Geophys. J. Int. (2015) **200,** 1046–1051

doi: 10.1093/gji/ggu443

GJI Geomagnetism, rock magnetism and palaeomagnetism

Robust estimators of palaeosecular variation

Neil Suttie,[1] Andrew Biggin[1] and Richard Holme[2]

[1] *Geomagnetism Laboratory, Department of Earth, Ocean and Ecological Sciences, University of Liverpool, Liverpool L69 3BX, United Kingdom. E-mail: N.Suttie@liverpool.ac.uk*

[2] *Department of Earth, Ocean and Ecological Sciences, University of Liverpool, Jane Herdman Laboratories,* 4 *Brownlow, St Liverpool L69 3GP, United Kingdom*

Accepted 2014 November 11. Received 2014 November 11; in original form 2014 January 29

SUMMARY

The Fisher distribution is central to palaeomagnetism but presents several problems when used to characterize geomagnetic field directions as observed in sequences of volcanic rocks. First, it introduces a shallowing effect when used to define the mean of any group of directional unit vectors. This is problematic because it can suggest the presence of persistent non-axial dipole components when none are present. More importantly, it fails to capture the observed 'long tail' in distributions of both directions and associated virtual geomagnetic poles in terms of angular distance from a central direction. To achieve a good fit to data, it therefore requires the introduction of a second distribution (and therefore the estimation of additional parameters) or the arbitrary removal of data. Here we present a new distribution to describe palaeomagnetic directions and demonstrate that it overcomes both of these problems, generating robust indicators of both the central direction (or pole position) and the spread of palaeomagnetic data as defined by unit vectors. Starting from the assumption that poles (or directions) have an expected colatitude, rather than a mean location, we derive the spherical exponential distribution. We demonstrate that this new distribution provides a good fit to palaeomagnetic data sets from seven large igneous provinces between 15 and 65 Ma and also those produced by numerical dynamo models. We also use it to derive a new shape parameter which may be used as a diagnostic tool for testing goodness of fit of models to data and use this to argue for a shift in geomagnetic behaviour between 5 and 15 Ma. Furthermore, we point out that this new statistic can be used to determine the most appropriate distribution to be used when constructing confidence limits for poles.

Key words: Probability distributions; Palaeomagnetic secular variation; Palaeomagnetism applied to geological processes.

1 INTRODUCTION

The most accessible record of the ancient magnetic field is provided by the remanent magnetization of igneous rocks. The palaeomagnetic directions from a locality can be used to infer its palaeolatitude, under the hypothesis of a geocentric axial dipole (GAD), while the distribution of the directions reflects the geomagnetic secular variation over the period that the rocks cooled (e.g. Cox 1970). This palaeosecular variation is increasingly viewed as one of the most important tools for studying changes in the behaviour of the geodynamo (McFadden *et al.* 1991; Biggin *et al.* 2008a,b; Smirnov *et al.* 2011). To make inferences regarding either the mean direction or the dispersion of palaeomagnetic poles, the distribution proposed by Fisher (1953) is commonly used.

The hypothesis that the geomagnetic field, when averaged over palaeomagnetic timescales, is a GAD is generally accepted, at least as a first order approximation. A strictly valid statement of the GAD hypothesis is difficult because it would require a definite time period, long enough to average out secular variation, but free from reversals and excursions. The preferred form of the GAD hypothesis depends on the application. When deriving a palaeopole from a group of palaeomagnetic directions, it is convenient to assume that the average direction, obtained from unit vectors, is the direction of a GAD field, although it has been shown that averaging unit vectors may introduce bias to the mean direction (Creer 1983). The bias tends to give shallow directions which can masquerade as non-dipole field components, particularly a zonal octopole.

Field models can be constructed that reproduce realistic features of the secular variation of the field, subject to the constraint of a time-averaged GAD, by so-called giant-Gaussian processes (Constable & Parker 1988). In these models the GAD constraint is fulfilled by letting all the Gauss coefficients of the field, other than the axial dipole, vary randomly around zero. It is possible to derive a sampling distribution for palaeomagnetic vectors from such models

(Khokhlov *et al.* 2006), but generally the models are based on recent data, from the past 5 Myr, and so are only valid for that time period. Often, it is argued that the frequency distribution of virtual geomagnetic poles (VGPs) is more isotropic than that of palaeomagnetic directions and can be reasonably approximated by a Fisher distribution (e.g. McElhinny & Merrill 1975). Detailed analysis of VGPs from the past 5 Ma led Harrison (2009) to conclude that they could be well described by a mixture of a Fisher distribution and a uniform distribution. The extra parameter introduced (describing the relative amount of the two distributions) makes this very difficult to use in practical parameter estimation. In such problems, where for instance the question is to determine the position of the geographic pole, relative to the site, the Fisher distribution remains the sampling distribution of choice, largely because there exist explicit formulae for both the mean direction (or pole) and the corresponding confidence limits. Here we consider a simple alternative distribution, the spherical exponential distribution, which like the Fisher distribution has a single precision parameter in addition to the central direction.

