Geophys. J. Int. (2001) **145,** 157–171

Towards a self-consistent approach to palaeomagnetic field modelling

A. Khokhlov,[1,2] G. Hulot[2] and J. Carlut[2]

[1] *International Institute of Earthquake Prediction Theory and Mathematical Geophysics* 79, b2, *Warshavskoe shosse* 113556 *Moscow, Russia*
[2] *Institut de Physique du Globe de Paris,* 4, *Place Jussieu,* 75252, *Paris, France. E-mail: gh@ipgp.jussieu.fr*

Accepted 2000 October 10. Received 2000 April 21; in original form 1999 December 23

SUMMARY

Recent studies of the palaeomagnetic field behaviour over the past 5 Myr rely on statistical analysis of mainly directional data. However, the data are quite sparse and ill-distributed, and directional parameters are non-linear functions of the local field, rendering such statistical analysis non-trivial. Up to now these difficulties have usually been ignored or removed by relying on simplifications (linearization, neglecting internal correlations, etc.) that are unfortunately not justified if the field contains some amount of complexity.

The purpose of the present paper is to present a rigorous statistical forward approach to palaeomagnetic field modelling. Starting from a statistical model of the field defined in terms of the statistics of its Gauss coefficients (along the lines pioneered by Constable & Parker 1988), we show how such a model may be exactly tested against any given data set, either on a local regional or a global scale. A method to implement this approach is outlined and examples based on published models are provided.

In particular we focus on the treatment of directional data, for which the method is most relevant. The corresponding local probability density functions are derived and shown to be non-Fisherian, which we note may be a significant source of artefacts for standard mean-field modelling. Although the method we propose is already useful in its present state, some slight improvements are possible in order to account for noise in the data better.

Key words: directional data, magnetic fields, modelling, palaeomagnetism, statistical models.

1 INTRODUCTION

It is now well established that to first order the geomagnetic field averaged over palaeomagnetic timescales behaves like a geocentric axial dipole. However, it is also well known since the pioneering work of Wilson (1970) that this simple picture cannot account for the presence of some small persistent offset in the time-averaged palaeomagnetic field as seen in both volcanic and sediment data of the past 5 Myr. Much of this offset, known as the far-sided effect, can be interpreted in terms of an additional small quadrupole axial g_2^0 component in the time-averaged field [as shown by Wilson (1971) and confirmed by e.g. Lee (1983) and Schneider & Kent (1990)]. However, it is still not clear how much additional structure must be present in the time-averaged field to account for the global set of palaeomagnetic data. Indeed, whereas Gubbins & Kelly (1993), Johnson & Constable (1995) and Kelly & Gubbins (1997) argued for the existence of persistent higher-order features, more recent studies suggest that many of these features could be unresolved (e.g. McElhinny *et al.* 1996; Johnson & Constable 1997; Carlut & Courtillot 1998). This led these authors to propose much simpler mean-field models.

As well as the average field, a significant multipolar field is also found that fluctuates on timescales of the order of a couple of centuries (e.g. Bloxham & Jackson 1992; Hongre *et al.* 1998). Volcanic palaeomagnetic samples from various regions are highly unlikely to be of the same age within that level of precision. This makes it possible to study the statistical behaviour of this fluctuating field by considering the volcanic data as statistical samples. These so-called palaeosecular variation (PSV) studies have mainly been carried out with the help of volcanic directional data (declination and inclination) and most often by transforming the raw data into virtual geomagnetic pole (VGP) positions (for a recent review see e.g. Merrill *et al.* 1996).

© 2001 RAS

Such studies are not as straightforward as one could wish for a variety of reasons. One is the limited quality of the data, due to experimental uncertainty. This translates into a within-site dispersion that is traditionally modelled in terms of a random Fisherian noise that adds on to the local geomagnetic field direction. This approach is not free of problems, especially when it comes to studying the statistics of VGPs rather than those of field directions (see e.g. Johnson & Constable 1996), and it is clear that more progress will have to be made with respect to this matter.

Another intrinsic source of problems is the possibility of tectonic movements significantly altering the original palaeosignal. These effects can be quite large and are not easy to detect, even though great care is usually taken to avoid inclusion of suspicious data within the data sets to be analysed. This problem will not be dealt within this paper, which focuses on yet another issue (for a recent discussion of tectonic effects, see e.g. McElhinny *et al.* 1996).

Even in a situation when the data could be assumed to be free of such problems (an assumption that is often made in PSV studies, despite what has just been stated above), one very critical problem remains: only a limited number of sites can be found at the Earth's surface where volcanic data can actually be sampled. In addition, these sites tend to be clustered. In most of the databases presently available (corresponding to the past 5 Myr), out of about 100 sites, only 40 to 50 regions at the Earth's surface can for instance be considered as independent, each region being unevenly represented by as few as 20 to as many as several hundred palaeomagnetic field samples (e.g. Quidelleur *et al.* 1994; Johnson & Constable 1996).

This uncomfortable situation has led virtually all authors (with the noteable exception of the very recent study of Constable & Johnson 1999) to study PSV by binning the raw data from all sites within a wider region, such as a band of latitude, and studying the way these data scatter about some expected value. Such procedures have a distinct advantage: they allow one to plot statistical quantities that are based on a large number of samples and are therefore recovered with reasonable confidence. However, they suffer two significant drawbacks. One is that these quantities are statistically meaningful only to the extent that the field is behaving in a way that is compatible with the way the data are being binned. The other is that these quantities are not trivially related to simple field characteristics (such as the ratio of dipole to non-dipole fields). These difficulties have been responsible for increasing numbers of PSV models, each model interpreting the data by relying on its own set of simplifying assumptions [models A, B, C, D, E, F, G and H; see Merrill *et al.* (1996)].

The first paper to recognize the need for a well-defined and general statistical framework for PSV studies was that of Constable & Parker (1988). These authors noted that the best parameter space to describe the palaeomagnetic field in statistical terms is simply the space of the Gauss coefficients currently used in geomagnetic field modelling (e.g. Langel 1987). This approach requires no special assumption and is most useful. It provides a sound means of comparing the present geomagnetic field to the palaeomagnetic field. It also provides a general framework within which the concepts of average and fluctuating parts of the field, as well as all statistical assumptions, can be defined in a straightforward manner. Finally, it provides a basis for predicting the average and statistical behaviour of any observable quantity. This, in fact, is what led Constable & Parker (1988) to propose the first consistent statistical model for both the average and fluctuating components of the field for the past 5 Myr (at periods of stable polarity)—a model they termed the 'giant Gaussian process' (GGP; more details are given in Section 2).

Subsequent papers (e.g. Kono & Tanaka 1995; Hulot & Gallet 1996; Quidelleur & Courtillot 1996; Kono 1997; Constable & Johnson 1999) have generalized this approach by both relaxing the simple axisymmetric assumptions of the original GGP and producing additional statistical predictions to be compared with the data. These studies have highlighted the power of the GGP approach. However, all of them have always only considered predictions for one field parameter at a time (inclination, declination, etc.) and almost always for sites binned in bands of latitude. This fails to take into account possible correlations between different types of data (such as inclination and declination, which are necessarily correlated when recovered from the same sample). It is also largely inadequate if the field does not have axisymmetrically invariant statistical properties, which is precisely the property of the field one would wish to test. Finally, it provides weakly discriminating results: many different statistical models can indeed map into similar effects (see e.g. Hulot & Gallet 1996; Constable & Parker 1988; Constable & Johnson 1999).

The purpose of the present paper is to show that many of these difficulties can be avoided by slightly altering the way the data are being compared to the PSV and average field models. Rather than starting from the data, selecting parameters and binning the raw data from different sites in order to produce second-generation data to be compared to the models, we will show that it is possible to start from a model (that is, a generalized version of the GGP model), produce the local statistical behaviour for the field at each site and deduce a local, regional or global measure of the adequacy of the model for the data. We will focus on the most common case, that is, when the data set is mainly, if not entirely, composed of directional data (such as those of Lee 1983; Quidelleur *et al.* 1994; Johnson & Constable 1996; McElhinny & McFadden 1997). This is the case for which the approach we develop could be most useful. Indeed, because of the non-linear relation between the local directional parameters and the magnetic field model (i.e. the Gauss coefficients), trying to recover the average field without properly including the statistics of the fluctuating field [as done by e.g. Gubbins & Kelly (1993) and Johnson & Constable (1995) and recently Carlut & Courtillot (1998)] or vice versa (as in Johnson & Constable 1996) can in fact be hazardous (see Hulot & Gallet 1996; Kono *et al.* 2000). Unlike these previous studies, our approach requires no approximation and makes it possible to address the properties of the two fields simultaneously.

In what follows, starting from any generalized GGP model (Section 2), we derive for each site the *exact* statistical distribution for the local field (Section 3) and its direction (Section 4) (inclination and declination *not* being considered independently). We describe a uniformization procedure, which makes it possible to compare the prediction of the model with any type of data on a local, regional or global basis. This makes it possible to bin the local statistical information into more global information without any approximation. This provides us with a means to test any statistical model (with any assumed mean and fluctuating field) against the full data set on both a local and a global scale (Section 5). In this respect our approach

can be viewed as an exact forward approach. We finally apply our method to test a number of published PSV and mean-field models and to illustrate the rather complex nature of the local statistical behaviour of the direction of the field (Section 6). As we shall show, this complex nature may indeed lead to significant artefacts when trying to recover the average field using current methods of inversion and purely directional data.

2 DESCRIBING THE PALAEOMAGNETIC FIELD IN TERMS OF A 'GENERALIZED GIANT GAUSSIAN PROCESS'

The concept of GGP was introduced by Constable & Parker (1988). It consists of taking advantage of the fact that at any given moment the Earth's main magnetic field is derived from a potential V that can always be expanded in terms of a spherical harmonic model,

$$V(\Re, \Theta, \Psi) = a \sum_{l=1}^{\infty} \sum_{m=0}^{l} \left(\frac{a}{\Re}\right)^{l+1} \left(g_l^m \cos m\Psi + h_l^m \sin m\Psi\right) \times P_l^m(\cos\Theta), \quad (1)$$

where g_l^m and h_l^m are the so-called Gauss coefficients and $\{\Re, \Theta, \Psi\}$ are the standard spherical coordinates (the distance from the Earth's centre, colatitude and longitude). (We specify these notations in order to distinguish the several spherical coordinate systems that we need to use.) Next it is assumed that the temporal evolution of the field during a given long time period (say between two reversals) can be described in terms of statistical fluctuations of the field about a mean field.

