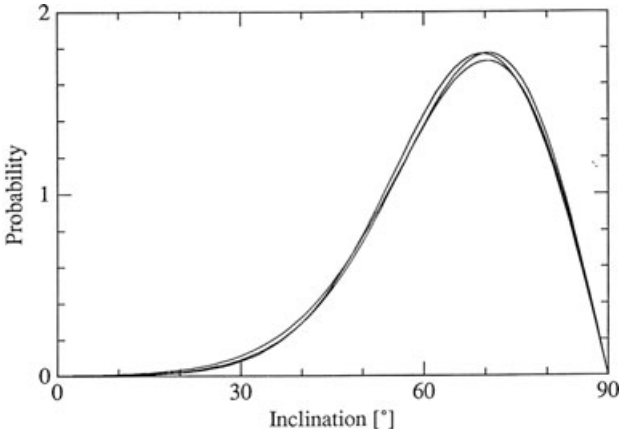

Figure 6. The distribution from eq. (2) of observed inclinations for three combinations of the true values (I_{true}, κ). The values are ($I_{\text{true}} = 75^\circ$, $\kappa = 8.6$), ($I_{\text{true}} = 80^\circ$, $\kappa = 10$), and ($I_{\text{true}} = 85^\circ$, $\kappa = 12$) (it is not important in this context to identify the curves). For these steep true inclinations and low κ it becomes impossible to extract information on both I_{true} and κ from a finite set of observed inclinations, and any attempt to do so will depend critically on the assumptions of the calculation method. From Arason (1991, fig. 5.12, p. 266).

The probability distribution for these cases is such that similar distributions may originate from a range of pairs of true inclinations and precision parameters. Fig. 6 shows an example of this effect, where three distributions of different pairs of (θ, κ) are almost identical. For such steep and dispersed true inclinations it is impossible to extract unique information on both the true inclination and precision parameter from a finite set of observed inclinations, and any attempt to do so will depend critically on the assumptions of the calculation method.

Therefore, we conclude that a mean inclination estimate on the edge, that is $I = \pm 90^\circ$, indicates that a unique solution does not exist, and in the absence of declination data the information to separate inclination and precision parameter is permanently lost. However, such a solution indicates that the true inclination is probably steep and the estimate of the precision parameter is likely to be a lower limit to the true value. To establish limits on the likely range of the true inclination for such steep and dispersed cases, one can use the single-parameter marginal likelihood method of Enkin & Watson (1996) to estimate a confidence interval of the inclination.

8 NUMERICAL EXAMPLE

As sample data, we use the paleomagnetic results analysed in numerical examples by Fisher (1953) and Briden & Ward (1966), listed in Table 1. The paleomagnetic samples were obtained by Jan

Table 1. Nine specimens from an Icelandic lava flow[a].

Specimen number	Declination (°)	Inclination (°)
631	343.2	66.1
632	62.0	68.7
633	36.9	70.1
634	27.0	82.1
635	359.0	79.5
636	5.7	73.0
642A	50.4	69.3
643A	357.6	58.8
644	44.0	51.4

[a]From Fisher (1953, Table 1, p. 304).

Table 2. Different methods used to estimate directional data statistics.

Method	Mean declination (°)	Mean inclination (°)	Precision parameter	95 per cent confidence limits (°)
Arithmetic mean	22.87	68.78	36.42	7.48
Fisher (1953)	24.27	70.89	35.08	8.81
		Inclination-only		
Briden & Ward (1966)	–	72	33	–
Kono (1980)	–	71.99	31.64	9.29
McFadden & Reid (1982)				
Original	–	68.79	34.62	9.25
Modified	–	70.95	34.62	+7.08, −11.40[a]
Enkin & Watson (1996)				
Gaussian estimate	–	71.87	25.13	7.47
Marginal likelihood	–	70.88	(26.88)	+13.96, −6.99[a]
Direct maximum likelihood				
Mathematica	–	71.85	32.45	–
Arason and Levi method of this study	–	71.85	32.45	9.17

[a]Confidence intervals not symmetric about the mean.

Hospers from the 1947 to 1948 lava flow of the Hekla volcano in Iceland (64.0°N, 19.7°W). In Table 2, we analyse these data using several methods. To identify small differences between the methods, we show the results with an accuracy of 0.01°. First, we show the arithmetic mean of the inclinations, declinations and a 95 per cent confidence limits on the mean inclination assuming normal distribution. Then we present the Fisher statistics estimates (Fisher 1953).

Using only the inclinations in Table 1, we calculate the mean inclination shown in Table 2 using the following methods: (1) The graphical estimate of Briden & Ward (1966, p. 137); (2) the estimate from Kono (1980); (3) estimates based on the methods of McFadden & Reid (1982), as originally published and subsequently modified; (4) estimates by Enkin & Watson (1996), both the Gaussian estimates and the single-parameter marginal likelihood method; we use the corresponding Gaussian estimate of the precision parameter because the marginal likelihood method has no estimate of the precision; (5) we then show a direct evaluation by the comprehensive mathematical software package, *Mathematica*; (6) finally, we show the Arason–Levi estimate following the method described in this study. Details of the convergence to our solution are shown in Appendix B.

The bivariate estimate of 95 per cent confidence limits of the Fisher statistics is not directly comparable to the other univariate estimates. Note also that the original McFadden–Reid method gives essentially the same mean inclination as the arithmetic mean.

Fig. 7 shows results of various methods for this particular numerical example, demonstrating that the Arason–Levi method accurately evaluates the maximum likelihood estimates for these inclination-only data.

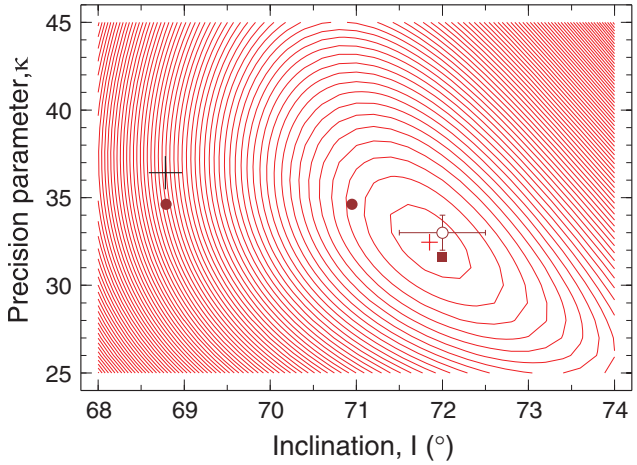

Figure 7. Contours of the log-likelihood function for the inclination data in Table 1, and results of various methods to identify its peak: black cross represents the arithmetic mean; open circle with error bars the Briden–Ward graphical method; filled circles McFadden–Reid, both their original method, close to the arithmetic mean, and the modified method; square the results of the Kono method; the Enkin–Watson estimate is not shown because they are not solving for this function; their estimate would be at the bottom of the graph; Finally, the red cross is obtained by the Arason-Levi method of this study, which we claim to represent an accurate estimate of the maximum likelihood.

9 COMPARISONS TO PREVIOUS METHODS

To assess the reliability and accuracy of the inclination-only methods, we used the 368 000 randomly generated Fisher-distributed data sets to estimate mean inclinations and precision parameters. For these data sets we calculated the Fisher mean, also using the declinations. For inclination-only data, the mean was calculated using the following methods: arithmetic mean; Kono (1980); McFadden & Reid (1982), both their original and modified methods; Enkin & Watson (1996); finally, we obtained maximum likelihood estimates by our robust technique.

Especially for cases of steep and dispersed data, the estimates provided by the previous methods are often displaced from the true peak of the likelihood function to systematically shallower inclinations. Moreover, mean inclination estimates by the original McFadden–Reid statistics, still used by some paleomagnetists, are nearly identical to the arithmetic mean; this method should be abandoned.

In Figs 8–10, we compare the distributions of 1000 estimates for one particular combination of the 368 cases, having true inclination, $I = 70°$, precision parameter, $\kappa = 20$, and sample number, $N = 100$. In scatter plots of Fig. 8, we show the distribution of the (I, κ) estimates for seven methods. The true mean ($I = 70°$, $\kappa = 20$) is shown as a large cross, and each of the 1000 estimates is marked