2 MAXIMUM ENTROPY DISTRIBUTIONS

In problems of elementary parameter estimation, a sensible choice of sampling distribution is one that has maximum entropy subject to whatever constraints are specified, where the entropy of a distribution $P(x)$ is defined as $-\sum_x P(x)\log P(x)$. It can be shown that, for given constraints, such a distribution is by far the most likely to be observed as a frequency distribution (Jaynes 2003). Simple univariate distributions include the uniform distribution which maximizes entropy for a quantity that lies in a given range, or the Gaussian distribution that maximizes entropy for a quantity that has an expected first and second moment. In the case of unit vectors, the Fisher distribution has maximum entropy of all distributions with a given mean direction. Assigning a Fisher distribution to palaeomagnetic directions therefore amounts to a simple statement of belief: the directions have a specified mean direction, usually taken as the direction of a GAD field. Similarly, when applied to VGPs, the belief that the mean VGP is coincident with the geographic pole leads to the assignment of a Fisher distribution, although it has been noted that this poorly represents the observed frequency distribution of VGPs for the past 5 Myr (Harrison 2009). Generally, the frequency distributions of both VGPs and directions have longer tails than a Fisher distribution, with a lot of the mass far from the central direction. Sometimes, the observations are trimmed to fit a Fisher distribution by either removing low-latitude VGPs or by a scheme proposed by Vandamme (1994), which iteratively removes outlying data until the remaining distribution fits some criteria. Rather than shoehorn data to fit a Fisher distribution, it would be preferable if all the available data could be retained, and a more appropriate sampling distribution developed, yet it is desirable to retain a maximum entropy formulation with no additional adjustable parameters. Clearly, to achieve this, the formulation of the GAD constraint must be adjusted. One possibility would be to constrain the total vector sum of the magnetic vectors to lie in a specified direction, but as palaeomagnetism usually only provides unit vectors, such a constraint would involve an additional adjustable parameter.

It has become commonplace in studies of palaeosecular variation to define the angular dispersion of the field in terms of the root mean squared angle of deviation, by analogy with the normal distribution. There is general agreement that the field over a given period can be characterized by the angular spread of either VGPs about the pole or directions around the GAD direction, usually as a function of observation latitude. In the case of VGPs, an obvious choice of sampling distribution would therefore be one that maximizes entropy subject to the constraint that the VGPs have expected latitude. A similar distribution can be given for directions. Imagine the angle between palaeomagnetic directions ($\mathbf{v}_i$) and the GAD direction ($\mathbf{u}$) are θ_i, and the expectation of θ is S'

$$\langle\theta\rangle = S', \tag{1}$$

where we have used the prime to distinguish S' from the conventional definition of S as the rms angular deviation:

$$S = \sqrt{\frac{1}{N}\sum_i \theta_i^2}. \tag{2}$$

Maximizing entropy subject to (1) leads to the spherical exponential distribution with precision parameter k_E

$$P(\theta)\mathrm{d}\theta = \frac{k_E^2+1}{1+\mathrm{e}^{-k_E\pi}}\mathrm{e}^{-k_E\theta}\sin\theta\,\mathrm{d}\theta. \tag{3}$$

If the distribution is circularly symmetric, the distribution of the vector $\mathbf{v}$ is found by dividing eq. (3) by $2\pi\sin\theta$, which gives

$$P(\mathbf{v}) \cong \frac{k_E^2+1}{2\pi}\mathrm{e}^{-k_E\theta}, \tag{4}$$

where it has been assumed that $e^{-k_E} \ll 1$.

The distribution has a longer tail than the Fisher distribution. The expectation of θ is given by

$$\langle\theta\rangle \cong \frac{2k_E}{2k_E+1} \cong \frac{2}{k_E}, \tag{5}$$

whereas for a Fisher distribution with precision parameter k_F,

$$\langle\theta^2\rangle \cong \frac{2}{k_F}. \tag{6}$$

The estimate of central direction given by the likelihood function (4) has the character of a median-type estimator; it minimizes the total angular deviation, rather than the squared deviation and is less sensitive to data far from the centre of the distribution. A drawback in assigning the sampling distribution (4) is that unlike the Fisher distribution there are no explicit formulae for either the central direction or its confidence limits. To evaluate the likelihood, we can use Bayes theorem.

$$P(\mathbf{u}|\mathbf{v},k) = \frac{P(\mathbf{v}|\mathbf{u},k)P(\mathbf{u},k)}{P(\mathbf{v},k)}, \tag{7}$$

where the subscript on k has been dropped.