From eq. (1), this temporal evolution can be described in terms of fluctuations of the 'model' vector $\mathbf{k} = \{\mathbf{g}, \mathbf{h}\}$ consisting of Gauss coefficients $\mathbf{g} = \{g_1^0, g_1^1, \ldots, g_l^m \ldots\}$, $\mathbf{h} = \{h_1^1, h_2^1, \ldots, h_l^m \ldots\}$ about some average model $\bar{\mathbf{k}} = \{\bar{\mathbf{g}}, \bar{\mathbf{h}}\}$. This 'model' vector is therefore assumed to fluctuate within what will be defined as the (infinite-dimensional) 'model space'. Assuming that these fluctuations can be described in terms of a short-term memory *stationary random Gaussian process* [which is consistent with both historical geomagnetic and archaeomagnetic variations; see e.g. Hulot & Le Mouël (1994) and Hongre *et al.* (1998)] and that any two palaeomagnetic observations are always separated by a period of time larger than the memory of the process (of the order of a couple of centuries), each palaeomagnetic datum can then be viewed as a local (both in time and in space) independent realization of a random Gaussian drawing of $\mathbf{k}$. Describing the palaeomagnetic field in terms of generalized GGP simply consists of identifying the first and second moments of the Gaussian statistics of $\mathbf{k}$ best predicting the observed statistics for such palaeomagnetic data. The mean and fluctuating fields are then characterized by the means $\mathrm{E}(g_i^j)$, $\mathrm{E}(h_k^l)$ and covariances $\mathrm{cov}(g_i^j, g_k^l)$, $\mathrm{cov}(h_i^j, h_k^l)$ and $\mathrm{cov}(g_i^j, h_k^l)$, i.e. by $\mathrm{E}(\mathbf{k})$ and $\mathrm{Cov}(\mathbf{k}, \mathbf{k}) = [\mathrm{cov}(k_\alpha, k_\beta)]$.

This description of the palaeomagnetic field generalizes the original GGP of Constable & Parker (1988), which was built upon two additional important simplifying assumptions: first that the Gauss coefficients could be considered as being independent of one another, and second that all Gauss coefficients sharing the same degree n would share the same variance σ_n. In other words, the original GGP further assumed that

$$\mathrm{cov}(h_i^j, h_k^l) = \mathrm{cov}(g_i^j, g_k^l) = \delta_{ik}\delta_{jl}\sigma_k^2, \quad (2)$$

$$\mathrm{cov}(g_i^j, h_k^l) \equiv 0. \quad (3)$$

This in fact amounts to assuming that the fluctuating field has no specific frame of reference [that is, that covariances do not depend on the geographic frame of reference for $\{\Re, \Theta, \Psi\}$; see Eckhardt (1984) and Hulot & Le Mouël (1994)]. This sounds like a sensible first-order approximation, but it is already known to be in error. Indeed, Hulot & Gallet (1996) have explicitely shown that no model satisfying eqs (2) and (3) can be found that can account for the VGP scatter curve corresponding to the past 5 Myr. In fact, all recent PSV models for that period include some amount of anisotropy within the degree 2 fluctuating field (e.g. Quidelleur & Courtillot 1996; Johnson & Constable 1996; Constable & Johnson 1999). This is the reason why in this paper the most general case will be considered.

3 GENERALIZED GGP BEHAVIOUR OF THE LOCAL MAGNETIC FIELD

For any point q with coordinates $(\Re, \Theta, \Psi)$, the local field vector is

$$\mathbf{B}(q) = \mathbf{B}(\Re, \Theta, \Psi) = -\mathrm{grad}\, V(\Re, \Theta, \Psi).$$

From eq. (1) this expression is linear with respect to Gauss coefficients g_l^m and h_l^m, i.e. $\mathbf{k}$. For each point q we therefore have a linear operator $A(q)$ that maps $\mathbf{k}$ onto $\mathbf{B}(q)$,

$$\mathbf{B}(q) = A(q)\mathbf{k}.$$

Within the generalized GGP, the probability distribution of $\mathbf{k}$ (or its finite-dimensional approximation) is assumed to be Gaussian. The probability distribution of the 3-D random vector $\mathbf{B}(q)$ will therefore also be Gaussian. It is defined by nine real parameters: three corresponding to the local value of the average field $\mathrm{E}(\mathbf{B})$ and six corresponding to a symmetric 3×3 covariance matrix $\mathrm{Cov}(\mathbf{B}, \mathbf{B})$. These parameters are trivially related to the mean model $\bar{\mathbf{k}} = \mathrm{E}(\mathbf{k})$ and the model covariance matrix $\mathrm{Cov}(\mathbf{k}, \mathbf{k})$ through

$$\mathrm{E}(\mathbf{B}(q)) = A(q)\mathrm{E}(\mathbf{k}) = A(q)\bar{\mathbf{k}}, \quad (4)$$

$$\mathrm{Cov}(\mathbf{B}(q), \mathbf{B}(q)) = A(q)\mathrm{Cov}(\mathbf{k}, \mathbf{k})A(q)^{\mathrm{T}}. \quad (5)$$

The GGP model therefore entirely and straightforwardly defines the distribution (and the corresponding probability density function) for the Gaussian field vector $\mathbf{B}(q)$ at any given point q.

4 GENERALIZED GGP BEHAVIOUR OF THE DIRECTION OF THE LOCAL MAGNETIC FIELD

To test a given generalized GGP model against a data set composed of directional data, the statistics describing the generalized GGP behaviour of the direction of the local magnetic field are needed. These statistics are the $\mathbf{B}_{(q)}$ statistics integrated over all

possible intensity values for $|\mathbf{B}(q)|$. Mathematically speaking, the problem therefore corresponds to deriving the distribution of directions of a Gaussian random vector, i.e. to deriving the pdf (probability density function) for $\mathbf{B}(q)/|\mathbf{B}(q)|$, knowing that the local vector $\mathbf{B}(q)$ satisfies a Gaussian distribution characterized by $\mathrm{E}(\mathbf{B}(q))$ and $\mathrm{Cov}(\mathbf{B}(q), \mathbf{B}(q))$. This distribution on the unit sphere will be referred to as a Gaussian directional distribution ($\mathscr{GD}$ distribution). As we show below, a general analytic expression for the corresponding pdf can in fact be derived.

4.1 Gaussian directional distribution density function

Consider the unit sphere Σ in $\mathbb{R}^3$ centred at the origin O. Suppose we are given a 3-D Gaussian random vector $\mathbf{U}$. Its probability density function may then be written as $f_U(\rho, \mathbf{s})$, where $\mathbf{s} \in \Sigma$ and ρ denotes the distance from the origin O. The projection of the random vector $\mathbf{U}$ along the rays from the origin O to the unit sphere Σ defines the random distribution we are searching for—the Gaussian directional distribution—on Σ. The corresponding probability density function simply reads

$$f_{\mathscr{GD}}(s) = \int_0^\infty f_U(\rho, s) d\rho \,. \tag{6}$$

In a local Cartesian coordinate system (u_1, u_2, u_3) with the same origin O, the Gaussian vector field $\mathbf{U}$ can be characterized by $\mathrm{E}(\mathbf{U}) = (\bar{U}_1, \bar{U}_2, \bar{U}_3)$ and $\mathrm{Cov}(\mathbf{U}, \mathbf{U}) = [\mathrm{cov}(U_i, U_j)]$. Let $\mathbf{\Lambda} = [\Lambda_{i,j}]$ be the inverse (and hence also symmetric) matrix of $\mathrm{Cov}(\mathbf{U}, \mathbf{U})$. With respect to these local Cartesian coordinates, the pdf of $\mathbf{U}$ is

$$f_U(u_1, u_2, u_3) = \sqrt{\frac{\det \mathbf{\Lambda}}{(2\pi)^3}} \exp\left[-\frac{1}{2} \sum_{i,j=1}^{3} \Lambda_{ij} (u_i - \bar{U}_i)(u_j - \bar{U}_j) \right]. \tag{7}$$

With the help of an appropriate choice of the local Cartesian coordinate system it is always possible to ensure that $\bar{U}_1 = \bar{U}_2 = 0$, $\bar{U}_3 = U$, so that $\mathrm{E}(\mathbf{U}) = (0, 0, U)$. Using the local spherical coordinate system

$$u_1 = \rho \sin\theta \cos\varphi \,, \quad u_2 = \rho \sin\theta \sin\varphi \,, \quad u_3 = \rho \cos\theta$$

in eq. (7), which we then use in eq. (6), finally leads to the expression of the pdf for the $\mathscr{GD}$ distribution,

$$f_{\mathscr{GD}}(\theta, \varphi) = \sqrt{\frac{\det \mathbf{\Lambda}}{(2\pi)^3}} \int_0^{+\infty} \mathrm{e}^{-\frac{1}{2}(a\rho^2 + b\rho + c)} \rho^2 d\rho \,, \tag{8}$$

where

$$a = a(\theta, \varphi), \; b = b(\theta, \varphi)$$

and

$$a = \Lambda_{11} \sin^2\theta \cos^2\varphi + \Lambda_{12} \sin^2\theta \sin 2\varphi + \Lambda_{13} \sin 2\theta \cos\varphi + \Lambda_{22} \sin^2\theta \sin^2\varphi + \Lambda_{23} \sin 2\theta \sin\varphi + \Lambda_{33} \cos^2\theta \,,$$

$$b = U(-2)(\Lambda_{13} \sin\theta \cos\varphi + \Lambda_{23} \sin\theta \sin\varphi + \Lambda_{33} \cos\theta) \,,$$

$$c = U^2 \Lambda_{33} \,.$$

The integral in eq. (8) may be evaluated (using e.g. the MATHEMATICA 3.0 software) as

$$\int_0^{+\infty} \mathrm{e}^{-\frac{1}{2}(a\rho^2 + b\rho + c)} \rho^2 d\rho = \frac{\mathrm{e}^{-c/2}}{2a^2} \left\{ \sqrt{\frac{\pi}{2}} \frac{(4a + b^2)\left[1 - \mathrm{Erf}\left(\frac{b}{2\sqrt{2a}}\right)\right] \mathrm{e}^{b^2/8a}}{2\sqrt{a}} - b \right\}, \tag{9}$$

and eq. (8) becomes

$$f_{\mathscr{GD}}(\theta, \varphi) = \frac{\mathrm{e}^{-c/2}}{2a^2} \sqrt{\frac{\det \mathbf{\Lambda}}{(2\pi)^3}} \left\{ \frac{\sqrt{\pi}(4a + b^2)\left[1 - \mathrm{Erf}\left(\frac{b}{2\sqrt{2a}}\right)\right] \mathrm{e}^{b^2/8a}}{2\sqrt{2a}} - b \right\}, \tag{10}$$

where the dependence on (θ, φ) is implicit in the functions a and b. This formula generalizes the result of Bingham (1983), who derived a series expansion of $f_{\mathscr{GD}}(\theta, \varphi)$ in the special symmetric case where $[\Lambda_{i,j}]$ is a scalar matrix.