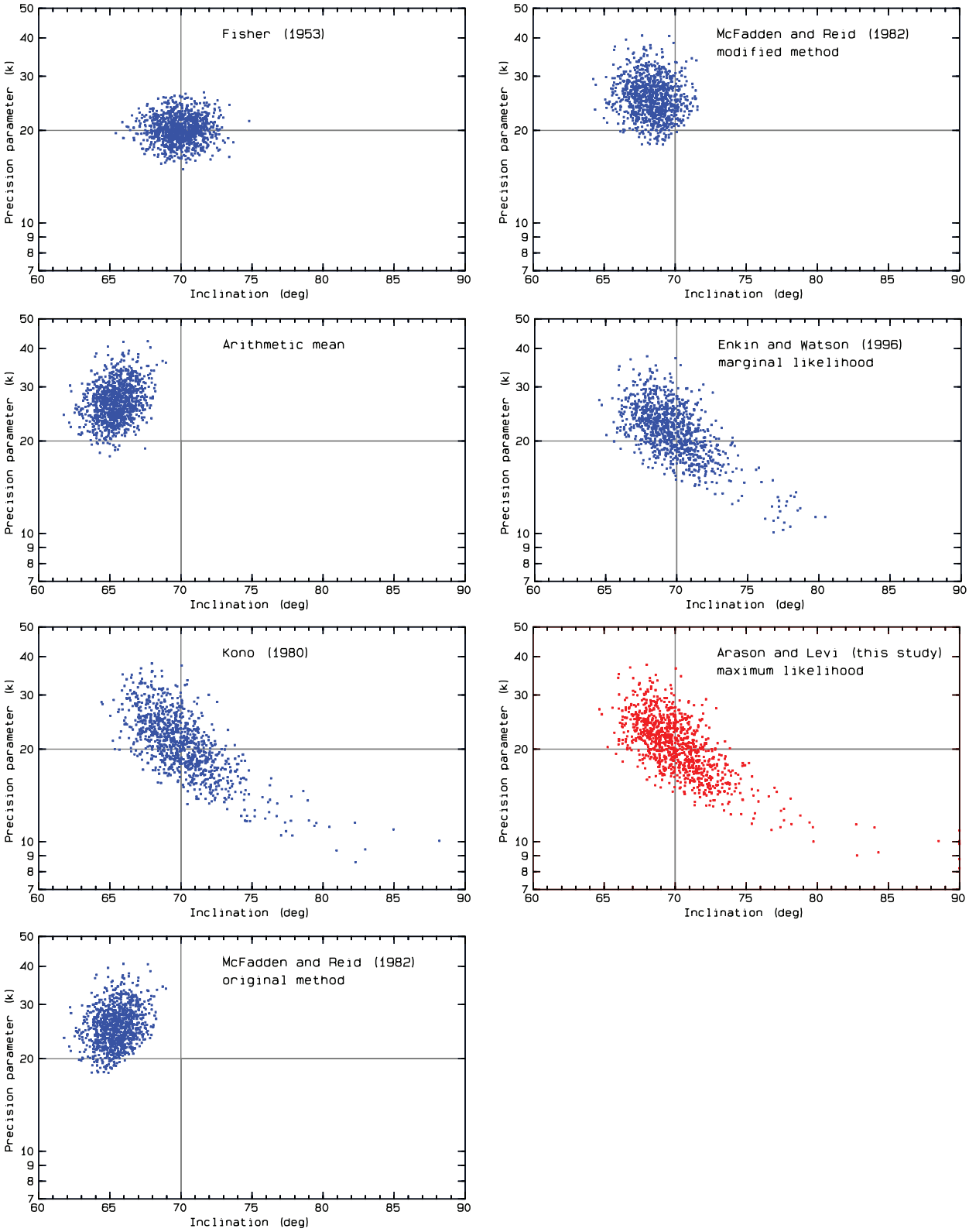

Figure 8. The distribution of mean inclination and precision parameter estimates for 1000 random Fisher-distributed data sets where $I = 70^{\circ}$, $\kappa = 20$, $N = 100$.

as a single dot. Displacement of the cloud of dots from the cross could indicate that something is wrong. The same data are shown in Figs 9 and 10 as histograms of the estimated mean inclinations and precision parameters.

Figs 8–10 show that the results of the Fisher-statistics (with declinations) are clustered symmetrically about the true mean. This is expected when both inclinations and declinations are known, and the Fisher-analysis should result in unbiased means. This indicates that

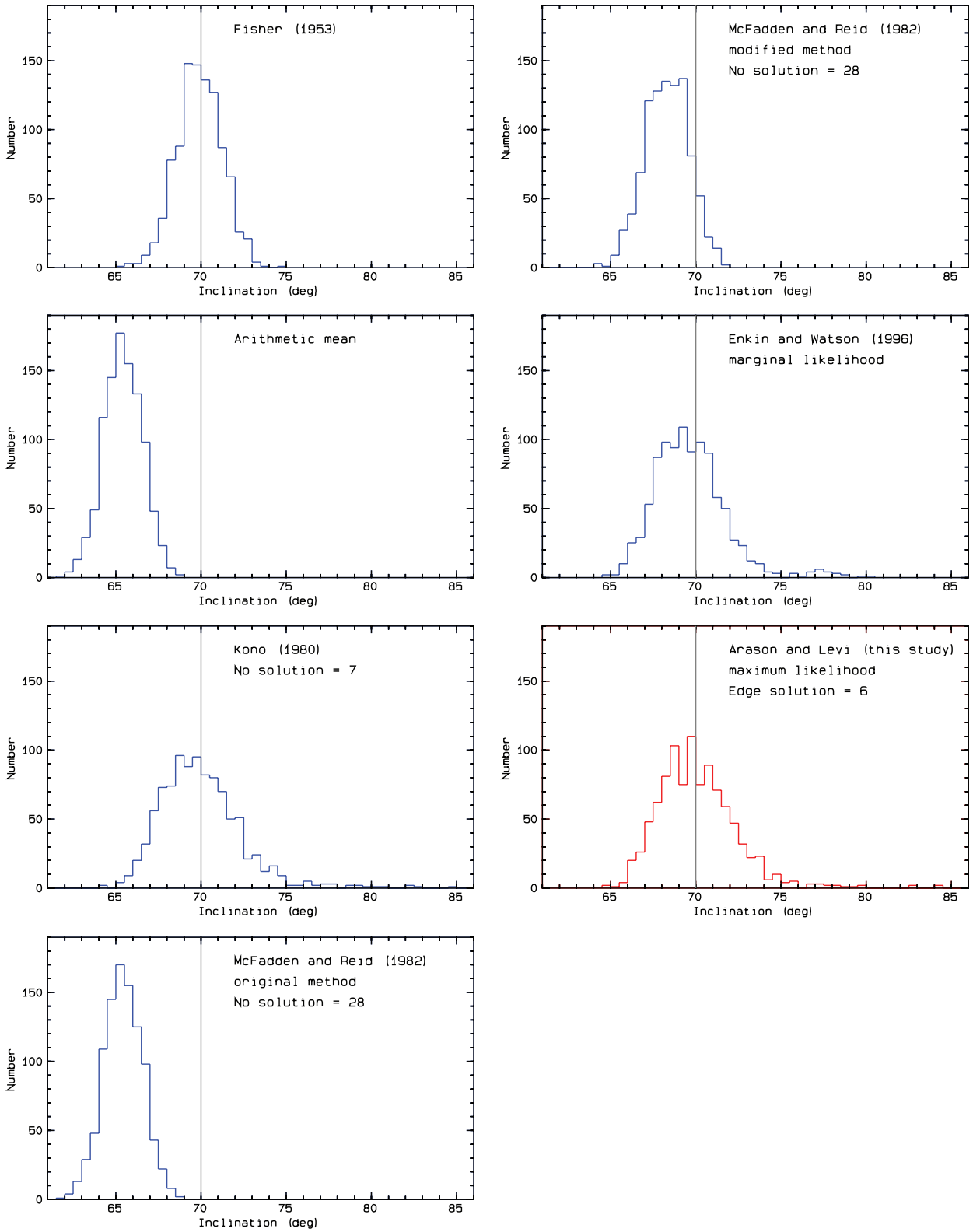

Figure 9. The distribution of 1000 mean inclination estimates for $I = 70^\circ$, $\kappa = 20$, $N = 100$.

the process of generating the Fisher-distributed random data sets is correct. The inclination means of the 1000 Fisher distributions, using the various methods, are shown in Table 3, which lists the arithmetic means for the 1000 inclination estimates and the geometric means of the precision parameter for one particular combination $I = 70^\circ$, $\kappa = 20$ and $N = 100$. The arithmetic mean of the data shifts the solutions to shallow inclinations and higher precision parameters. The Kono method gives solutions centred about the true value, but there is an apparent leakage towards the vertical ($I = 90^\circ$) and lower precision parameters; also, the Kono method has no-solution