As the expected deviation and precision k are unknown, k must be eliminated by integration:

$$P(\mathbf{u}|\mathbf{v}) = \int_{k_{\min}}^{k_{\max}} P(\mathbf{u}|\mathbf{v},k)P(k)\mathrm{d}k. \tag{8}$$

The prior probability of k [$P(k)$] can be assigned by considering likely limits of palaeosecular variation. If we assign upper and lower limits to S' of 4° and 40°, we find $2 < k < 30$ to be a suitable range for the prior. Rather than assigning a uniform prior within this range, we assign $P(k) \propto \frac{1}{k}$.

This choice, often called a Jeffreys' prior, can be used when there are well-defined limits to the range of a scale parameter, so that the probability can be normalized. It was suggested by Jeffreys (1932) and was vigorously attacked by Fisher (1934). The motivation for using it is that it obviates the need to consider whether precision (k), mean deviation (S') or mean squared deviation (S^2), etc. is the parameter of interest, because any power of k has the same prior distribution (i.e. the logarithm of k is uniformly distributed). The explicit form of prior is then

$$P(k) = (k \log 15)^{-1}. \qquad (9)$$

The distribution of $\mathbf{u}$ can now be found numerically. Because the likelihood is such a sharply peaked function of the data, the choice of prior is of no practical importance and the results would not be noticeably changed if a uniform prior was assigned. In fact, all parameter estimation involves adopting some prior distribution for the parameter of interest whether this is made explicit or not; often it is uniform (e.g. when calculating maximum likelihood estimates), but this is not usually stated.

3 ESTIMATING POLES

Here, we compare the effectiveness of the spherical exponential distribution and the Fisher distribution at determining the position of the pole using the statistical field model TK03 (Tauxe & Kent 2004). It is known that taking the Fisher average of unit vectors can lead to bias towards shallow inclinations (Creer 1983). Using the statistical field model of Constable & Johnson (1999), the effect of averaging unit vectors was shown to produce an apparent inclination anomaly that would be well fit by an axial octupole of about 2 per cent of the axial dipole (Johnson *et al.* 2008). A similar shallowing of inclinations when unit vectors are averaged was observed in the field model TK03 (Tauxe & Kent 2004). While the average effect is quite small, in individual cases the mean pole may fall some way from the expected position. It is of particular interest to see how often the much used α_{95} value of the Fisher distribution actually contains the expected pole.

One hundred vectors were drawn from the field model TK03 (Tauxe & Kent 2004) and the Fisher mean direction calculated, along with its 95 per cent confidence limits. Over many repetitions at random locations on the globe, the 95 per cent confidence limits for the mean were found to exclude the direction of a GAD field about 20 per cent of the time. Generally, the mean direction was shallower than the GAD field. 95 per cent confidence regions were also calculated using the likelihood function of the spherical exponential distribution, and these were found to include the GAD field direction even when the Fisher mean was too shallow. In 100 trials, with 100 vectors drawn from TK03, the 95 per cent confidence limits failed to contain the GAD direction only once. In Fig. 1, some examples are shown. By including a g_3^0 term in the statistical model, it can be shown how using the spherical exponential sampling distribution allows the degree of shallowing due to geomagnetic field behaviour to be recovered (Fig. 1d). In this case, the average direction of the full magnetic vectors lies close to the region of maximum probability for $\mathbf{u}$ calculated using a spherical exponential distribution, and it would be possible to correctly infer the contribution of the axial octupole.

From this it can be seen that adopting a spherical exponential sampling distribution all but eliminated the effect of inclination shallowing due to averaging of unit vectors and so allows not only more accurate poles to be derived, but also will reveal whether observed shallow inclinations in real palaeomagnetic data genuinely reflect geomagnetic field behaviour. As an example, consider the 390 directions from 15-Myr-old lava flows in northwest Iceland which yielded a mean direction about 5° shallower than would be expected at the latitude (Kristjansson *et al.* 2003). We found close agreement between the 95 per cent confidence limits calculated using a Fisher and a spherical exponential distribution for these data therefore, it can be concluded that averaging of unit vectors is not the reason for shallow inclinations in this data set. Fig. 2 shows the cumulative distributions of the Iceland data along with the best-fitting Fisher spherical exponential distributions. The close agreement between the spherical exponential distribution and the data suggests that low latitude VGPs are part of the normal secular variation and should not be treated differently from other data.