4.2 Application to the local magnetic field

In the case we are currently interested in, the statistical behaviour of the field over the Earth's surface is defined by the generalized GGP parametrized by $\mathrm{E}(\mathbf{k})$ and $\mathrm{Cov}(\mathbf{k}, \mathbf{k})$ as described in Section 2. For any given site q, the parameters describing the behaviour of $\mathbf{B}(q)$ can then be computed as a function of $\mathrm{E}(\mathbf{k})$ and $\mathrm{Cov}(\mathbf{k}, \mathbf{k})$ (see eqs 4 and 5). The behaviour of the local magnetic field and its direction at this site q are thus entirely described by the pdfs $f_U(u_1, u_2, u_3)$ and $f_{\mathscr{GD}}(\theta, \varphi)$ in eqs (7) and (10), provided that

(i) we choose the local spherical coordinate system such that the direction $\theta = 0$ corresponds to that of the local average field $\mathrm{E}(\mathbf{B}(q))$ (produced by eq. 4);
(ii) we set $U = |\mathrm{E}(\mathbf{B}(q))|$;
(iii) we set $\mathbf{\Lambda}$ equal to the inverse matrix of $\mathrm{Cov}(\mathbf{B}(q), \mathbf{B}(q))$.

In practice, when a standard north, east and vertical set of local directions is being used (or any other useful local frame), this further means that one should first perform a proper frame rotation before applying eqs (7) and (10). Such a frame rotation can, however, easily be implemented numerically. For any generalized GGP the pdfs of the local magnetic field and its direction at any given location q can thus directly be deduced once the parameters $\mathrm{E}(\mathbf{k})$ and $\mathrm{Cov}(\mathbf{k}, \mathbf{k})$ have been specified. The corresponding pdfs are denoted $[f_U]_q(\rho, \mathbf{s})$ and $[f_{\mathscr{GD}}]_q(\theta, \varphi)$ hereafter.

5 TESTING A WORLDWIDE DATA SET AGAINST A GENERALIZED GGP

5.1 General considerations

Testing the data set against a generalized GGP model [that is, given values of $\mathrm{E}(\mathbf{k})$ and $\mathrm{Cov}(\mathbf{k}, \mathbf{k})$] would be quite straightforward if for each site q we had enough data to compare the

observed density of data to the predicted pdfs $[f_U]_q(\rho, \mathbf{s})$ and $[f_{\mathscr{GD}}]_q(\theta, \varphi)$. However, in practice the data are usually unevenly distributed among various sites and the amount of available data for a given site may sometimes be quite small. The problem thus amounts to assessing the simultaneous likelihood of the data distribution observed at various locations in terms of the local pdfs predicted by the generalized GGP model against which the data are to be tested.

This can be done by modifying a standard procedure often used to test a data set $\{x_i\}$ against a univariate density function $f(x)$; we refer to this as 'uniformization' (as we know of no conventional name for this procedure). This procedure consists of transforming the data set $\{x_i\}$ into a new data set $\{t_i\}$ defined on [0, 1] by

$$t_i = \int_{-\infty}^{x_i} f(x)dx\,. \tag{11}$$

Testing the data set $\{x_i\}$ against $f(x)$ is then equivalent to testing $\{t_i\}$ against a uniform distribution on [0, 1]. This kind of procedure is usually used implicitly. It is also often used in its inverse form (producing $\{x_i\}$ from $\{t_i\}$) as an algorithm to produce data sets conforming to the density function $f(x)$ when a numerical uniform random generator over [0, 1] is already available (see e.g. Press *et al.* 1996, and references therein). Here it needs to be modified to account for the fact that the density functions are multivariate [$[f_U]_q(\rho, \mathbf{s})$ and $[f_{\mathscr{GD}}]_q(\theta, \varphi)$].

For $[f_{\mathscr{GD}}]_q(\theta, \varphi)$ this could be done by modifying the procedure in such a way that the transformed data set would have to be tested against a uniform distribution on the unit sphere rather than on [0, 1]. However, this procedure would be specific to $[f_{\mathscr{GD}}]_q(\theta, \varphi)$ and would stop us from possibly combining data sets composed of both directional and three-component data [for which $[f_U]_q(\rho, \mathbf{s})$ is relevant].

To avoid this difficulty we decided to keep the original idea of transforming the data set into a new 1-D data set on [0, 1] to be tested against a uniform distribution. This '1-D uniformization' of multidimensional data is described in the next section. It has one major advantage within the palaeomagnetic context that we are dealing with: because each palaeomagnetic datum is assumed to be an independent realization of the generalized GGP, it will indeed be possible to uniformize each type of data independently for each site using the relevant [type of data- and site-dependent $[f_U]_q(\rho, \mathbf{s})$ or $[f_{\mathscr{GD}}]_q(\theta, \varphi)$] functional form of the generalized GGP and combine as many transformed data as we wish. Thus, 1-D uniformization allows us to carry out tests at all levels: locally, on a site by site basis; regionally, by combining sites within a region; globally, by combining all sites.

However, 1-D uniformization has some drawbacks. It maps a multidimensional distribution into a 1-D distribution. This has two consequences, first, the procedure is non-unique, and second, for any chosen procedure there will always be a degree of freedom for the data set not to be compatible with the original multivariate density function even if it appears to be compatible with a uniform distribution after uniformization. However, if after 1-D uniformization the data set fails to be compatible with a uniform distribution on [0, 1], then we may state that it is incompatible with the original multivariate density function. In this respect we may say that the procedure is conservative in the usual statistical sense. As we shall see, it turns out that tests based on this procedure are in fact quite sensitive if the 1-D uniformization is well chosen.

5.2 1-D uniformization of a multidimensional data set

As stated above, 1-D uniformization of a multidimensional data set $\{x_i\}$ ($x_i \in X$) to be tested against a pdf $f_\alpha(x)$ aims at producing a 1-D data set $\{t_i\}$ over [0, 1] such that if $\{x_i\}$ is compatible with $f_\alpha(x)$ then $\{t_i\}$ is compatible with a uniform pdf over [0, 1]. Such a procedure can be defined in the following way. First we introduce a one-parameter family $\{U_t\}$, $t \in [0, 1]$ of subsets $U_t \subset X$ such that

$$\begin{aligned} &P\{\alpha \in U_t\} \equiv \int_{U_t} f_\alpha(x)dx = t\,, \\ &U_0 = \varnothing\,, \quad U_1 = \Sigma\,, \\ &x \le y \Longleftrightarrow U_x \subseteq U_y\,, \end{aligned} \tag{12}$$

where $P\{\alpha \in U_t\}$ stands for the probability of the random variable α falling within U_t. Having such a family we can next transform the data set $\{x_i\}$ into $\{t_i\}$ by applying the following rule: if $x_i \in U_{\bar{t}}$ and for all $t' < \bar{t}$ we have x_i in the complement of $U_{t'}$, then set $t_i = \bar{t}$. Less formally this means that we seek t_i such that x_i belongs to the boundary of U_{t_i}. As can easily be checked, this rule defines a 1-D uniformization that achieves what we want. In the case of 1-D data this procedure is exactly the standard uniformization procedure eq. (11) for testing against a univariate pdf.

In the multidimensional case, it is clear that many different families $\{U_t\}$ satisfying eq. (12) can be found (reflecting the fact that we map a multidimensional space to a line). This is how different procedures can in fact be defined. However, only one family can be defined without needing additional information [other than the one provided by the $f_\alpha(x)$ pdf]. [For more details explaining the properties of these families $\{U_t\}$, see e.g. Kullback (1968), Chapter 2, who discusses these within the context of information theory.]

This special family is the one for which each U_t is bounded by an isovalue of f_α. It satisfies

$$U_t = \{x | f_\alpha(x) \ge C(t)\}\,, \tag{13}$$

where $C(t)$ is a monotonically decreasing positive real function of $t \in (0, 1)$ entirely defined by the equality $P\{\alpha \in U_t\} = t$. As can easily be checked, using the family of eq. (13) leads to a 1-D uniformization defined by a very simple rule: for each sample x_i of $\{x_i\}$, define

$$t_i = P\{x | f_\alpha(x) \ge f_\alpha(x_i)\} = \int_{\{x | f_\alpha(x) \ge f_\alpha(x_i)\}} f_\alpha(x)dx\,. \tag{14}$$

In practice, the right-hand side integral of eq. (14) is then evaluated numerically. Standard software such as MATHEMATICA 3.0 allows this to be done. However, the computation is quite time-consuming and a specific algorithm had to be developed to achieve the computation in a more efficient way.

Formula (14) defines the 1-D uniformization that we use in the following.