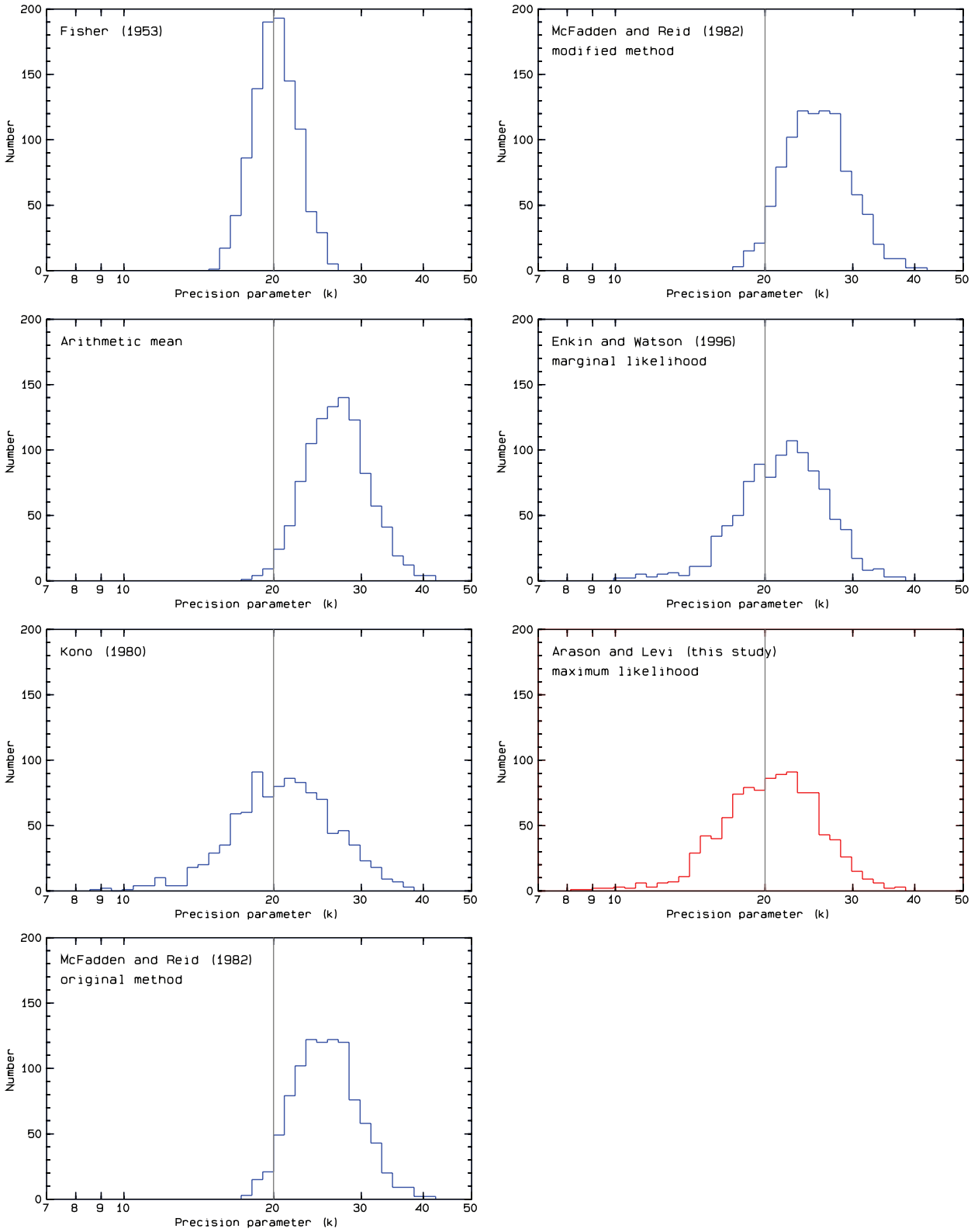

Figure 10. The distribution of 1000 precision parameter estimates (logarithmic scale) for $I = 70^\circ$, $\kappa = 20$, $N = 100$.

for seven cases. The McFadden–Reid method has no-solution for 28 data sets. The original McFadden–Reid method gives similar results as the arithmetic mean, whereas the modified McFadden–Reid method is noticably better; the mean inclination is less shallow, and the precision parameter remains the same. The Enkin–Watson method requires using the marginal likelihood method for all these data sets, but we use the corresponding Gaussian estimates for the precision parameter. The Arason–Levi method of this study has a distribution similar to the Kono and Enkin–Watson results; and for the six excluded data sets the maximum likelihood is vertical.

Table 3. Mean of the 1000 estimates shown in Figs 8–10 for the combination ($I = 70^\circ$, $\kappa = 20$, $N = 100$).

	Mean inclination, I (°)	Precision parameter
Fisher (1953)	69.9	20.2
Arithmetic mean	65.4	26.8
Kono (1980)	70.0	20.9
McFadden & Reid (1982)		
Original	65.4	25.5
Modified	68.3	25.5
Enkin & Watson (1996)		
Marginal likelihood	69.7	(21.6)[a]
Arason & Levi method of this study	70.1	20.5

[a]See text on estimate of precision parameter.

Fig. 11 shows average inclination shallowing versus true inclination for the various values of the precision parameter, κ. Each panel has four curves representing precision parameters, κ, of 10, 20, 40 and 100 (least shallowing). Each point was calculated as the average of 1000 estimates of the mean inclination by the methods for random Fisher-distributed data sets where the sample number was fixed at $N = 100$. The arithmetic mean and the McFadden–Reid methods produce the most shallowing. The Arason–Levi method causes the least shallowing of the mean inclination, and for $\kappa > 40$ there is no significant mean-shallowing for inclinations up to 80°. Results of the McFadden–Reid method(s) seem to be similar to the arithmetic mean. For the McFadden–Reid and Kono estimates 'no-solution' results were not included in the mean calculations. It may be somewhat surprising to see an apparent bias in the Fisher-estimates at 90°. Although the Fisher-estimates are themselves unbiased, arithmetic mean of one thousand Fisher mean-inclinations for $I = 90^\circ$ is slightly biased (shallow) due to the arithmetic mean bias. In Fig. 12, we show this effect, where we compare average inclination estimates of 1000 data sets, where $\kappa = 20$ and $N = 100$. The red data represent Fisher mean-inclinations of the 1000 data sets using both mean inclinations and declinations, whereas the blue data show arithmetic means of the 1000 Fisher mean-inclinations. The shallowing at 90° is about 1.6°.

Fig. 13 shows a comparison of the curves of the inclination-only methods in Fig. 11 for $\kappa = 20$ and $N = 100$. The largest inclination bias arises from the original McFadden–Reid method, which is identical to the arithmetic mean for this range. The methods of Enkin–Watson, Kono and Arason–Levi have only a very small inclination bias, and the Arason–Levi method produces the least bias. For comparison with the other methods, the solutions at the edges ($I = \pm 90^\circ$) by our method were not included in the means.

We compared the means of I and κ of the inclination-only methods for the 368 combinations of I, κ and N. For true inclinations in the range ±75° to ±90°, our method gives results closest to the true inclination in 79 per cent of the 128 combinations, whereas Enkin–Watson is closest in 16 per cent and Kono in 5 per cent. For true inclinations ±60° or ±70°, our method was closest to the true inclination in 70 per cent of the 64 combinations, Kono 20 per cent, Enkin–Watson 6 per cent and the modified McFadden–Reid in 5 per cent of the combinations. For lower inclinations, there is less difference between the methods. For the mean estimates of the precision parameters for the 368 combinations, the closest to the true value is given by the Enkin–Watson method in 77 per cent of cases, whereas our method is closest in 23 per cent of the combinations.

The magnitude of the inclination bias of the various methods as a function of the true inclination is summarized in Table 4, where we combined positive and negative inclinations. For each value of true inclination there are 16 combinations of κ and N. For true inclinations of ±60°, ±50° and ±40° there are 96 (I, κ, N)-combinations. We calculated the arithmetic mean of the 1000 estimates of the mean inclinations. We then found the absolute difference between the true inclination and these 96 averages. Table 4 shows that for true inclinations of ±40° to ±60°, 48 of the 96 averages, were within about 0.1–0.2° of the true mean for all the inclination-only methods except the arithmetic mean and the original McFadden–Reid method, where the shallowing bias was close to 1.3°. These values should be compared to the apparent bias for Fisher-statistics, which includes a mixture of random statistical fluctuations and a component of the arithmetic mean bias for the 1000 averages. For true inclinations up to 75°, there is almost no inclination bias for the Arason–Levi method of this study. Even for steeper inclinations, our method may give reliable results. We also tried grouping the data sets by $\theta\sqrt{\kappa}$ as suggested by Enkin & Watson (1996). For the three groups $\theta\sqrt{\kappa} > 400$, $200 < \theta\sqrt{\kappa} < 400$ (or $150 < \theta\sqrt{\kappa} < 400$ when $N > 30$) and $\theta\sqrt{\kappa} < 200$ (or $\theta\sqrt{\kappa} < 150$ when $N > 30$), the Arason–Levi method gives lower inclination bias than the Enkin–Watson procedures. For these combinations, the Arason–Levi method has the lowest inclination bias of all the inclination-only methods.

The relative difficulty in applying the methods of Kono, Arason–Levi or the Enkin–Watson Gaussian estimates is comparable; they all require a fairly simple iteration process through some equations. However, the numerical approach of the Enkin–Watson single-parameter marginal likelihood method is more complicated and computer intensive. The Arason–Levi method gives much better results than the Enkin–Watson Gaussian estimates, and slightly better results (Fig. 13) than the Kono method and the much more complicated Enkin–Watson single-parameter marginal likelihood method.

The new Arason–Levi method accurately calculates the maximum likelihood estimates of mean inclinations from inclination-only data. Comparisons of the results with other methods is very favourable to our new maximum likelihood method. On average, our method gives estimates of mean inclinations, least biased to shallow values.

10 DISCUSSION

This study shows that the McFadden–Reid method, which is still widely used, often leads to significant inclination shallowing. The original method has an apparent error and the resulting mean inclinations are almost identical to the arithmetic mean. A modification to the method is only somewhat better. In our opinion the McFadden–Reid method should be abandoned.

For steep and dispersed inclinations, the Kono and McFadden–Reid methods often do not have a solution. This problem seems more pronounced for the McFadden–Reid method. Some of these cases result in a vertical solution by our method.