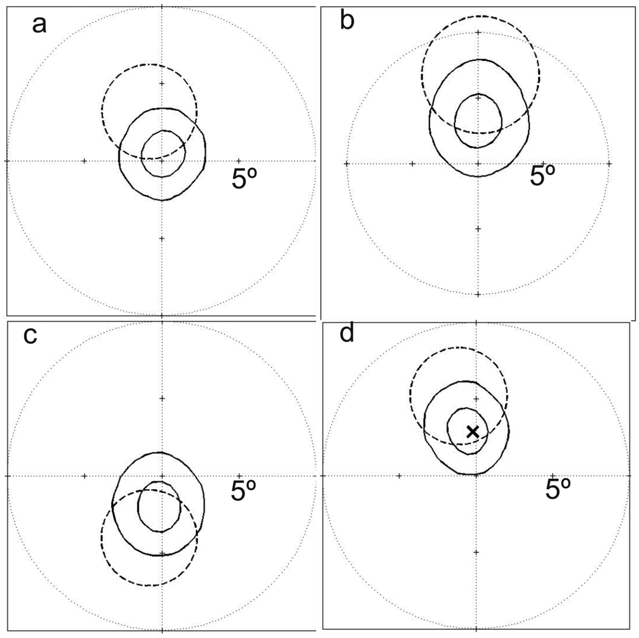

Figure 1. Azimuthal equidistant projections showing inferred central direction of groups of 100 vectors drawn from TK03. Each projection is centred on the GAD direction, and the Fisher 95 per cent confidence limits (dashed line) and 95 and 67 per cent confidence limits derived from the spherical exponential distribution are shown. In (d), an axial octupole of 5 per cent of the axial dipole is added, and the expected direction marked with a cross.

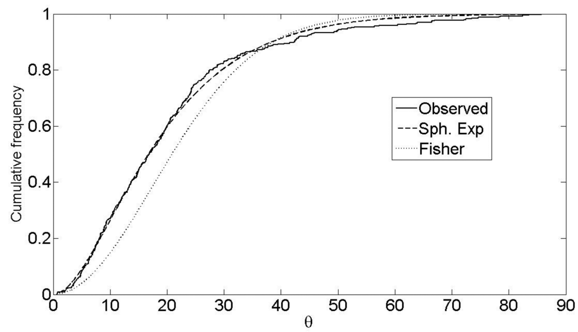

Figure 2. Cumulative frequency of angle (θ) between individual flow direction and mean direction for the northwest Iceland data of Kristjansson *et al.* (2003). Also shown are the cdfs for the best-fitting Fisher (dotted) and spherical exponential (dashed) distributions.

4 THE DISTRIBUTION OF POLES IN LARGE IGNEOUS PROVINCES

Large igneous provinces provide ideal testing grounds for statistical analyses of palaeomagnetic directions because they provide a large number of lava flows emplaced over a period that is short enough to minimize the effects of continental drift. We selected data recent studies of Cenozoic large igneous provinces that employ the best palaeomagnetic practice (see Table 1) and have at least 40 independent flow means. We considered all data to be equally valid, although

Table 1. Locations, age and number of directions along with derived statistics for each of the studies used.

Location	N	Age (Ma)	Kolmogorov prob exp		Kolmogorov prob Fisher		Likelihood ratio (dB)		Reference
			Directions	Poles	Directions	Poles	Directions	Poles	
Iceland	390	∼15	0.61	0.01	0.00	0.00	+205	+9	Kristjansson (2002)
Kerguleen	258	24–30	0.21	0.31	0.00	0.00	+145	+75	Camps *et al.* (2007)
Yemen	69	28–30	0.84	0.60	0.01	0.00	+25	+41	Riisager *et al.* (2005)
Ethiopia	65	∼30	0.45	0.50	0.50	0.02	−3	+6	Rochette *et al.* (1998)
Greenland	55	55–61	0.49	0.06	0.07	0.85	+12	−10	Riisager *et al.* (2003)
Faroes	43	55–58	0.62	0.62	0.41	0.36	+4	+5	Riisager *et al.* (2002)
Deccan	179	65	0.95	0.27	0.00	0.00	+76	+105	Vandamme *et al.* (1991)[a]

[a]Only directions with α_{95} less than 5° have been taken from the Deccan compilation of Vandamme *et al.* (1991).

where authors had put neighbouring flows into directional groups, only the group means were taken and in the case of the Deccan compilation (Vandamme *et al.* 1991) we restricted the data to those with an α_{95} of less than 5°. The studies selected all provide high quality estimates of the field direction; in the Kerguelen compilation of Camps *et al.* (2007) for instance, the average number of samples per flow is 7.6 and the average value of α_{95} is 5.5°. For each case we consider which of the two sampling distributions discussed above is most consistent with the data. In palaeomagnetism it is customary to use hypothesis tests and quantile–quantile plots to examine a distribution of poles where the distribution of colatitude and longitude are tested independently (e.g. Riisager *et al.* 2002). We note in passing that having uniformly distributed longitude does not imply that a distribution is circularly symmetric; there are many ways in which circular asymmetry can be introduced while maintaining a uniform distribution of azimuth, for instance by making k_E in eq. (4) a function of azimuth. Here we consider only the distribution of angle of deviation without considering circular symmetry.