5.3 Application to the palaeomagnetic data set

When dealing with the palaeomagnetic data set, $f_\alpha(x)$ may equally well represent the 3-D statistics $[f_U]_q$ of the field at some location q where we have full-field data (in which case $X=\mathbb{R}^3$) or the 2-D $\mathcal{GD}$-distribution statistics [i.e. $[f_{\mathcal{GD}}]_q(\theta, \varphi)$] at some other location where we only have directional data available (in which case X is the unit sphere Σ). However, for the purpose of the present paper, which aims more at illustrating the method than at carrying out a systematic study of all available palaeomagnetic data, only the most common case of directional data is considered in detail. 1-D uniformization is then implemented with the help of eq. (14), where $f_\alpha(s)$ is taken to be $[f_{\mathcal{GD}}]_q(\theta, \varphi)$, as given by eq. (10).

For each site q and after 1-D uniformization, we are left with a set of uniformized data $\{t_i\}$ to be tested against the uniform distribution over [0, 1]. Since a variety of tests exist to check such a uniform distribution (Knuth 1981), testing a generalized GGP can ultimately be done quite straightforwardly. Binning uniformized data from as many sites as we wish makes it very easy to carry out these tests on a local, regional or global scale. The paragraph below summarizes the whole procedure.

5.3.1 Summary

The practical testing of a given directional datum against a candidate generalized GGP finally amounts to the following four steps, given a data set from various sites, E(**k**) and Cov(**k**, **k**) and defining the generalized GGP.

Step 1. For each site q find the centre and moments of the Gaussian random vector $\mathbf{B}_q=(\mathrm{grad}\,V)_q=\mathrm{grad}\,V(\Re_q, \Theta_q, \Psi_q)$ as described in Section 3.

Step 2. For each site q compute the probability density function $[f_{\mathcal{GD}}]_q(\theta, \varphi)$ according to eq. (10).

Step 3. For each site q carry out the 1-D uniformization of the $\{s_i\}_q$ data set and derive the uniformized $\{t_i\}_q$ data set as described in this section.

Step 4. Join as many $\{t_i\}_q$ as needed into one local, regional or global uniformized data set $\{t_i\}$ and apply to it the criterion of one's choice in order to test whether this sample represents a uniform random distribution over [0, 1] or not.

6 DISCUSSION

In this section, we wish to provide simple examples of application of the theoretical considerations developed in the previous sections. It is not our intention to be exhaustive, but rather to demonstrate the usefulness of the approach we suggest. We first point out that if we are ready to believe that the field can be described in terms of a generalized (or even classical) GGP—this, we should stress again, remains an arbitrary assumption—then the pdf of the direction of the field can have quite complex non-Fisherian properties. We show that this may lead to significant artefacts when trying to construct mean-field models with the help of averages of pure directional data in the way palaeomagnetists currently do. To illustrate the point further and to provide examples of practical applications of our method, we finally test a number of recently published PSV and time-averaged field models against a small but well-controlled data set covering the Bruhnes chron and corresponding to volcanic directional data acquired at 26 sites distributed worldwide (see Fig. 1 for details). It is shown that the method proves sufficiently efficient to warrant further applications to palaeomagnetic field modelling.

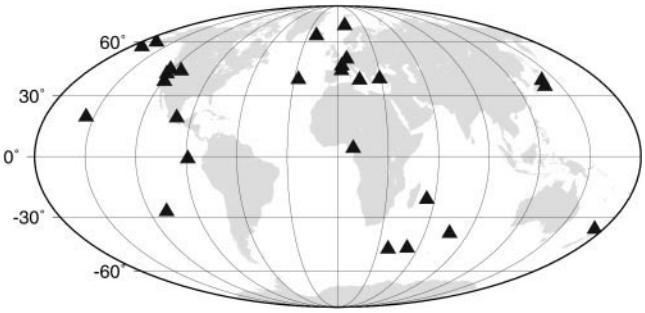


Figure 1. Locations of the 26 sites of the Bruhnes data set.

The data set we rely on has been extracted from the Quidelleur *et al.* (1994) database and is available on the Web (updated in 1999) at http://www.ipgp.jussieu.fr/obs/data/paleomag/var-secu/. The various models we test are

(i) the original model of Constable & Parker (1988) (hereafter CP88);

(ii) model C1 of Quidelleur & Courtillot (1996) (hereafter QC96);

(iii) model CJ98 recently proposed by Constable & Johnson (1999).

Additional models, CP88-0, QC96-0 and CJ98-0, derived from these three models by reducing the average field to its single axial dipole component, are also considered for illustration purposes. Table 1 gives an exact description of each of these models.

6.1 Non-Fisherian properties of the local directional probability density function

Two models have been considered as starting points to illustrate the complex non-Fisherian behaviour of the local directional pdf, QC96 and CJ98. These two models are typical of the GGP models that have been proposed. The fluctuations of the field are characterized by covariances of the forms eqs (2) and (3)—and thus assume no correlation between the Gauss coefficients and the same variance σ_k for Gauss coefficients sharing the same degree k—except for the degree 2 coefficients

Table 1. Mean field and variances of the GGP models tested in this study. The parameter α defines the $\sigma_n^m=\sigma_n=[\alpha(c/a)^n/[(n+1)(2n+1)]^{1/2}]$ for $n\geq 3$. All tests have been carried out by considering contributions up to degree $n=7$ (higher-degree contributions appear to be negligible). All values in μT.

	CP88	QC96	CJ98	CP88-0	QC96-0	CJ98-0
$E(g_1^0)$	−30.0	−30.0	−30.0	−30.0	−30.0	−30.0
$E(g_2^0)$	−1.8	−1.2	−1.5	0	0	0
σ_1^0	3.0	3.0	11.72	3.0	3.0	11.72
σ_1^1	3.0	3.0	1.67	3.0	3.0	1.67
σ_2^0	2.14	1.3	1.16	2.14	1.3	1.16
σ_2^1	2.14	4.3	4.06	2.14	4.3	4.06
σ_2^2	2.14	1.3	1.16	2.14	1.3	1.16
α	27.7	27.7	15.0	27.7	27.7	15.0

in the case of QC96 and the degree 1 and 2 coefficients in the case of CJ98, which are also assumed uncorrelated but have variances σ_k dependent on the order of the Gauss coefficient. These dependences were originally introduced as a first-order attempt to improve the fit to the set of data-derived PSV indicators (VGP scatter curves for the past 5 Myr), which models without such dependence, such as CP88, unquestionably fail to account for (see e.g. Hulot & Gallet 1996). Both models also assume a mean field composed of only an axial dipole and an axisymmetric quadrupole. This axisymmetric quadrupole was introduced to account for the fact that the data tend to show a slight average offset in the inclinations, known as the far-sided effect (Wilson 1970; see also Constable & Parker 1988 and Quidelleur *et al.* 1994), which is widely thought to be the most reliable feature within the mean field of the past 5 Myr (e.g. Carlut & Courtillot 1998). Thus, QC96 and CJ98 may be viewed as two typical GGP models accounting for what are thought to be the most robust features in the data for the past 5 Myr.

Figs 2 and 3 show examples of pdfs for the two models QC96-0 and CJ98-0, derived from the two previous models QC96 and CJ98 by reducing the average field to only its axial dipole component [i.e. by setting $\mathrm{E}(g_2^0)=0$]. The reason we chose these two models (rather than the original models) will soon become clear. In these figures, which are plotted for theoretical sites in the Northern Hemisphere, the convention is as follows: for each site, the unit sphere is Lambert-projected, the centre point (pointing downwards on the page) corresponds to the direction produced by a pure axial dipole field, and two axes representing the S–N and E–W directions are also plotted (north being towards the top of the page, east being towards the right). The ellipses are pdf isovalues representing the bounds within which 0 per cent (central point), 10 per cent, ..., 90 per cent of the directions are expected to lie.

Because the models are statistically axisymmetric [this follows from the facts that the variances for the g_n^m and h_n^m are identical and the assumed mean field is axisymmetric (Hulot & Le Mouël 1994)], the pdfs only depend on the latitudes of the sites and not