Enkin & Watson (1996) weighted the likelihood function with a Bayesian factor. This weighting gives better constraints of the solution for very dispersed and steep inclinations. By using Bayesian statistics they use the *a priori* knowledge that shallow inclinations should be more common than steep ones. This higher weight to low inclinations will result in shallower mean inclinations than by unweighted maximum likelihood estimation. Furthermore, they presented three different approaches depending on the dispersion and steepness of the data: arithmetic mean, Gaussian estimate and

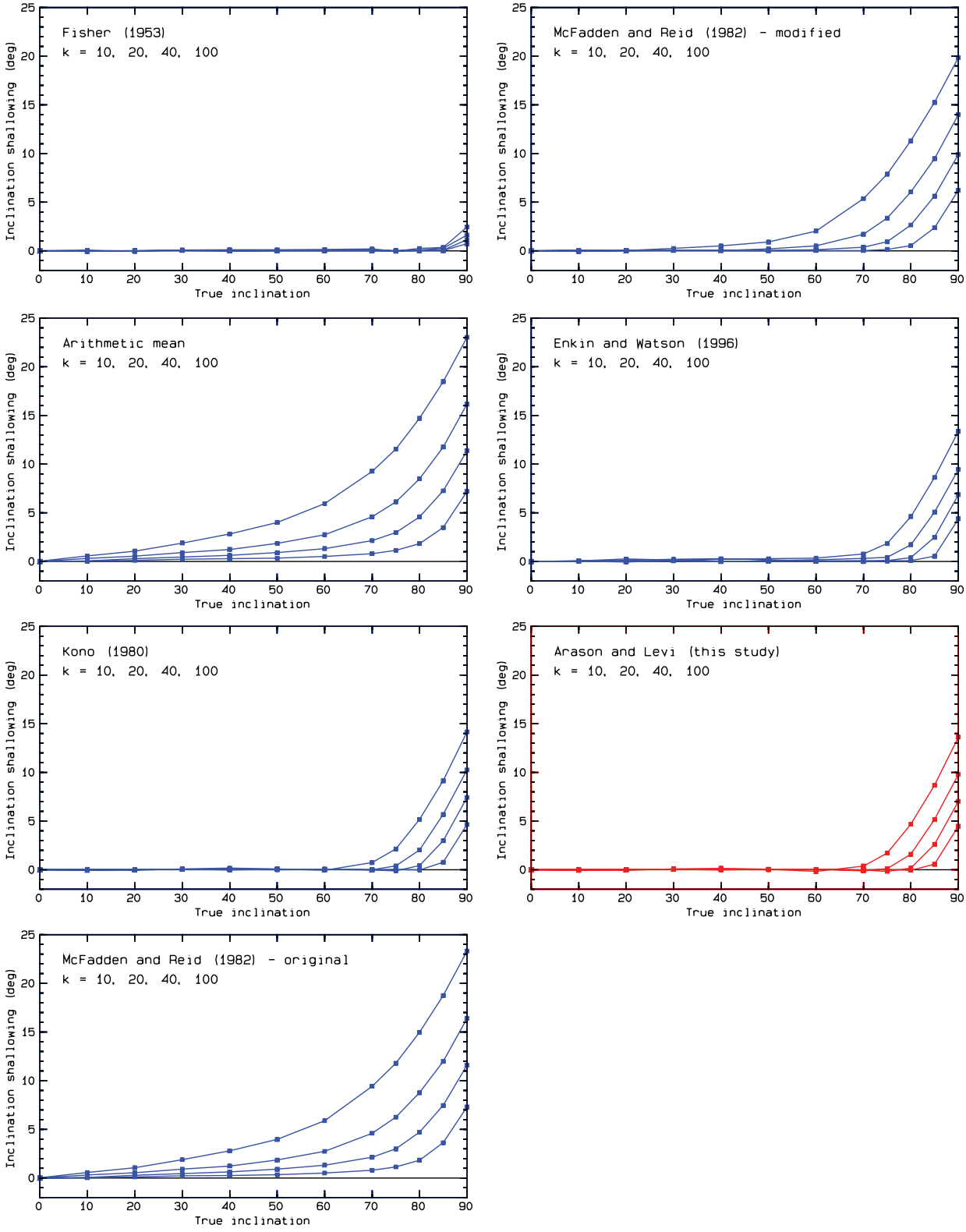

Figure 11. The comparison of the inclination-only methods. Each point represents a mean value of 1000 estimates. The ideal method should result in no inclination shallowing. Identical results are for negative inclinations.

marginal likelihood. Results near their empirically chosen boundaries can yield different mean inclinations, depending on the approach chosen. Enkin & Watson (1996) use too few iterations in their numerical examples to locate the best solution, which leads to less accurate results. The Bayesian view is successful in better constraining the solution and, especially, in estimating the precision parameter. Our calculations show that their estimates of precision parameters are often closer to the true value than estimates by the

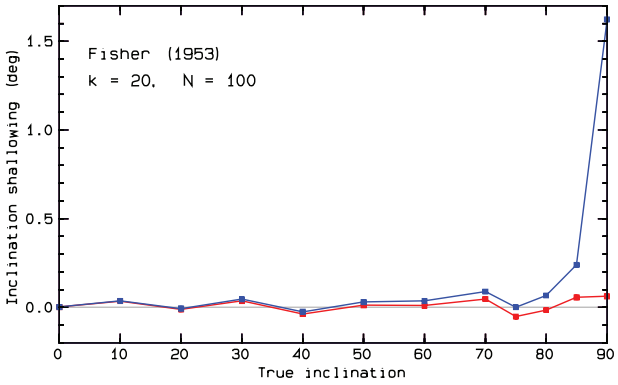

Figure 12. Comparison of Fisher and arithmetic mean inclinations. Each point represents the mean of 1000 inclination-mean estimates. See text.

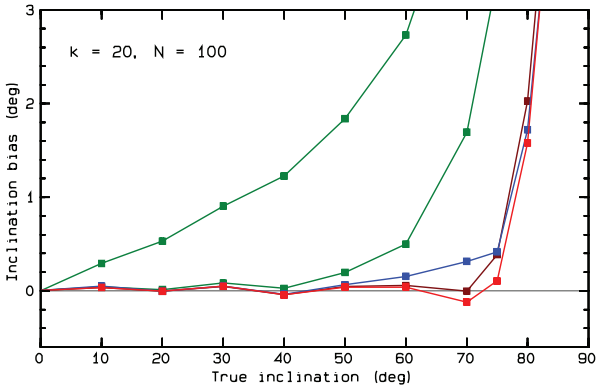

Figure 13. Comparison of the curves in Fig. 11, concentrating on the variations up to 3° inclination shallowing bias. The two green curves show the estimates of the original and modified McFadden–Reid methods. For this range, the original McFadden–Reid method (higher green curve) gives identical results to the arithmetic mean. The blue curve represents the Enkin–Watson estimates and the brown curve the Kono estimates. The red curve showing least bias for steep inclinations is due to the Arason–Levi method presented in this study.

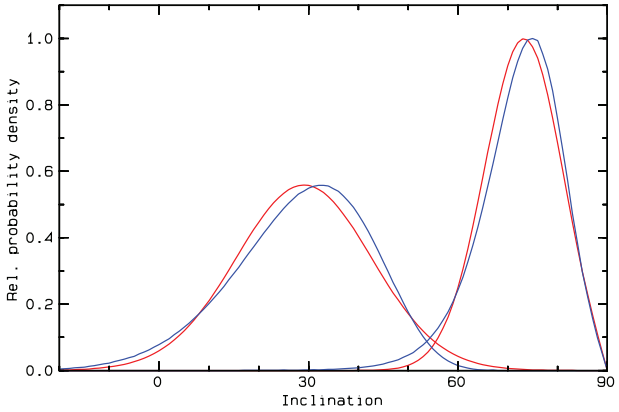

Figure 14. Comparison of the distribution of inclinations for Fisher-distributed VGPs (blue curves for I_{GAD} of 30° and 75°) and Fisher-distributed directions (red curves from eq. 2). The area under all the curves is the same.

Arason–Levi method. However, we think it is important to use the same self-consistent method for all data sets, rather than using *ad hoc* conditions to select between different methods. Also, we consider that it is more important to obtain unbiased estimates of mean inclinations than unbiased estimates of the precision parameter.

All methods for analysing inclination-only data assume that the directions are Fisher-distributed. Through the dipole equation and spherical trigonometry, one can transform paleomagnetic directions to virtual geomagnetic poles (VGPs) and vice versa. If either of the distributions (poles or directions) is isotropic the other becomes oval (e.g. Arason & Levi 1997). Studies have shown that VGPs from lavas, which record the instant geomagnetic field, are more isotropic than the observed directions. This may not be true for data from sediments, where there is more within specimen averaging of the secular variation.

We are interested in comparing the distributions of inclination-only data for Fisher-distributed VGPs and Fisher-distributed directions. In Fig. 14, the blue curves show the probability density of the inclinations for Fisher-distributed VGPs and the red curves show the same for Fisher-distributed directions, where the

Table 4. Comparisons of the inclination-bias of the methods for various true inclinations.

	True inclination				
	0–30°	40–60°	70–75°	80–85°	90°
Number of (I, κ, N)-combinations	112	96	64	64	32
Fisher (1953)	0.07	0.09	0.18	0.50	3.3
	Inclination-only				
Arason and Levi method of this study	0.07	0.09	0.29	4.1	9.6
Kono (1980)	0.07	0.09	0.60	4.6	10.2
Enkin & Watson (1996)	0.11	0.19	0.95	4.6	9.9
Gaussian estimate	0.08	0.12	0.63	4.4	10.1
Marginal likelihood	0.07	0.19	0.95	4.6	9.9
McFadden & Reid (1982)					
Modified method	0.08	0.16	1.8	6.4	11.7
Original method	0.3	1.3	3.7	7.9	13.0
Arithmetic mean	0.3	1.3	3.5	7.5	12.6

Shallowing bias for the various methods in different ranges of the inclination. The data show median absolute inclination difference, in degrees, between the true inclination and averages of 1000 estimates for several (I, κ, N)-combinations. Ranges of the true inclination include both positive and negative values.

mean inclinations are about 30° and 75°. The distribution of inclinations assuming Fisher-distributed directions was obtained using eq. (2), whereas the inclination distribution assuming Fisher-distributed VGPs was numerically determined by transforming the Fisher-distribution through spherical trigonometry and the dipole equation. The Fisher-distributed VGPs (blue curves) were assumed to have angular standard deviation, $\mathrm{ASD_{VGP}} = 12.1^\circ$ ($\kappa_{\mathrm{VGP}} = 44.8$), observed at latitude 16.1°N ($I_{\mathrm{GAD}} = 30.0^\circ$), and $\mathrm{ASD_{VGP}} = 18.5^\circ$ ($\kappa_{\mathrm{VGP}} = 19.1$), seen at latitude 61.8°N ($I_{\mathrm{GAD}} = 75.0^\circ$). The chosen ASD values are based on a latitude dependent model of Harrison (2009). For comparison, we show distributions of inclinations (red curves) assuming Fisher-distributed directions using eq. (2), with $I = 30^\circ$, $\kappa_{\mathrm{I}} = 17.2$ and $I = 75^\circ$, $\kappa_{\mathrm{I}} = 48.8$. The values for the precision parameters were chosen so that the height of the peak of the area normalized red curves were identical to the corresponding blue curve. Fig. 14 shows that the distributions of inclination-only data are very similar, independent of whether the poles or the directions are Fisher-distributed. The peak is slightly displaced; this bias of the mean was quantified by Arason & Levi (1997). Furthermore, there are small but systematic differences in the shape of the curves. The blue curves are more skewed, with relatively scarce very high values and values slightly lower than the mean, and relatively abundant very low values and values slightly higher than the mean. However, for ordinary sample sizes of inclination-only data, one can probably not distinguish whether the VGPs or directions are better represented by a Fisher distribution. Because the distributions are so very similar, the assumption of Fisher-distributed directions in all the inclination-only methods is justifiable.