In order to estimate the distribution of θ under a spherical exponential distribution the central direction is taken as the direction of maximum likelihood as defined by eq. (8). This was achieved by random sampling over the appropriate area of a unit sphere.

A popular hypothesis test used to compare distributions is the Kolmogorov test. Strictly speaking, it should not be used when the parameters of a distribution have been estimated from the data, but we give the Kolomogorov probabilities in Table 1 for each data set under both distributions with the understanding that the true probability would be somewhat lower. An alternative test is to compare the two distributions directly by applying a likelihood ratio test to both the directions and VGPs for the data sets in Table 1. In each case the likelihood of the data is calculated under the hypothesis of the best-fitting Fisher (L_F) and spherical exponential distribution (L_E). The ratio of likelihoods is given in decibels as 10 $\log_{10}(L_E/L_F)$, following a suggestion of Jaynes (2003), where a positive value supports a spherical exponential distribution and a negative value supports a Fisher distribution. A ratio of less than 3 dB has little significance, up to 10 dB shows a slight preference and over 20 dB is strong evidence for one model over the other. The ratios are shown in Table 1. The point of using a logarithmic scale for comparing ratios is clear as they vary by over 30 orders of magnitude. Converting the directions to VGPs is seen to decrease the likelihood ratio in some cases, but most data are fit significantly better by a spherical exponential distribution in either direction or pole space. An interesting point is that the largest data sets are seen to be very well fit by the spherical exponential distribution, but not by the Fisher distribution. This observation demonstrates the problem of using hypothesis tests on small data sets; if a small number (less than about 15) of spherical exponential data points are generated then any common statistical test will prefer the hypothesis that they are Fisher distributed. This is because the Fisher distribution, having shorter tails, will tend to return a higher likelihood for small numbers of data because the likelihood is more sharply peaked. This may make it difficult to choose the appropriate sampling distribution for small sets of data.

5 PALAEOSECULAR VARIATION

The spherical exponential distribution naturally leads to choosing the mean angle of deviation as the estimate of palaeosecular variation:

$$S' = \frac{1}{N}\sum_i \theta_i. \tag{10}$$

By taking the simple mean deviation, rather than squaring the angular deviations, S' is less affected by occasional outlying directions than S (eq. 2) and so can be considered robust to outliers. This largely removes the need for various schemes to trim excursional and transitional directions from palaeomagnetic data (e.g. Vandamme 1994). The squared deviation (S^2) traditionally studied as an indicator of PSV is the sum of two contributions. If there is a distribution of angles from a central direction (θ_i), with mean S' then the variance of θ is given by

$$V(\theta) = S^2 - S'^2. \tag{11}$$

When applied to VGPs, the statistic S^2 is therefore the sum of the square of the expected angular deviation from the pole and its variance. Dividing V by S'^2 gives a normalized measure of the variance (V'), which can be thought of as a measure of shape. For the spherical exponential distribution V' can be calculated by integrating by parts and is found to be 1/2, for large k. Random sampling sets of 50 directions generated using a spherical exponential distribution suggests that $0.35 < V' < 0.65$ in 80 per cent of cases. For a Fisher distribution, V' is largely independent of the precision κ and tends to be less than for the previous case with $0.2 < V' < 0.34$ 80 per cent of the time. With the exception of Ethiopia ($V' = 0.34$), for all of the directions from LIPs in Table 1 $0.4 < V' < 0.6$, as would be expected if they conform to spherical exponential distributions Given that the shape statistic V' is around $\frac{1}{2}$ in these LIPs, it provides a useful statistical feature to seek in field models and numerical dynamo simulations. Statistical field models inspired by the 'giant Gaussian process' (Constable & Parker 1988) include QC96 (Quidelleur & Courtillot 1996) and CJ98 (Constable & Johnson 1999) as well as TK03 (Tauxe & Kent 2004). Groups of 50 vectors are drawn at a random location for each model and vectors more than 90° from the mean direction have their direction reversed, as V' will be sensitive to occasional angles greater than 90°. For TK03, $0.41 < V' < 0.92$ 80 per cent of the time, with an average value of about 2/3. For QC96, V' lies between 0.22 and 0.48 in 80 per cent of cases and for CJ98, V' is much higher, lying between 0.5 and 2.4

Table 2. Control parameters for the three simulations and descriptive statistics of the field directions. Statistics were calculated from directions (rather than poles) and the range of the middle 80 per cent for each simulation is given.