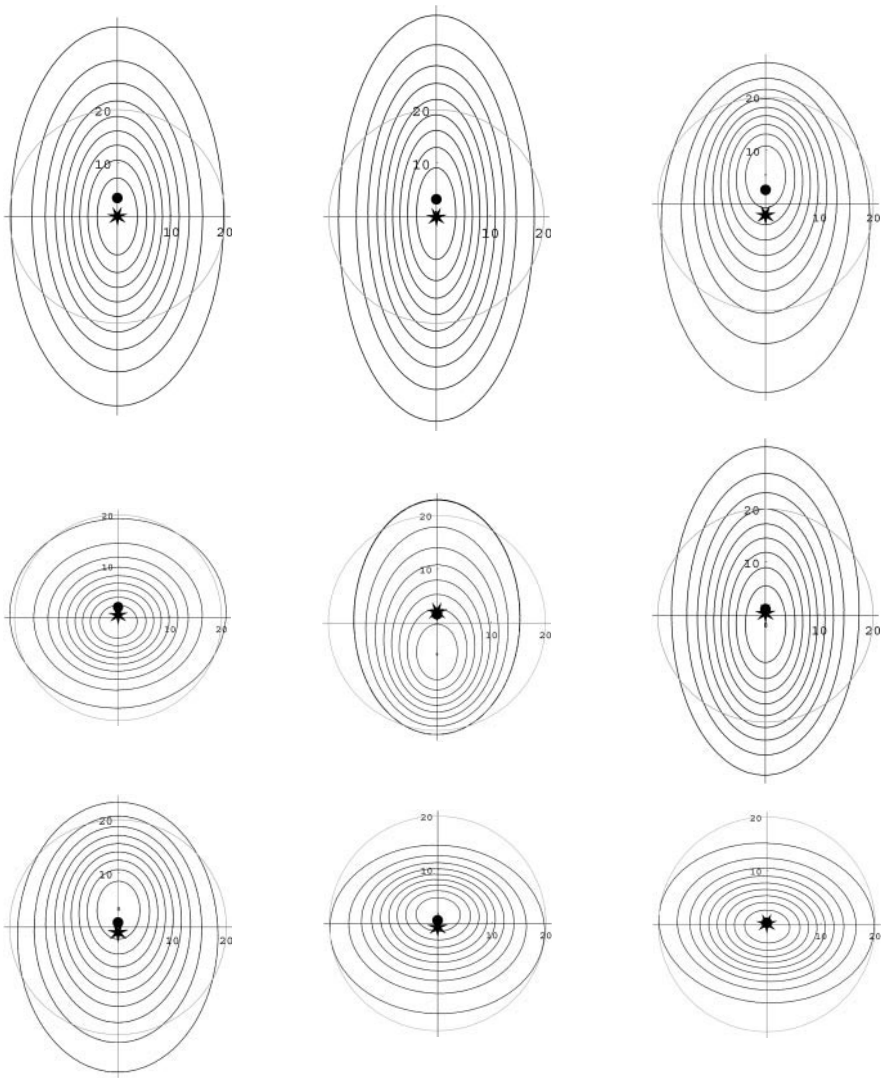


Figure 2. Local pdfs for model QC96-0 for theoretical sites at northern latitudes 0°, 10°, 20°, ... 80° (line by line, left to right, starting from upper left). The axes are labelled in degrees, the circles indicating a 20° angle with respect to the central (pure axial dipole) direction. The ellipses are pdf isovalues representing bounds within which 10, 20, ..., 90 per cent of the directions are expected to lie. The stars mark the theoretical limits of the local average directions and the black dots mark the directions produced by the mean field of QC96. Note that the angle for the theoretical limit can reach 1.2° in this case. Figs 2, 3 and 4 may be viewed in colour in the online version of the journal (www.blackwell-synergy.com).

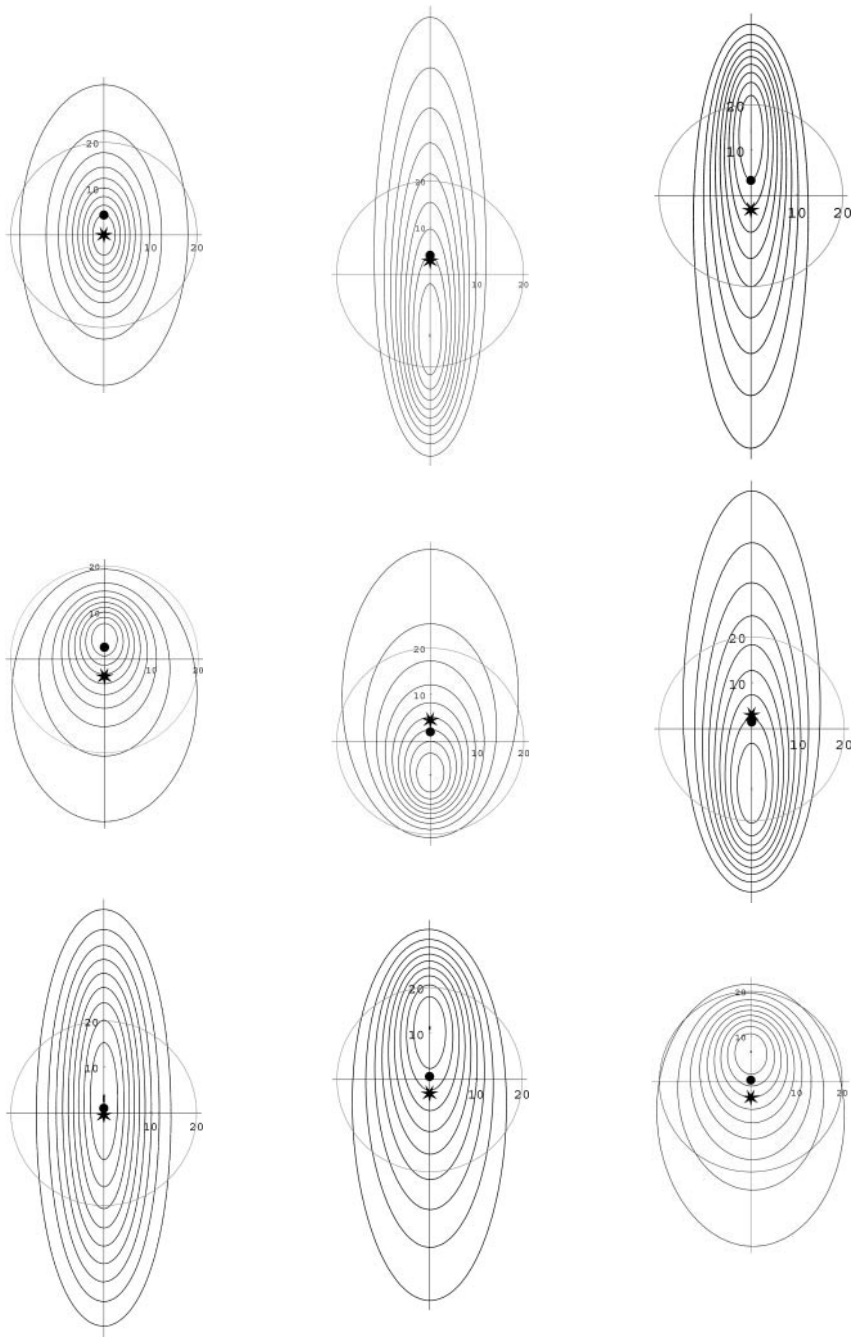


Figure 3. Same as Fig. 2 but for model CJ98-0 (pdf and stars) and mean field of CJ98 (dots). Note that the theoretical limit of the local average direction (stars) can now reach 5° (see central diagram).

on their longitudes. This is why longitudes are not specified in Figs 2 and 3. A similar symmetry holds with respect to the equator for the two models QC96-0 and CJ98-0. This is because the second-order moments for these GGPs are invariant with respect to this symmetry [the effect of this symmetry is only to change the sign of the Gauss coefficients with $(n-m)$ odd, and since no correlations are assumed, the second-order moments remain unchanged under the symmetry] and because the average field (an axial dipole) is itself antisymmetric with respect to this symmetry. Only sites with northern latitudes are shown in Figs 2 and 3. An additional symmetry properties holds locally, again as a direct consequence of the symmetries assumed within these two GGPs (which are axisymmetric, as already mentioned, and invariant under a longitudinal plane symmetry, as the reader can easily verify): for each site, the pdfs are clearly symmetric with respect to the local S–N axis (see Figs 2 and 3). Although we do not intend to discuss this point any further here, it is important to emphasise the fact that these symmetries are the direct consequence of the symmetries assumed within the GGPs. Any departure of the actual data distribution from these symmetries could thus provide very valuable indications that the field might behave in a non-symmetric manner [something which could then possibly be modelled in terms of anisotropy and/or correlations between Gauss coefficients, as first suggested by Hulot & Gallet (1996) and partially discussed by Constable & Johnson (1999)].

Despite the symmetries just mentioned, Figs 2 and 3 show that the local pdfs produced by the two models QC96-0 and CJ98-0 are non-Fisherian. They do not display the circular pattern typical of a Fisherian distribution. They do not even display a simple enough behaviour for it to possibly be approximated by more elaborate distributions such as the bivariate extension of the Fisher statistics (see e.g. Kent 1982; Fisher *et al.* 1987; Le Goff 1990). In fact, except for special situations (such as for sites located at or very near to the equator), the pdfs display a skewed behaviour. Isovalues of the pdfs display ellipse-like shapes but these ellipses are not centred on a common point. The effect is quite spectacular for most of the site latitudes, such as mid-latitudes, that are particularly relevant in practice (corresponding to latitudes of actual sampling sites). Not taking this effect into account may thus lead to a misleading assessment of the fit of a given GGP to the actual data distribution, even for models as simple as those considered here.

6.2 Consequences for mean-field modelling

The non-Fisherian behaviour of the local directional pdfs has important consequences for the possibility of recovering the mean-field model $\bar{\mathbf{k}}$ about which the GGP is assumed to fluctuate. Such mean-field modelling is usually performed with the help of averaged directional data, assuming these averages are unbiased estimators of the local direction of the mean field $\bar{\mathbf{k}}$ (see e.g. Gubbins & Kelly 1993; Johnson & Constable 1995; Johnson & Constable 1997; Kelly & Gubbins 1997; Carlut & Courtillot 1998). This assumption is compatible with the assumption of a Fisherian local directional pdf, that is, with the distribution these authors more or less explicitly assume for the data. Unfortunately, as we shall now explicitly show with the help of the two previous examples QC96-0 and CJ98-0, it is not compatible with the non-Fisherian local pdfs that arise as a consequence of a generalized GGP.

The theoretical limit of the local average direction that palaeomagnetists rely on (constructed by taking the direction of the sum of the observed unit vectors) is the direction of the average of the unit vectors weighted by the pdf. This average direction can easily be computed for each of the local pdfs shown in Figs 2 and 3 (stars). As can be seen (especially in Fig. 3), it usually does not correspond to the direction of the mean field $\bar{\mathbf{k}}$ (which is that of an axial dipole in the present instance and thus corresponds to the centre of the Lambert projection on each figure). The local average direction is thus a biased estimator of the direction produced by the mean field. The disagreement can be quite large, sometimes reaching several degrees. It is a direct consequence of the non-Fisherian behaviour of the local pdfs. (Note that this behaviour produces a most likely direction—central ellipse on each part of Figs 2 and 3—that also significantly differs from both the pdf average direction and the direction of the mean field).

Clearly, not taking this bias into account could lead one to misinterpret the departure of the reconstructed direction (with respect to the direction produced by the mean axial dipole) in terms of some structure within the mean field. This would not be a problem if the true structure could be expected to produce a signal significantly larger than the bias we see here. Unfortunately, this may not be the case. This again can be illustrated with the help of our examples. Consider the mean fields of the two original models QC96 and CJ98. As already mentioned, these mean fields are composed of an axial dipole and an axial quadrupole, the quadrupole term being considered as the single most significant non-axial dipole structure within the mean field of the past 5 Myr. Using just these mean fields and computing the direction they would produce at each site of Figs 2 and 3 (black dots) shows that these directions depart from the direction produced by a pure axial dipole by amounts that are comparable to the bias produced by the non-Fisherian pdfs of QC96-0 and especially CJ98-0.