In paleomagnetism, there are many possible sources of directional errors; for example, compaction-induced inclination shallowing in sediments, non-vertical coring, tectonic tilting and sample orientation. Such errors may be comparable to or greater than the correction to the arithmetic mean by the method of this study. However, one should always strive to use the best methods available to compensate for potential errors.

11 CONCLUSIONS

It can be difficult to evaluate the terms comprising the log-likelihood function and its derivatives with respect to θ and κ. Previous studies have made various approximations to the problem, leading to inaccurate estimates, which are systematically biased to shallow inclinations (e.g. Arason 1991; Arason & Levi 1995; Levi & Arason 2006).

We succeeded in analytically cancelling the exponential elements from the likelihood function, and we can now calculate its value anywhere in the parameter space, for any inclination-only data set, and with full accuracy. Furthermore, we can now calculate the partial derivatives of the likelihood function, also with desired accuracy, and locate the maximum likelihood without the assumptions required by previous methods. We have developed a simple method that accurately evaluates the maximum likelihood estimates of the mean inclination and precision parameter for inclination-only data.

The information to separate the mean inclination from the precision parameter will be lost for very steep and dispersed data sets. Although the likelihood function always has a maximum value, for some dispersed and steep data sets with few samples, the likelihood function has its highest value on the boundary of the parameter space, that is $I = \pm 90^\circ$, but having a relatively well-defined dispersion. Our simulations indicate that this occurs quite frequently for certain data sets, and that relatively small perturbations in the data will drive the maximum to the boundary. We interpret this to indicate that, for such data sets, the information needed to separate the mean inclination and the precision parameter is permanently lost.

The new Arason–Levi method for determining unbiased mean inclinations from inclination-only data is based on maximum likelihood estimation. The Arason–Levi method should be preferred because it is more accurate than the previous methods, particularly for steep inclinations. It is self-consistent for all data sets, employing a simple iteration procedure. It is very robust, and, of the available methods, it gives the least biased estimates of mean inclinations. Computer program codes for our maximum likelihood method for inclination-only data, and a convenient web-calculator for the occasional user, can be found at: http://www.vedur.is/∼arason/paleomag/

ACKNOWLEDGMENT

The authors thank L. Kristjánsson, C.G.A. Harrison, R.J. Enkin and C.G. Langereis for reviews and helpful comments on the paper.

REFERENCES

Arason, P., 1991. Paleomagnetic inclination shallowing in deep-sea sediments, *PhD thesis*. Oregon State University, Corvallis, Oregon, U.S.A., 363 pp.

Arason, P. & Levi, S., 1995. *Comparison of Statistical Methods in the Analysis of Inclinations-Only Paleomagnetic Data,* International Union of Geodesy and Geophysics, XXI General Assembly, Boulder, Colorado, USA, 2–14 July 1995, Abstracts Week A, A182–A183 (http://www.vedur.is/∼arason/paleomag/iugg95.pdf).

Arason, P. & Levi, S., 1997. Intrinsic bias in averaging paleomagnetic data, *J. Geomag. Geoelectr.,* **49,** 721–726.

Briden, J.C. & Ward, M.A., 1966. Analysis of magnetic inclinations in borecores, *Pure appl. Geophys.,* **63,** 133–152.

Clark, R.M., 1983. Estimation of parameters in the marginal Fisher distribution, *Austr. J. Stat.,* **25,** 227–237.

Clark, R.M., 1988. An evaluation by simulation of alternative estimators for the marginal Fisher distribution, *J. appl. Stat.,* **15,** 235–246.

Cox, A. & Gordon, R.G., 1984. Paleolatitudes determined from paleomagnetic data from vertical cores, *Rev. Geophys.,* **22,** 47–72.

Enkin, R.J. & Watson, G.S., 1996. Statistical analysis of palaeomagnetic inclination data, *Geophys. J. Int.,* **126,** 495–504.

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

Fisher, N.I., Lewis, T. & Embleton, B.J.J., 1987. *Statistical Analysis of Spherical Data,* 329 pp, Cambridge Univ. Press, London.

Gradshteyn, I.S. & Ryzhik, I.M., 1980. *Table of Integrals, Series, and Products,* Academic Press, Orlando, 1160 pp.

Harrison, C.G.A., 1974. The paleomagnetic record from deep-sea sediment cores, *Earth Sci. Rev.,* **10,** 1–36.

Harrison, C.G.A., 2009. Latitudinal signature of Earth's magnetic field variation over the last 5 million years, *Geochem. Geophys. Geosyst.,* **10,** Q02012, doi:10.1029/2008GC002298.

Kono, M., 1980. Statistics of paleomagnetic inclination data, *J. geophys. Res.,* **85,** 3878–3882.

Langevin, P., 1905. Magnétisme et théorie des électrons (in French), *Ann. Chim. Phys.,* **5,** 70–127.

Levi, S. & Arason, P., 2006. *Comparisons of Inclination-only Statistical Methods,* American Geophysical Union, Fall Meeting, San Francisco, California, USA, 11–15 December 2006, Eos Trans. AGU, **87**(52), Fall Meet. Suppl., Abstract GP21B–1313 (http://www.vedur.is/∼arason/paleomag/agu2006_comp.pdf).

McElhinny, M.W. & McFadden, P.L., 2000. *Paleomagnetism: Continents and Oceans,* Academic Press, New York, 386 pp.

McFadden, P.L. & Reid, A.B., 1982. Analysis of paleomagnetic inclination data, *Geophys. J. R. astr. Soc.,* **69,** 307–319.

Olver, F.W.J., 1972. Chapter 9. Bessel functions of integer order, in *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,* ed. Abramowitz, M. & Stegun, I.A., Dover Publications, Mineola, 1047 pp. (http://www.math.sfu.ca/∼cbm/aands/toc.htm)

Peirce, J.W., 1976. Assessing the reliability of DSDP paleolatitudes, *J. geophys. Res.,* **81,** 4173–4187.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T., 1989. *Numerical Recipes, The Art of Scientific Computing (Fortran version),* Cambridge University Press, Cambridge, 702 pp.

Westphal, M., Gurevitch, E.L. & Pozzi, J.P., 1998. On the estimation of the true mean inclinations when declinations are unknown, *Geophys. Res. Lett.,* **25,** 1569–1572.

APPENDIX A: EVALUATION OF THE FUNCTIONS

To solve the maximum likelihood problem for inclination-only data, one needs to be able to accurately calculate several functions for any combination of κ and θ: $0 \leq \kappa$ and $0^\circ \leq \theta \leq 180^\circ$. This can be done with a combination of analytical cancellations and a careful choice of approximations of the functions for certain regions of the parameters.

A.1 The hyperbolic Bessel functions

We need to evaluate the hyperbolic Bessel functions of zero and first order, $I_0(x)$ and $I_1(x)$, but it is sufficient to determine the ratios $I_0(x)/\mathrm{e}^x$ and $I_1(x)/I_0(x)$, and these ratios can be calculated more accurately than the Bessel functions themselves.

The hyperbolic Bessel functions are sometimes called the modified Bessel function of first kind, and they can not be expressed as a finite combination of elementary functions. At zero, the functions have the values, $I_0(0) = 1$ and $I_1(0) = 0$, and as x increases the Bessel functions increase exponentially (Figs A1 and A2). The Bessel function $I_0(x)$ can be expanded as an infinite series (e.g. Gradshteyn & Ryzhik 1980, eq. 8.447.1, p. 961). Here we are only interested in positive values of x; we note the relations

$$I_0(-x) = I_0(x) \quad \text{and} \quad I_1(-x) = -I_1(x). \tag{A1}$$

$I_1(x)$ is related to the derivative of $I_0(x)$ by

$$I_1(x) = I_0'(x). \tag{A2}$$

To compute the Bessel function $I_0(x)$ and the ratio $I_1(x)/I_0(x)$, we use the approximations of Press *et al.* (1989), which are based on Olver (1972, eqs 9.8.1–4).