Model	Rayleigh number	Ekman number	Thermal Prandtl	Magnetic Prandtl	S' (°)	V'
p1	120	0.0065	1	20	16.2–26.2	0.54–0.89
p2	1350	0.001	1	5	19.0–28.4	0.40–0.67
p3	2500	0.001	1	5	31.1–46.6	0.23–0.47

80 per cent of the time, and it is not uncommon to see groups of vectors for which the traditional estimate of angular dispersion, S, is 50 per cent higher than the robust estimate S'. These models were produced to simulate the secular variation over the past 5 Myr, so it is not surprising that they produce shapes of distributions that are at odds with the Cenozoic LIPs, but the shape of the distribution as given by V', provides an additional control on field models if enough data exists to constrain it. Also, it has been noted previously that much of the latitudinal dependence of the dispersion of VGPs can be accounted for by making the harmonics vary as a function of order (Harrison 2006), which none of the models considered here do.

Increasingly there has been interest in studying palaeosecular variation with respect to the output of numerical dynamo simulations run under varying boundary conditions (Aubert *et al.* 2009, 2010; Lhuillier & Gilder 2013). Here we take directions from three dynamo simulations produced by Parody-JA2.2 (Dormy *et al.* 1998; Aubert *et al.* 2008) and consider the shape of the directions produced by each as described by V'. The control parameters for each model are given in Table 2. Note that the Rayleigh number (Ra) is as defined in Christensen *et al.* (2001) and m1 is very similar to model 2 of Lhuillier *et al.* (2013). All three of the models had non-conducting cores of present day size and were sampled over several reversals. Statistics were derived by taking the directions of 100 field vectors randomly through the series at a random location on the globe. Reversed directions were rotated by 180° and S' and V' were calculated. This was repeated 200 times; Table 2 gives the middle 80 per cent range of each statistic. The magnitude of the secular variation is large compared to the Earth; S' does not exceed 21° for any of the LIPs in Table 1. This seems to be a feature of reversing dynamo simulations (see Lhuillier *et al.* 2013) and producing simulations with realistic magnitudes of VGP dispersion remains a challenge. However differences in the shapes of the directional distributions are also seen, as witnessed by V'. The values of V' seen in m2 closely resemble those seen in LIPs, averaging around 0.55, while increasing the forcing in m3 causes a great increase in the dispersion of directions, but actually reduces V'. We make no claims regarding the parameter space that might produce Earth-like simulations, but simply note that V' provides a convenient measure of the shape of a distribution which might be used in comparisons between simulations and palaeomagnetic data. Furthermore, it can be used as quick method of determining the most appropriate sampling distribution to use when finding confidence limits for poles, when the data sets are reasonably large.

6 DISCUSSION AND CONCLUSIONS

The statistic V', introduced in the last section, is a measure of the shape of a distribution of vectors and may be useful for discriminating between models. Interestingly, although it can be rather sensitive to even a single outlying datum, when applied to real palaeomagnetic data sets from LIPs, V' tends to be reasonably consistent, tending to the sort of values that are seen in a spherical exponential distribution.

The traditional estimate of PSV (S) is the quadrature sum of two parts, S', the mean of θ, and V, the variance of θ. In this study, data were not arbitrarily excluded from the analysis. Exclusion is often justified as an attempt to make the data fit a Fisher distribution (e.g. Vandamme 1994), but if a simple one-parameter maximum entropy distribution adequately describes the data, then the logic of trimming extreme values must be questioned. The values of V' seen here are all well within the expected range for the spherical exponential distribution, and this seems to be a highly appropriate sampling distribution to adopt in palaeomagnetism. In addition it gives rise to naturally robust estimators for both the central direction and dispersion of a group of vectors, so that there is less of a temptation to discard data. Given the effort that goes into palaeomagnetic field initiatives and subsequent measuring, this must be regarded as a considerable bonus. Instead of discarding data which fails to fit a Fisher distribution, a simple estimate of V' could be used by palaeomagnetists to determine the most appropriate distribution to use when making inferences from a set of directions.