At this point, it may be argued that the mean field and the bias do not always produce shifts in the same direction (as is the case for instance for northern latitudes of 20° and 30° in the case of model CJ98) and that sometimes the bias does not produce any effect whereas the mean field does (as is the case for equatorial sites) and vice versa (at high latitude). It may also correctly be argued that in the Southern Hemisphere the non-axial dipole mean field would produce an effect in the opposite direction to that produced in the Southern Hemisphere (because of the equatorial symmetry of a g_2^0 field). In this respect, we must firmly acknowledge that if sites over the entire Earth's surface were being used, it would not be possible to misinterpret the biases produced by the non-Fisherian pdfs considered here in terms of a mean g_2^0 component. This, in fact, is good news. It suggests that interpreting the worldwide inclination bias in terms of a persistent g_2^0, as done by most authors when relying on the best data set currently available, is a reasonable thing to do. However, if the sites are not so well distributed (as we shall see, this is the case for the data set considered here, which is strongly biased towards northern sites) or if one intends to look for more structure than the axial quadrupole component in the mean field, the non-Fisherian biases produced by a generalized GGP could clearly be misinterpreted in terms of some mean-field structure.

6.3 Examples of global and local statistical tests

That the issues raised in the previous sections should not be overlooked can also be seen directly in the course of testing GGP models in the way we advocate. In this last section, we provide explicit examples of the rigorous statistical testing method described in Section 5 and test the six GGP models CP88, QC96, CJ98 and CP88-0, QC96-0, CJ98-0 against the database previously described.

For each of these models, we followed the four steps described in Section 5, and tested the final uniformized data set $\{t_i\}$ over the [0, 1] segment with the help of two standard tools, the KS (Kolmogorov–Smirnov) test and the χ^2 test (see e.g. Press *et al.* 1996).

The KS test is an objective test that does not require external parameters. However, it has one well-known drawback: it is more sensitive to departures of the data $\{t_i\}$ from a uniform distribution at the middle of the segment [0, 1] than near its margins, hence we also decided to carry out a χ^2 test . This second test is less objective in the sense that some degree of arbitrariness is left to the user. However, this precisely provides us with the possibility of having a test more sensitive to departures near the margins of the distribution. Of course, many other tests could be carried out. However, as we shall see, these two 'complementary' tests already provide us with very useful results.

For the purpose of this illustrative study we used a χ^2 test with five degrees of freedom, based on the division of the [0, 1] segment into six boxes, namely, (0, 0.1), (0.1, 0.25), (0.25, 0.5),

(0.5, 0.75), (0.75, 0.9), (0.9, 1). This choice puts more emphasis on the (0, 0.1) and (0.9, 1) boxes, while still ensuring that enough data values will eventually fall in each box for the test to be valid.

In practice, for each model to be tested and for each site q with data available, we computed the local directional pdf on a numerical grid defined on the local unit sphere Σ (steps 1 and 2). This then made it possible to define numerically the isovalues for $[f_{\mathscr{G}\mathscr{D}}]_q(\theta, \varphi)$ on the unite sphere, and to carry out the numerical integration involved in eq. (14). Thus step 3 (uniformization $\{s_i\}$ into $\{t_i\}$) was fulfilled.

At this stage, a given model could then easily be tested locally by carrying out both the KS and χ^2 tests with the corresponding single set of uniformized data $\{t_i\}_q$. Alternatively, the same model could also be tested regionally by stacking $\{t_i\}_q$ for various sites q before carrying out the tests. Finally, global tests could also be carried out by stacking all the $\{t_i\}_q$ (step 4). However, for conciseness and given the illustrative purpose of the present section, only a few local tests and the global tests for each of the six models have actually been carried out.

Figs 4 and 5 provide illustrations of the local test carried out for the six models at the French site. This site has been chosen because it lies in a band of latitude that is relatively common within our data set (see Fig. 1) and because the number of data points is also representative of the available data at a reasonably good site. As can be seen in Fig. 4, the general aspect of the local pdfs predicted by the six models is fairly dependent on the variances that are assumed, but apparently

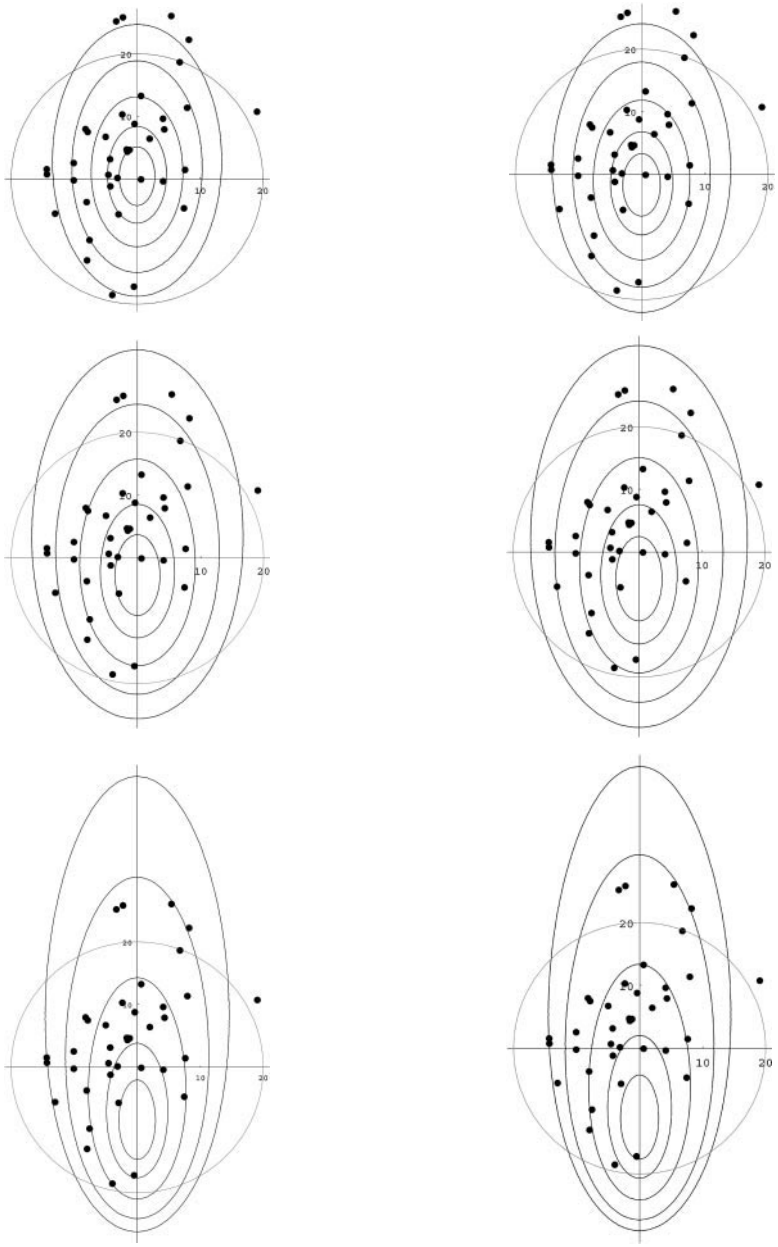


Figure 4. Local pdfs (together with the data) for the models tested at the French site: CP88 (top left) and CP88-0 (top right), QC96 (middle left) and QC96-0 (middle right), CJ98 (bottom left) and CJ98-0 (bottom right). Ellipses are pdf isovalues representing bounds within which 10, 25, 50, 75 and 90 per cent of the directions are expected to lie.

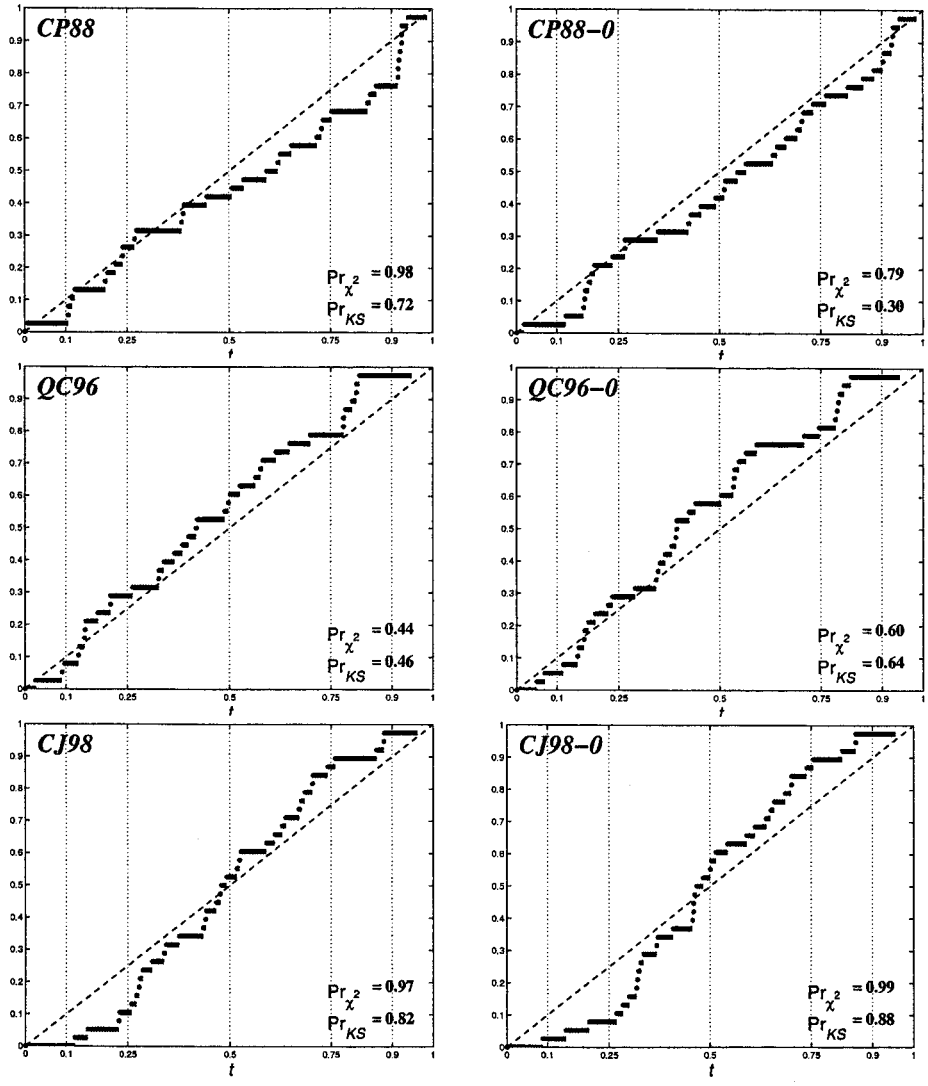


Figure 5. Empirical distributions of the uniformized data for the various models tested at the French site. Also shown are the relevant values of the probabilities for the χ^2 and KS tests (taken from Table 2).

insensitive to the weak mean model people assume is required to explain the data. These figures already give a hint of how good each model may be with respect to the site considered. Indeed, isovalues of the local pdf shown in these figures represent the bounds within which 10, 25, 50, 75 and 90 per cent of the directions are predicted to lie, given the model under consideration. As the reader can easily verify, this means that the χ^2 test can be implemented in a straightforward manner by just counting the number of points falling between two successive 'ellipses'.