We define the functions P, Q, U and V as

$$P(x) = p_0 + p_1 t^2 + p_2 t^4 + p_3 t^6 + p_4 t^8 + p_5 t^{10} + p_6 t^{12}, \tag{A3}$$

$$Q(x) = q_0 + q_1 t^{-1} + q_2 t^{-2} + q_3 t^{-3} + q_4 t^{-4} + q_5 t^{-5} + q_6 t^{-6} + q_7 t^{-7} + q_8 t^{-8}, \tag{A4}$$

$$U(x) = x\left(u_0 + u_1 t^2 + u_2 t^4 + u_3 t^6 + u_4 t^8 + u_5 t^{10} + u_6 t^{12}\right), \tag{A5}$$

$$V(x) = v_0 + v_1 t^{-1} + v_2 t^{-2} + v_3 t^{-3} + v_4 t^{-4} + v_5 t^{-5} + v_6 t^{-6} + v_7 t^{-7} + v_8 t^{-8}, \tag{A6}$$

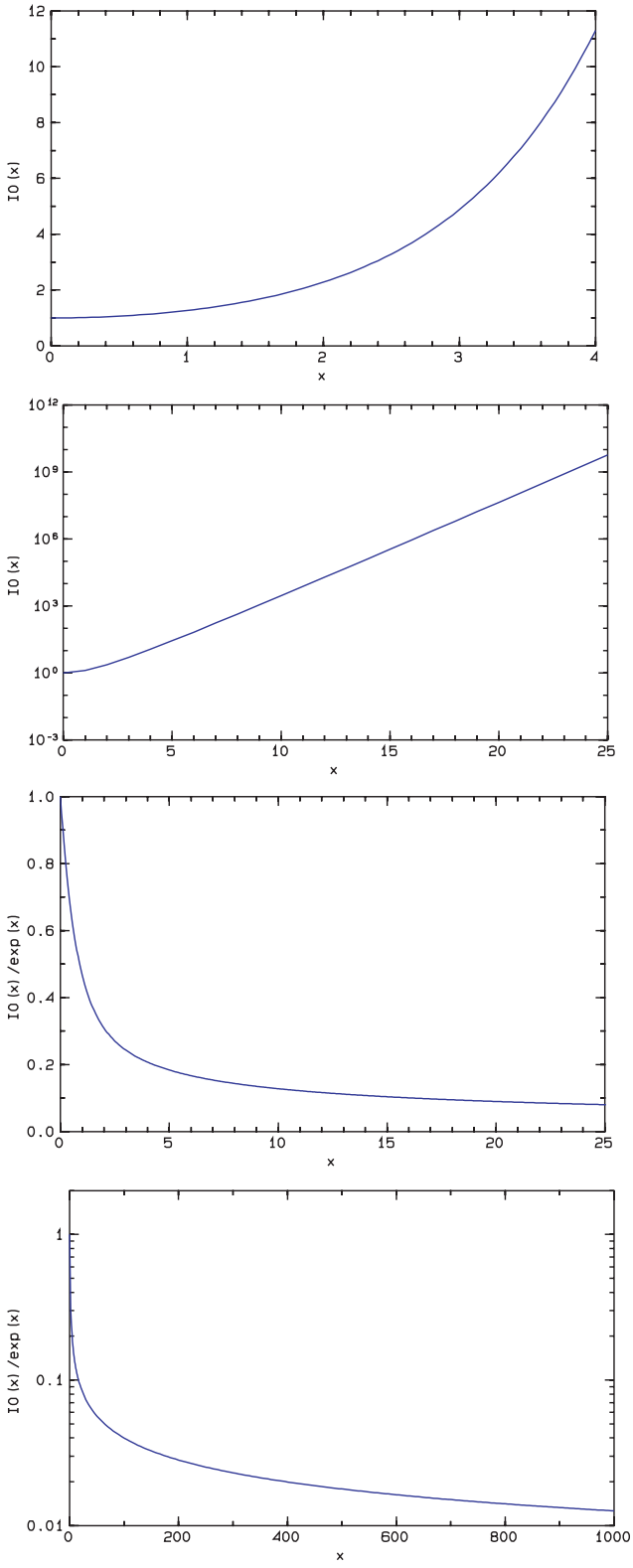

Figure A1. The hyperbolic Bessel function $I_0(x)$ increases exponentially and becomes numerically difficult to evaluate for high values of x. However, the ratio $I_0(x)/\mathrm{e}^x$ is easier to handle.

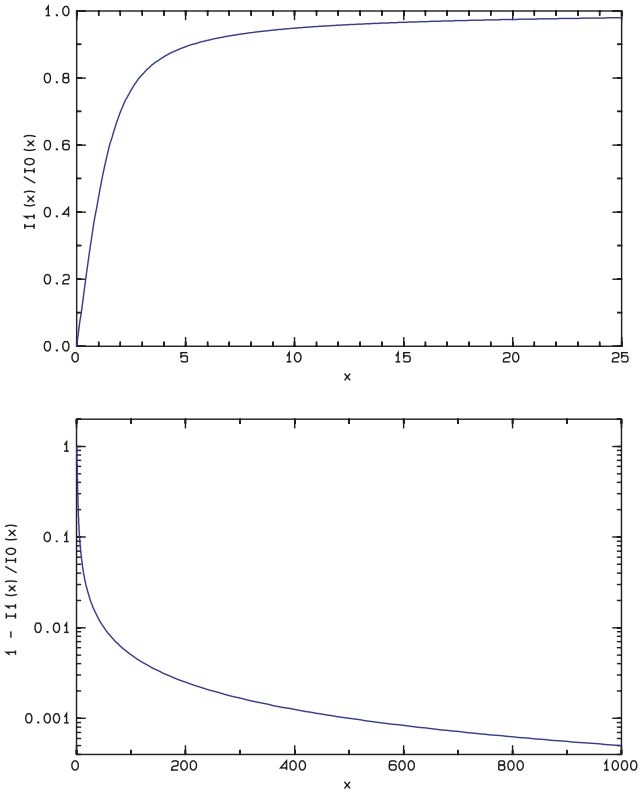

Figure A2. The hyperbolic Bessel function $I_1(x)$ increases exponentially, similar to $I_0(x)$, and becomes numerically difficult to evaluate for high values of x. However, the ratio $I_1(x)/I_0(x)$ is easier to handle.

where

$$t = x/3.75, \tag{A7}$$

$$p_i = [1, 3.5156229, 3.0899424, 1.2067492, 0.2659732, 0.0360768, 0.0045813], \tag{A8}$$

$$q_i = [0.39894228, 0.01328592, 0.00225319, -0.00157565, 0.00916281, -0.02057706, 0.02635537, -0.01647633, 0.00392377], \tag{A9}$$

$$u_i = [0.5, 0.87890594, 0.51498869, 0.15084934, 0.02658733, 0.00301532, 0.00032411], \tag{A10}$$

$$v_i = [0.39894228, -0.03988024, -0.00362018, 0.00163801, -0.01031555, 0.02282967, -0.02895312, 0.01787654, -0.00420059]. \tag{A11}$$

We evaluate the Bessel functions for the range $0 \leq x < 3.75$:

$$I_0(x) = P(x) + \varepsilon, \tag{A12}$$

where the error term is $|\varepsilon| < 1.6 \times 10^{-7}$ and

$$I_1(x) = U(x) + \varepsilon, \tag{A13}$$

where the error term is $|\varepsilon| < 0.3 \times 10^{-7}$.

The Bessel ratio for the range $0 \leq x < 3.75$ is calculated by

$$I_1(x)/I_0(x) \approx U(x)/P(x). \tag{A14}$$

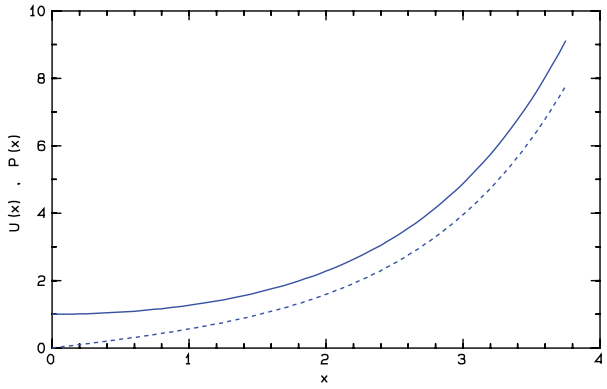

Figure A3. The functions $P(x)$ (solid) and $U(x)$ (dashed) for the range $0 \leq x < 3.75$. The functions $P(x)$ and $U(x)$ were used to evaluate the hyperbolic Bessel functions $I_0(x) = P(x)$ and $I_1(x) = U(x)$ in this range.

Fig. A3 shows the functions $P(x)$ and $U(x)$ [i.e. $I_0(x)$ and $I_1(x)$], for low values of x, $0 < x < 3.75$. Both the functions $P(x)$ and $U(x)$ and their ratio are easily calculated for any value of x in this range using eqs (A3), (A5), (A7), (A8), (A10) and (A14).

We evaluate the Bessel functions for the range $x \geq 3.75$:

$$I_0(x) = \left(\frac{e^x}{\sqrt{x}}\right)(Q(x) + \varepsilon), \tag{A15}$$

where the error term is $|\varepsilon| < 1.9 \times 10^{-7}$ and

$$I_1(x) = \left(\frac{e^x}{\sqrt{x}}\right)(V(x) + \varepsilon), \tag{A16}$$

where the error term is $|\varepsilon| < 2.2 \times 10^{-7}$.

The Bessel ratio for the range $x \geq 3.75$ is calculated by

$$I_1(x)/I_0(x) \approx V(x)/Q(x). \tag{A17}$$

Fig. A4 shows the functions $Q(x)$ and $V(x)$, for higher values of x, $3.75 < x < 25$. Both the functions $Q(x)$ and $V(x)$ and their ratio are easily calculated for any value of x in this range using eqs (A4), (A6), (A7), (A9), (A11) and (A17). In calculating the ratio $I_1(x)/I_0(x)$ by eq. (A17), the square root and the exponential elements of eqs (A15) and (A16) are cancelled. The functions $Q(x)$ and $V(x)$ and their ratio are easily calculated for any value of $x \geq 3.75$.