As a diagnostic tool, V' can also help discriminate between field models. Palaeomagnetic data from the past 5 Myr has not been considered here, as this has been dealt with in depth elsewhere (Quidelleur *et al.* 1994; Johnson & Constable 1996; McElhinny *et al.* 1996; Johnson *et al.* 2008; Harrison 2009). In fact, it is unlikely that any distribution will improve upon that suggested of Harrison (2009) for the past 5 Myr, which yielded nearly perfect fits. There it was shown that the distribution of poles from latitudinal bands could be described by a mixture of a Fisher distribution and poles whose latitude was uniformly distributed. It is interesting to see how the shapes of these VGP distributions compare to the earlier Cenozoic LIPs considered here. As the fitted distributions describe the data well, we can simply take the models of Harrison (2009) and calculate V' for each one. It is found that V' is about 1 for the lower latitude bands and decreases to about 0.6 at higher latitudes. This would suggest that there is a real difference in the pattern of PSV seen over the last 5 Ma compared to that seen in earlier Cenozoic LIPs. The values of V' seen in the models of Harrison (2009) are clearly higher than those seen in spherical exponential distributions and are close to those seen in simulation p1 (Table 2). The spherical exponential distribution would not fit the poles from low latitude bands given by Harrison (2009) but does describe those from both Yemen and Ethiopia. The shape of distribution of magnetic vectors produced by dynamo simulation appears to be sensitive to the strength of forcing given by the Rayleigh number and consideration of this statistical feature may help determine what parts of parameter space will produce Earth-like simulations and what might be controlling changes in the behaviour of the geomagnetic field. Finally, it is worth reiterating that the spherical exponential distribution arose naturally by considering a group of vectors having an expected angle from a reference direction. The fact that the resulting distribution fits observations well suggests that this is a more accurate description of the time-averaged field than that of a group of vectors having a definite mean direction.

ACKNOWLEDGEMENTS

The authors wish to thank Julien Aubert for help in obtaining the dynamo modelling results. The manuscript benefited from helpful reviews by Chris Harrison and two anonymous reviewers. This work was funded by NERC standard Grant NE/H021043/1.

REFERENCES

Aubert, J., Aurnou, J.M. & Wicht, J., 2008. The magnetic structure of convection-driven numerical dynamos, *Geophys. J. Int.,* **172,** 945–956.

Aubert, J., Labrosse, S. & Poitou, C., 2009. Modelling the palaeo-evolution of the geodynamo, *Geophys. J. Int.,* **179,** 1414–1428.

Aubert, J., Tarduno, J.A. & Johnson, C.L., 2010. Observations and models of the long-term evolution of earth's magnetic field, *Space Sci. Rev.,* **155,** 337–370.

Biggin, A.J., Strik, G.H.M.A. & Langereis, C.G., 2008a. Evidence for a very-long-term trend in geomagnetic secular variation, *Nat. Geosci.,* **1,** 395–398.

Biggin, A.J., van Hinsbergen, D.J.J., Langereis, C.G., Straathof, G.B. & Deenen, M.H.L., 2008b. Geomagnetic secular variation in the Cretaceous normal superchron and in the Jurassic, *Phys. Earth planet. Inter.,* **169,** 3–19.

Camps, P., Henry, B., Nicolaysen, K. & Plenier, G., 2007. Statistical properties of paleomagnetic directions in Kerguelen lava flows: implications for the late Oligocene paleomagnetic field, *J. geophys. Res.: Solid Earth,* **112,** B06102, doi:10.1029/2006JB004648.

Christensen, U.R. *et al.*, 2001. A numerical dynamo benchmark, *Phys. Earth planet. Inter.,* **128,** 25–34.

Constable, C.G. & Johnson, C.L., 1999. Anisotropic paleosecular variation models: implications for geomagnetic field observables, *Phys. Earth planet. Inter.,* **115,** 35–51.

Constable, C.G. & Parker, R.L., 1988. Statistics of the geomagnetic secular variation for the past 5 m.y, *J. geophys. Res.,* **93,** 11 569–511 581.

Cox, A., 1970. Latitude dependence of the angular dispersion of the geomagnetic field, *Geophys. J. R. astr. Soc.,* **20,** 253–269.

Creer, K.M., 1983. Computer synthesis of geomagnetic palaeosecular variations, *Nature,* **304,** 695–699.

Dormy, E., Cardin, P. & Jault, D., 1998. MHD flow in a slightly differentially rotating spherical shell, with conducting inner core, in a dipolar magnetic field, *Earth planet. Sci. Lett.,* **160,** 15–30.

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

Fisher, R.A., 1934. Probability likelihood and quantity of Information in the logic of uncertain inference, *Proc. R. Soc. Lond., A,* **146,** 1–8.

Harrison, C.G.A., 2006. Variation of spherical harmonic power as a function of harmonic order for Earth's core and crustal magnetic field and for Mars' crustal field, *Geochem. Geophys. Geosyst.,* **7,** Q10014, doi:10.1029/2006GC001334.

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.

Jaynes, E.T., 2003. *Probability Theory: The Logic of Science,* ed. Larry Bretthorst, G., Cambridge Univ. Press.