Alternatively, and also to implement the KS test properly, we may further plot the six empirical cumulative distribution functions (cdfs) obtained after carrying out 1-D uniformization. This is shown in Fig. 5, where the theoretical cdf of a uniform distribution is also plotted. Then, implementing the KS test simply consists of looking for the maximum vertical discrepancy between the empirical and theoretical cdfs.

In the present case, both tests (χ^2 and KS) lead to similar conclusions. However, it is important to note that this is not always exactly the case. This, in fact, is a direct illustration of the complementary nature of the two tests. Take one example for which the discrepancy is especially striking: CJ98-0. For this example the probability of rejecting the model is 99 per cent from the point of view of the χ^2 test, and only 88 per cent from the point of view of the KS test. This is because the KS test fails to see the systematic departure of the empirical cdf with respect to the linear theoretical cdf. By contrast, the χ^2 test sees this as a very strong lack of observed values in the first (0, 0.1) and (0.1, 0.25) boxes, together with a symmetric strong excess of values in the last (0.75, 0.9) and (0.9, 1) boxes. In this respect the model is thus indeed very unlikely to be compatible with the data. Clearly, this also means that in testing models, one should always consider the most unfavourable probability.

This being said, we can now state that our results show that some models are consistent with the data at a good level of confidence (QC96, QC96-0, CP88-0), while others (CP88, CJ98, CJ98-0) are clearly not (Figs 4 and 5 and Table 2). It is particularly striking that removing the axial quadrupole from the mean field usually leads to a slightly worse fit but not always. CP88-0, for instance, leads to a better fit than CP88. This, in fact, is a direct illustration of the fact mentioned in the previous section that the (slightly, in the present instance)

Table 2. Local and global results of the χ^2 and Kolmogorov–Smirnov tests. We show the pairs of associated probabilities using the χ^2 (left value) and Kolmogorov–Smirnov tests (right value) for the model to be incompatible with the data.

	N	CP88	QC96	CJ98	CP88-0	QC96-0	CJ98-0
France	38	0.98, 0.72	0.44, 0.46	0.97, 0.82	0.79, 0.30	0.60, 0.64	0.99, 0.88
Hawaii	112	0.98, 1.00	0.79, 0.76	1.00, 1.00	1.00, 1.00	0.42, 0.66	1.00, 1.00
Reunion	44	0.26, 0.50	0.03, 0.00	0.80, 0.28	0.71, 0.90	0.13, 0.10	0.85, 0.71
All 26	944	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00	1.00, 1.00

non-Fisherian properties of the local pdf may be enough to account for the bias in direction usually interpreted in terms of a persistent mean-field component. Many other local tests have been carried out, which we do not describe in detail (see Table 2 for typical results).

The main fact to come out of these additional tests is that the quality of the fit of each model to the data is strongly site-dependent. These somewhat contradictory results simply testify to the fact that none of the six models we considered manages to account globally for the whole data set in a satisfying way. There seems always to exist at least a few sites where any given model fails to account for the data.

Turning to the global test gives us a global statistical measure of this difficulty. For such a task no figure similar to Fig. 4 can be shown. However, we may of course draw the equivalent of Fig. 5, that is, the empirical cdfs obtained after carrying out 1-D uniformization and binning all uniformized data sets into a single global data set to be tested against the uniform distribution. This is shown in Fig. 6. Both the χ^2 and KS tests can then be carried out.

As can be seen in Table 2, it turns out that indeed none of the six models scores well enough in the tests for it to be compatible with the full data set. All models are thus strictly incompatible with this data set, at least within the set of assumptions we decided to rely on (we will return to this important point). The best we can say is that some models are apparently not as bad as others. This can be seen from both Fig. 6 and Table 3, which, rather than probabilities, shows the observed values χ^2 and D for the χ^2 and KS tests (for more technical details see Press *et al.* 1996).

Focusing first on the three published models CP88, QC96 and CJ98, we note that these three models obtain very similar scores. The KS test suggests that QC96 is slightly worse, but the χ^2 test seems to give a slightly more discriminating result. It is interesting to note that the relative χ^2 scores lead to the conclusion that authors are progressively coming out with models better and better corrected for drawbacks identified in earlier models. That CP88 is the worst of these three models is no big surprise. As previously mentioned, this model is among the simplest GGP models one can create but belongs to a family of models that are well known to be seriously incompatible with a variety of effects observed in data-derived indicators (see e.g. Kono & Tanaka 1995; Hulot & Gallet 1996). That model QC96 does 'slightly better' is also quite sensible. This model was indeed specifically designed to correct for the drawbacks observed in CP88 (Quidelleur & Courtillot 1996). The same of course also applies to the CJ98 model (Constable & Johnson 1999), which is the best with respect to both tests.

Far less expected in view of earlier studies is the fact that CJ98-0 does so well. It is actually the best model we tested (i.e. the 'least bad'). This result shows that a model with a pure axial dipole mean field can provide a better fit then a model with more mean-field features. Note that this also questions the need for a persistent axial quadrupole field, within the context of the data set and assumptions considered here. Non-Fisherian behaviour of the local pdfs for this model appears to account for the present data set at a level that is not good (remember that none of the models is 'satisfactory'), but which is nevertheless better than any of the models considered here. This is because the sites we considered are unevenly distributed and often lie at latitudes such that the non-Fisherian behaviour actually produces biases in a favourable direction (in the way described in the previous section). Models CP88-0 and QC96-0 do not lead to a similar improvement with respect to the CP88 and QC96 models. Setting the axial quadrupole mean field to zero in that case in fact leads to a dramatic deterioration of the models (Fig. 6 and Table 3). This is the perfect complement of the previous result. Indeed, for these models the assumed variances are such that the non-Fisherian pdfs produce much weaker biases (see Fig. 2). In that case the g_2^0 is needed to account better for the data. These results again illustrate the need to take the non-Fisherian behaviour of local pdfs into account properly before making claims concerning the mean field. They also show that mean-field and PSV modelling cannot be carried out independently.

Finally, it should be stressed that comparing models that so poorly account for the data is anyway not satisfactory. This finally brings us back to what seems to be the main conclusion of this set of rigorous local and global statistical tests: none of the considered models appears to be compatible *with the full data set we considered within the set of assumptions made*. This last point is important because two strong assumptions concerning the data set have also been made: first, each data

Table 3. Global results of the χ^2 and Kolmogorov–Smirnov tests. We show the pairs of input statistics for the χ^2 test (left value) and the Kolmogorov–Smirnov test (right value).