For convenience, we define the Bessel-exponent-free function $B(x)$

$$\frac{I_0(x)}{e^x} \approx B(x) = \frac{P(x)}{e^x} \quad \text{for } 0 \leq x < 3.75, \tag{A18}$$

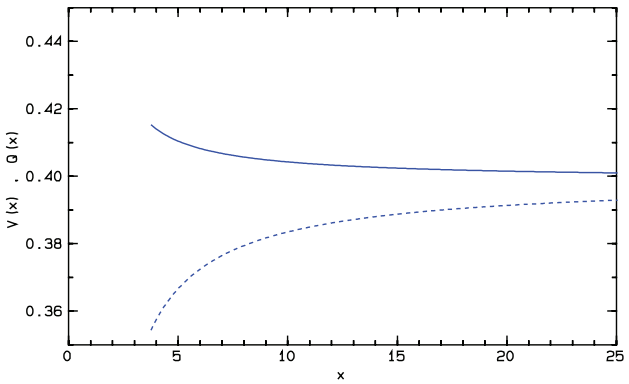

Figure A4. The functions $Q(x)$ (solid) and $V(x)$ (dashed) for $x \geq 3.75$. Both functions approach $(2\pi)^{-\frac{1}{2}} \approx 0.3989$ asymptotically. These were used to evaluate the hyperbolic Bessel functions.

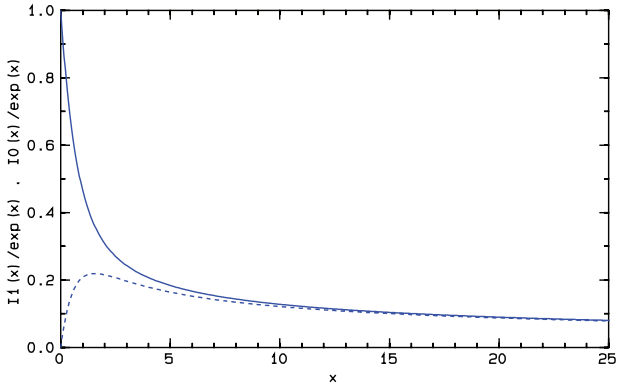

Figure A5. The ratios $I_0(x)/e^x$ (solid) and $I_1(x)/e^x$ (dashed).

$$\frac{I_0(x)}{e^x} \approx B(x) = \frac{Q(x)}{\sqrt{x}} \quad \text{for } 3.75 \le x, \tag{A19}$$

and the Bessel-ratio function $R(x)$

$$\frac{I_1(x)}{I_0(x)} \approx R(x) = \frac{U(x)}{P(x)} \quad \text{for } 0 \le x < 3.75, \tag{A20}$$

$$\frac{I_1(x)}{I_0(x)} \approx R(x) = \frac{V(x)}{Q(x)} \quad \text{for } 3.75 \le x. \tag{A21}$$

Fig. A5 shows the ratios $I_0(x)/e^x$ along with $I_1(x)/e^x$, for x in the range $0 < x < 25$. Fig. 4 shows the functions $B(x)$ and $R(x)$ (i.e. $I_0(x)/e^x$ and $I_1(x)/I_0(x)$). By defining the functions $B(x)$ and $R(x)$ we have eliminated the exponential elements of the Bessel functions and are able to calculate, with full accuracy, the essential part of the Bessel functions and their ratio for any value of x. For $x \ge 0$, these functions take values in the range $0 < B(x) \le 1$ and $0 \le R(x) < 1$. In evaluating the Bessel functions and their ratios with the functions $B(x)$ and $R(x)$, we avoid the exponential elements to keep full accuracy for all values of x and therefore avoid overflow problems in our calculations. This same exponential element of the Bessel function can also be analytically eliminated from the log-likelihood function.

A.2 The hyperbolic cotangent

The derivative of the log-likelihood function includes a hyperbolic cotangent term, coth κ. The hyperbolic cotangent can be written as

$$\coth x = \frac{e^x + e^{-x}}{e^x - e^{-x}}. \tag{A22}$$

Fig. A6 shows the hyperbolic cotangent function. Because of overflow problems, numerical evaluation by eq. (A22) is problematic for high and low values of x, so we evaluate the hyperbolic cotangent by

$$\coth x \approx \frac{1}{x} + \frac{1}{3}x - \frac{1}{45}x^3 + \frac{2}{945}x^5 \quad \text{for } 0 < x < 0.01, \tag{A23}$$

$$\coth x = \frac{1 + e^{-2x}}{1 - e^{-2x}} \quad \text{for } 0.01 \le x \le 15, \tag{A24}$$

$$\coth x \approx 1 \quad \text{for } 15 < x. \tag{A25}$$

The error in using the approximation of eq. (A23) in the range $0.0001 \le x \le 0.01$ is less than 10^{-13}, and in using eq. (A25), less than 2×10^{-13}.

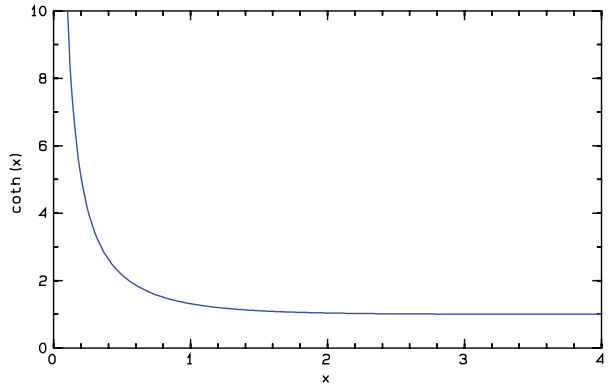

Figure A6. The hyperbolic cotangent function, coth(x). Close to $x = 0$ it behaves like $1/x$ and for $x > 3$ it approaches 1.

A.3 Evaluation of the log-likelihood function

We evaluate the log-likelihood function $h(\theta, \kappa)$ in three parts

$$h(\theta, \kappa) = A_1 + A_2 + A_3, \tag{A26}$$

where

$$A_1 = N \ln\left[\frac{\kappa}{2 \sinh \kappa}\right], \tag{A27}$$

$$A_2 = \sum_{i=1}^{N} (\kappa \cos\theta \cos\theta_i + \ln[I_0(\kappa \sin\theta \sin\theta_i)]), \tag{A28}$$

$$A_3 = \sum_{i=1}^{N} \ln[\sin\theta_i]. \tag{A29}$$

A.4 The first part of the log-likelihood function

To evaluate A_1 of the log-likelihood function $h(\theta, \kappa)$ for any value $0 \le \kappa$, we express the hyperbolic sine function as

$$\sinh \kappa = (e^{\kappa} - e^{-\kappa})/2. \tag{A30}$$

The hyperbolic sine function and its exponential behaviour is shown in Fig. A7. Because of its exponential growth and potential overflow problems, it is difficult to accurately evaluate this function for high values of κ. Therefore, it is necessary to isolate the exponential element of the function.

For values of κ in the range, say 1–10, eqs (A27) and (A30) can be easily evaluated with a calculator or a computer program. However, when κ approaches zero or high values, there may be difficulties evaluating the exponential elements (or ln 0). For high values of κ (e.g. $\kappa > 100$), ordinary calculators and programming languages will overflow or give very inaccurate values for sinh κ and e^{κ}. However, it is possible to combine the elements in eq. (A27) to analytically cancel the problematic terms, so the function A_1 can be accurately calculated for any value of κ

$$A_1 = N \ln\left[\frac{\kappa}{2 \times (e^{\kappa} - e^{-\kappa})/2}\right], \tag{A31}$$

$$A_1 = N[\ln(\kappa) - \ln(1 - e^{-2\kappa}) - \kappa]. \tag{A32}$$

Eq. (A32) can be used to accurately calculate the function A_1, at least in the range $0.01 \le \kappa \le 15$. For values of κ close to zero, we

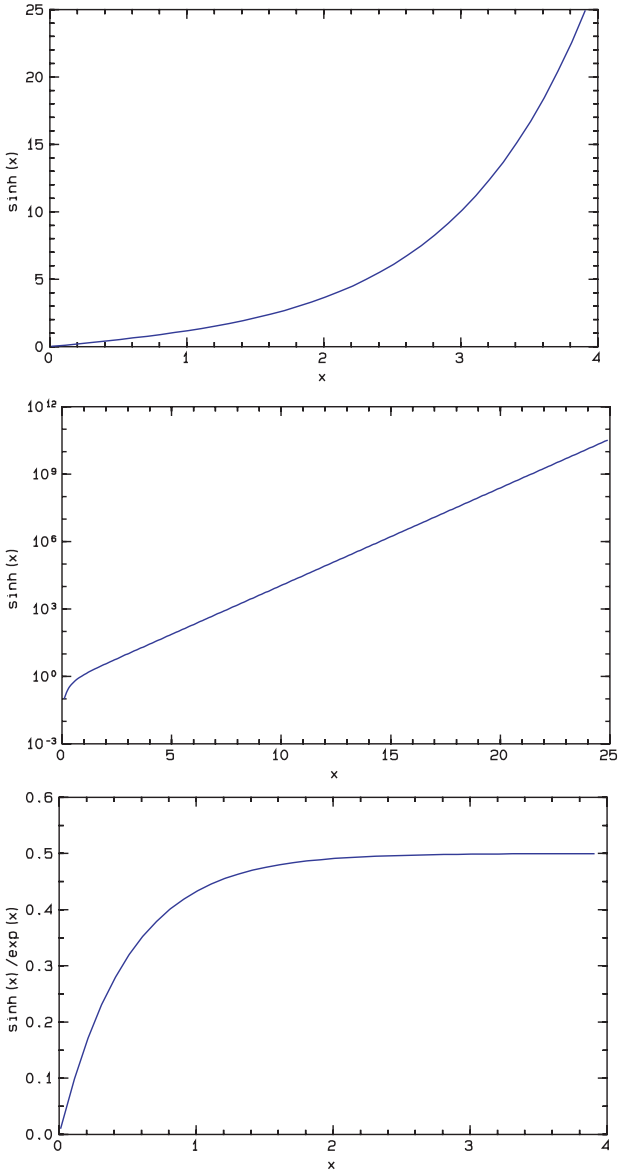

Figure A7. The hyperbolic sine function, sinh(x), grows exponentially and becomes difficult to accurately evaluate for high values of x. However, the function sinh(x)/e^x is numerically very simple.