Jeffreys, H., 1932. On the theory of errors and least squares, *Proc. R. Soc. Lond., A,* **138,** 48–55.

Johnson, C.L. & Constable, C.G., 1996. Palaeosecular variation recorded by lava flows over the past five million years, *Phil. Trans. R. Soc., A,* **354,** 89–141.

Johnson, C.L. *et al.*, 2008. Recent investigations of the 0–5 Ma geomagnetic field recorded by lava flows, *Geochem. Geophys. Geosyst.,* **9,** Q04032, doi:10.1029/2007GC001696.

Khokhlov, A., Hulot, G. & Bouligand, C., 2006. Testing statistical palaeomagnetic field models against directional data affected by measurement errors, *Geophys. J. Int.,* **167,** 635–648.

Kristjansson, L., 2002. Estimating properties of the paleomagnetic field from Icelandic lavas, *Phys. Chem. Earth,* **27,** 1205–1213.

Kristjansson, L., Hardarson, B.S. & Aundunsson, H., 2003. A detailed palaeomagnetic study of the oldest (≈15 Myr) lava sequences in Northwest Iceland, *Geophys. J. Int.,* **155,** 991–1005.

Lhuillier, F. & Gilder, S.A., 2013. Quantifying paleosecular variation: insights from numerical dynamo simulations, *Earth planet. Sci. Lett.,* **382,** 87–97.

Lhuillier, F., Hulot, G. & Gallet, Y., 2013. Statistical properties of reversals and chrons in numerical dynamos and implications for the geodynamo, *Phys. Earth planet. Inter.,* **220,** 19–36.

McElhinny, M.W., McFadden, P.L. & Merrill, R.T., 1996. The time-averaged paleomagnetic field 0–5 Ma, *J. geophys. Res. B: Solid Earth,* **101,** 25 007–25 027.

McElhinny, M.W. & Merrill, R.T., 1975. Geomagnetic secular variation over the past 5 m.y, *Rev. geophys. Space Phys.,* **13,** 687–708.

McFadden, P.L., Merrill, R.T., McElhinny, M.W. & Sunhee, L., 1991. Reversals of the Earth's magnetic field and temporal variations of the dynamo families, *J. geophys. Res.,* **96,** 3923–3933.

Quidelleur, X. & Courtillot, V., 1996. On low-degree spherical harmonic models of paleosecular variation, *Phys. Earth planet. Inter.,* **95,** 55–77.

Quidelleur, X., Valet, J.P., Courtillot, V. & Hulot, G., 1994. Long-term geometry of the geomagnetic field for the last five million years: an updated secular variation database, *Geophys. Res. Lett.,* **21,** 1639–1642.

Riisager, P., Riisager, J., Abrahamsen, N. & Waagstein, R., 2002. New paleomagnetic pole and magnetostratigraphy of Faroe Islands flood volcanics, North Atlantic igneous province, *Earth planet. Sci. Lett.,* **201,** 261–276.

Riisager, J., Riisager, P. & Pedersen, A.K., 2003. Paleomagnetism of large igneous provinces: case-study from West Greenland, North Atlantic igneous province, *Earth planet. Sci. Lett.,* **214,** 409–425.

Riisager, P., Knight, K.B., Bajer, J.A., Peate, I.U., Al-Kadasi, M., Al-Subbary, A. & Renne, P.R., 2005. Paleomagnetism and ^{40}Ar/^{39}Ar Geochronology of Yemeni Oligocene volcanics: implications for timing and duration of Afro-Arabian traps and geometry of the Oligocene paleomagnetic field, *Earth planet. Sci. Lett.,* **237,** 647–672.

Rochette, P. *et al.*, 1998. Magnetostratigraphy and timing of the Oligocene Ethiopian traps, *Earth planet. Sci. Lett.,* **164,** 497–510.

Smirnov, A.V., Tarduno, J.A. & Evans, D.A.D., 2011. Evolving core conditions ca. 2 billion years ago detected by paleosecular variation, *Phys. Earth planet. Inter.,* **187,** 225–231.

Tauxe, L. & Kent, D.V., 2004. A simplified statistical model for the geomagnetic field and the detection of shallow bias in paleomagnetic inclinations: Was the ancient magnetic field dipolar?, in *Timescales of the Paleomagnetic Field,* pp. 101–115, AGU.

Vandamme, D., 1994. A new method to determine paleosecular variation, *Phys. Earth planet. Inter.,* **85,** 131–142.

Vandamme, D., Courtillot, V., Besse, J. & Montigny, R., 1991. Paleomagnetism and age determinations of the Deccan traps (India): results of a Nagpur-Bombay traverse and review of earlier work, *Rev. Geophys.,* **29,** 159–190.