expand the second term as a power series resulting in

$$\ln(1-\mathrm{e}^{-2\kappa}) \approx \ln(\kappa) + \ln(2) + \ln\left(1-\kappa+\frac{2}{3}\kappa^2-\frac{1}{3}\kappa^3+\frac{2}{15}\kappa^4-\frac{8}{45}\kappa^5\right). \tag{A33}$$

The first terms of eqs (A32) and (A33) can be analytically cancelled for κ approaching zero. The function A_1 can be calculated with an error of less than 10^{-10} in the range $0 \le \kappa < 0.01$ by

$$A_1 \approx -N\ln(2) - N\ln\left(1-\kappa+\frac{2}{3}\kappa^2-\frac{1}{3}\kappa^3+\frac{2}{15}\kappa^4-\frac{8}{45}\kappa^5\right) - N\kappa. \tag{A34}$$

This representation allows us to calculate the function at zero: $A_1(0) = -N\ln(2) \approx -0.6931\,N$.

For values of $\kappa > 15$, the second term in eq. (A32) becomes negligible (with an error of less than 10^{-10}) and the function can be calculated by

$$A_1 \approx N\ln(\kappa) - N\kappa. \tag{A35}$$

In summary, to retain full accuracy, we calculate the first part of the log-likelihood function by

$$A_1 \approx N\left[-\ln(2) - \ln\left(1-\kappa+\frac{2}{3}\kappa^2-\frac{1}{3}\kappa^3+\frac{2}{15}\kappa^4-\frac{8}{45}\kappa^5\right) - \kappa\right], \tag{A36}$$

for $0 \le \kappa < 0.01$

$$A_1 = N[\ln(\kappa) - \ln(1-\mathrm{e}^{-2\kappa}) - \kappa] \quad \text{for } 0.01 \le \kappa \le 15, \tag{A37}$$

$$A_1 \approx N[\ln(\kappa) - \kappa] \quad \text{for } 15 < \kappa. \tag{A38}$$

The essential elements of the function A_1 are shown in Fig. A8, that is the function $x + A_1/N$. The functions $\ln(\kappa)$ and $\ln(\sinh\kappa)$ cancel at $\kappa = 0$, and the function can be calculated with full accuracy for all $0 \le \kappa$.

A.5 The second term of the log-likelihood function

To compute the second term, A_2, of the log-likelihood function $h(\theta, \kappa)$ for any value $0 \le \kappa$ and $0^\circ \le \theta \le 180^\circ$ and any given data set of θ_i, $i = 1, \ldots, N$, we use the function $B(x)$ in eqs (A18) or (A19), dependent on the value $x = \kappa\,\sin\theta\,\sin\theta_i$:

$$A_2 = \sum_{i=1}^{N}(\kappa\cos\theta\cos\theta_i + \ln[\exp(\kappa\sin\theta\sin\theta_i)B(\kappa\sin\theta\sin\theta_i)]), \tag{A39}$$

$$= \sum_{i=1}^{N}(\kappa\cos\theta\cos\theta_i + \kappa\sin\theta\sin\theta_i + \ln[B(\kappa\sin\theta\sin\theta_i)]), \tag{A40}$$

which can be simplified using the cosine difference formula

$$\cos(x-y) = \cos x\,\cos y + \sin x\,\sin y \tag{A41}$$

to yield

$$A_2 = \sum_{i=1}^{N}(\kappa\cos(\theta_i-\theta) + \ln[B(\kappa\sin\theta\sin\theta_i)]). \tag{A42}$$

Eq. (A42) is easily evaluated with full accuracy for all values of κ, θ and θ_i, using the function $B(x)$ in eq. (A18) or (A19). We note that $A_2(\theta, \kappa = 0) = 0$ and the log-likelihood function can be evaluated and becomes independent of θ at the boundary $\kappa = 0$.

A.6 The third term of the log-likelihood function

The third term, A_3, of the log-likelihood function is constant for any given data set and should not affect the location of the maximum of the likelihood function. There is a problem if any of the observed values is exactly vertical, that is $\theta_i = 0^\circ$ or 180° (ln 0). Otherwise A_3 can be directly calculated.

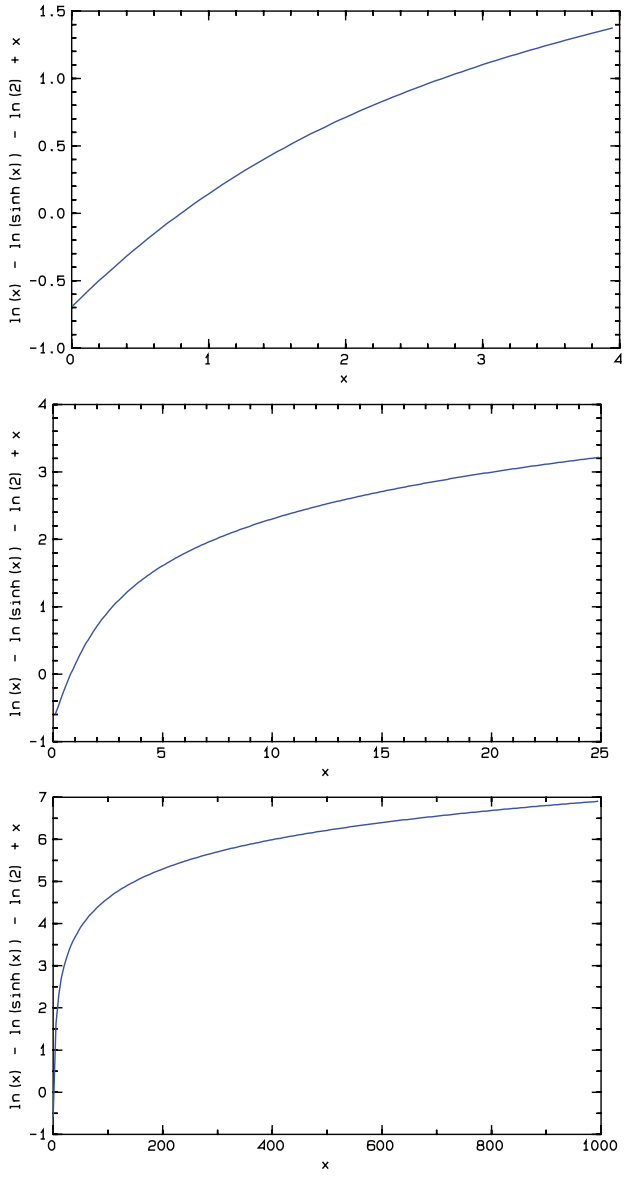

Figure A8. Essential elements of the function $A_1(x)$. Note that $\ln(x)$ and $\ln(\sinh x)$ cancel each other at $x = 0$.

APPENDIX B: NUMERICAL ITERATION PROCESS

The following list shows the convergence of the iteration process for the data in Table 1. We start with the initial guess of the arithmetic mean $I = 68.78^\circ$ ($\theta = 21.22^\circ$), $\kappa = 36.42$. For this pair, the value of the log-likelihood function is $h = 3.83$.

Iter	θ	κ	Log-likelihood
0	21.22	36.42	3.83
1	18.93	35.59	4.12
2	18.57	34.75	4.129
3	18.44	34.13	4.132
4	18.36	33.68	4.1336
5	18.31	33.36	4.1344
10	18.189	32.66	4.13533
20	18.153	32.467	4.1353855
30	18.15129	32.4554	4.1353857157
40	18.151171	32.45473	4.135385716284
44	18.151166	32.45471	4.1353857162853

After 44 iterations the difference between successive steps was less than our limit ($\Delta\theta < 0.000\ 001^\circ$ and $\Delta\kappa/\kappa < 0.000\ 001$). The solution indicates that the maximum likelihood estimate of the mean inclination is $I = 71.85^\circ$ ($\theta = 18.15^\circ$), which is 3.07° steeper than the arithmetic mean.

Now we find the maximum value on the three edges. First, we find a solution for $I = +90^\circ$, then $I = -90^\circ$, and finally calculate the value of the log-likelihood function at $\kappa = 0$.

Iter	θ	κ	Log-likelihood
5	0	12.651665	3.76
5	180	0.520573	−21.04
0	90	0	−16.32

The value of the log-likelihood function on the three edges is considerably lower than for our solution in the interior.

The highest value of the log-likelihood function is from the first iteration process, it satisfies our convergence and robustness checks, and our maximum likelihood solution is $I = 71.85^\circ$, $\kappa = 32.45$.