	CP88	QC96	CJ98	CP88-0	QC96-0	CJ98-0
All 26	39.0, 0.06	25.4, 0.08	19.5, 0.06	841.5, 0.29	1658., 0.40	18.5, 0.06

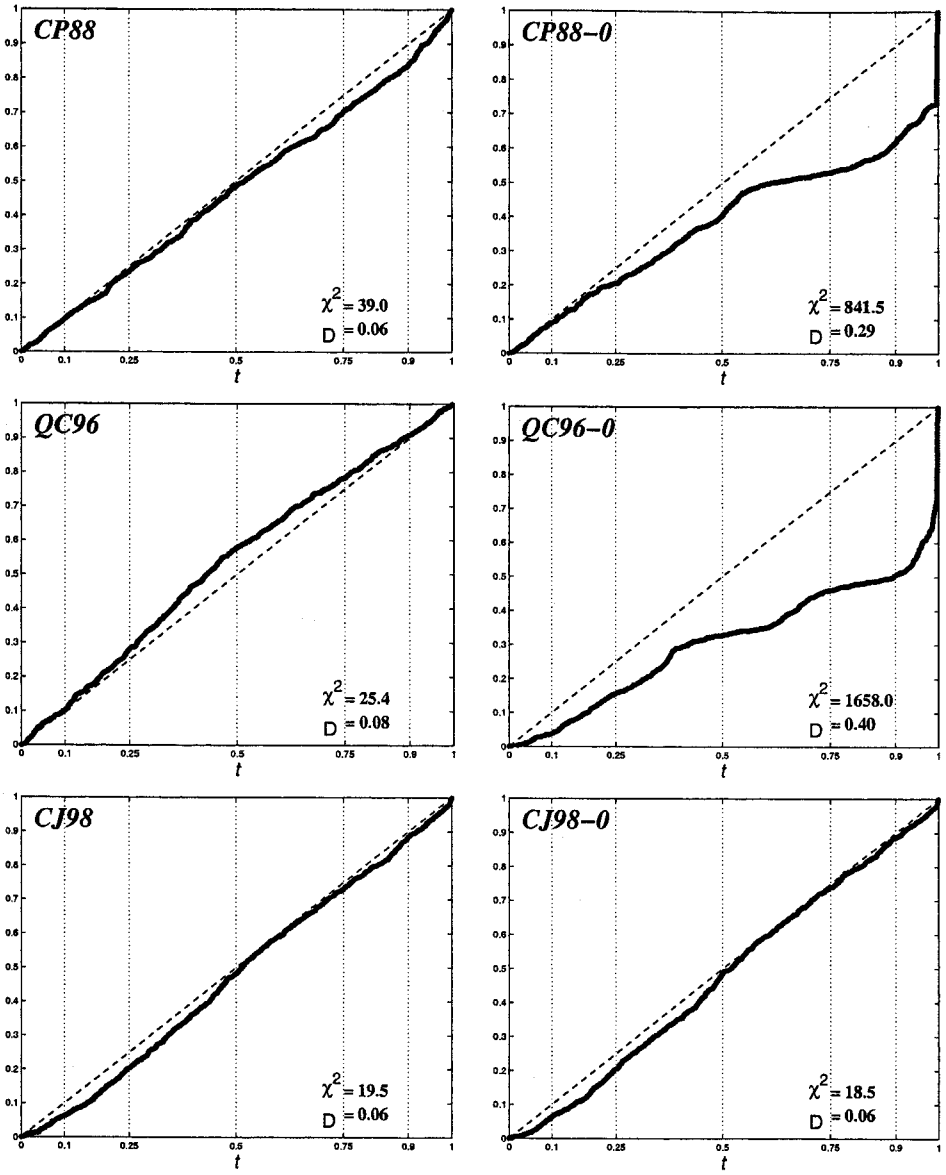


Figure 6. Empirical distributions of the uniformized global data for the various models. Also shown are the input values for the χ^2 and KS tests (taken from Table 3).

point is statistically independent of any other data point (because of the time elapsed between the setting of the corresponding lava flows), and second, the data are infinitely accurate. Three reasons might thus explain the failure of the models: the models are inadequate, or the data are sometimes correlated, or the errors in the data are so large that we need to take them into account. All three reasons are likely to combine to produce the overall negative result found. The problem of data correlation could be addressed within the framework of our method. Indeed, such a correlation would *a priori* be the result of lava samples taken from different flows (on a common site) that are erroneously thought to be non-contemporaneous. The problem would then be specific to this site and could be detected by the failure of the local test for virtually all models. The problem of noise in the data is less obvious. It would require that the rigorous statistical approach presented here be slightly generalized to deal with imperfect data. Such a generalization, which is beyond the scope of the present paper, would improve the scores of all models in the tests, but would be unlikely to modify the general ordering of the models from best to worst (unless, as in the case of CJ98 and CJ98-0, the models obtain very close scores). The best models considered here (or better models) would then hopefully pass the tests in a truly satisfactory way.

7 CONCLUSIONS

The purpose of the present paper was to provide a fully consistent way of addressing the problem of modelling the palaeomagnetic field in terms of the now quite well accepted (and here slightly generalized) concept of GGP. We showed that for any given generalized GGP model, it was possible to compute the exact pdf not only for the full field, but also for the direction of the field at any given site at the Earth's surface. We explained how by forwarding in such a way the statistical information of a model to the local sites, it was possible to

make local assessments of the validity of any given model. We also explained how such local statistical information could be binned together to test the regional or worldwide validity of such a GGP model.

The approach we propose has several advantages over those that have been used in the recent literature. First, it is fully consistent with the conceptual framework used to define a GGP model. Second, it only involves a statistically rigorous transformation of the data, and the data may thus be tested in their raw form (avoiding uncontrolled propagation of errors in the data-derived indicators most authors rely on). Third, it requires no approximation of any kind. Fourth, it may be used with any kind of 'instantaneous' palaeomagnetic data.

In the process of explaining our method, one particularly important result was derived that we again wish to stress: local directional pdfs produced by generalized GGP models are usually significantly non-Fisherian. As a consequence, even very simple models only involving a pure axial dipole mean field can produce local directions whose average (defined by taking the direction of the average of the unit vectors, as is common practice for mean-field modelling) may differ from the direction of the true mean field. The bias introduced has been shown to be possibly comparable to the observed shift in average direction, which is usually interpreted in terms of some mean-field structure. This means that the standard procedure most palaeomagnetists rely on to produce mean-field models from directional data (especially when starting from instantaneous lava samples) is inconsistent with the generalized GGP concept. Models produced in this way may not reflect the true mean field of the best generalized GGP model describing the data. In fact, these models are likely to be seriously affected by the non-Fisherian biases of the local directional pdfs, the amplitude of which is governed by the secular variation of the field (that is, the variances of the GGP model) and not by the true mean field. As has been shown, this would be the case if, for instance, the best generalized GGP model describing the data had variances similar to those involved in the CJ98 model. Great caution should therefore be taken in interpreting such mean-field models.

Unfortunately, what our results also show is that no good (i.e. fully consistent) procedure to recover or test a mean-field model starting from pure directional data without simultaneously considering the variances describing the GGP model seems to be readily available. Some progress might be possible, but this would require introducing some justified approximation. This is a possibility that is worth pursuing and that we intend to consider in detail in future work.

In the mean time, the best we can do is to test the full GGP model against a (partly or fully) directional data set by relying on the method we propose here, possibly using this forward method as a tool for a trial-and-error approach or a Monte Carlo approach to recover both the mean field and the variances characterizing the GGP model. The very complete and carefully checked database of McElhinny & McFadden (1997) could, for instance, be used to carry out a much more systematic study than the one presented here for illustrative purposes. Alternatively, one could also decide to focus mainly on the much smaller data set composed of full magnetic data (direction and intensity), in the spirit of the recent study of Kono *et al.* (2000). In that case, one can take advantage of the fact that the statistics of the local field produced by a GGP model are those of a pure 3-D Gaussian distribution (see Section 3). The observed mean local field then being an unbiased estimate of the signal the mean field would produce at the site (see eq. 4), these observed means could be used in a consistent way to recover the mean-field model independently of the variances involved in the GGP model [the modelling procedure is then essentially the same as the one used for producing modern geomagnetic field models; see e.g. Langel (1987)]. As noted by Kono *et al.* (2000), such a data set could also provide some useful information concerning the variances of the GGP model, independently of the value of the mean model. This is because of eq. (5), which linearly relates observed local covariances to the model covariances. Unfortunately, as described by Kono *et al.* (2000), the inverse problem defined by eq. (5) is far less stable then the classic inverse problem defined by eq. (4). It thus eventually turns out that little more than a modest (but consistent) mean-field model can safely be recovered through mathematical inversion. Such a mean field should, however, be considered as the best starting point for constructing consistent generalized GGP models of the real palaeomagnetic field.

ACKNOWLEDGMENTS

We thank V. Courtillot for fruitful discussions and C. Constable for a very helpful review. This work was partly supported by grant INTAS/CNES 97-1048. INSU 'Interieur de la Terre' programme contribution No. 254 and IPGP contribution No. 1724.

REFERENCES

Bingham, C., 1983. A series expansion for angular Gaussian distribution, Appendix C, in *Statistics on Spheres,* ed. Watson, G., pp. 226–231, Wiley-Interscience, New York.

Bloxham, J. & Jackson, A., 1992. Time-dependent mapping of the magnetic field at the core-mantle boundary, *J. Geophys. Res.,* **97,** 19 537–19 563.

Carlut, J. & Courtillot, V., 1998. How complex is the Earth's average magnetic field?, *Geophys. J. Int.,* **134,** 527–544.

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

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

Eckhardt, D.H., 1984. Correlations between global features of terrestrial fields, *Math. Geol.,* **16,** 155–171.

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

Gubbins, D. & Kelly, P., 1993. Persistent patterns in the geomagnetic field over the past 2.5 Myr, *Nature,* **365,** 829–832.

Hongre, L., Hulot, G. & Khokhlov, A., 1998. An analysis of the geomagnetic field over the past 2000 years, *Phys. Earth planet. Inter.,* **106,** 311–315.

Hulot, G. & Gallet, Y., 1996. On the interpretation of virtual geomagnetic pole (VGP) scatter curves, *Phys. Earth planet. Inter.,* **95,** 37–53.

Hulot, G. & Le Mouël, J.-L., 1994. A statistical approach to the Earth's main magnetic field, *Phys. Earth planet. Inter.,* **82,** 167–183.

Johnson, C. & Constable, C., 1995. The time-averaged field as recorded by lava flows over the past 5 Myr, *Geophys. J. Int.,* **122,** 489–519.

Johnson, C. & Constable, C., 1996. Paleosecular variation recorded by lava flows over the last 5 Myr, *Phil. Trans. R. Soc. Lond.,* **A354,** 89–141.

Johnson, C. & Constable, C., 1997. The time-averaged geomagnetic field: global and regional biases for 0–5 Ma, *Geophys. J. Int.,* **131,** 643–666.

Kelly, P. & Gubbins, D., 1997. The geomagnetic field over the past 5 million years, *Geophys. J. Int.,* **128,** 1–16.

Kent, J.T., 1982. The Fisher-Bingham distribution on the sphere, *J. R. stat. Soc.,* **B44,** 71–80.

Knuth, D., 1981. *The Art of Computer Programming,* 2nd edn, Addison-Wesley, Reading, MA.

Kono, M., 1997. Paleosecular variation in field directions due to randomly varying Gauss coefficients, *J. Geomag. Geoelectr.,* **49,** 615–631.

Kono, M. & Tanaka, H., 1995. Mapping the Gauss coefficients to the pole and the models of paleosecular variation, *J. Geomag. Geoelectr.,* **47,** 115–130.

Kono, M., Tanaka, H. & Tsunakawa, H., 2000. Spherical harmonic analysis of paleomagnetic data, the case of linear mapping, *J. geophys. Res.,* **105,** 5817–5833.

Kullback, S., 1968. *Information Theory and Statistics,* Dover, New York.

Langel, R.A., 1987. The main field, in *Geomagnetism,* Vol. 1, pp. 249–512, ed. Jacobs, J.A., Academic Press, London.

Lee, S., 1983. A study of the time-averaged paleomagnetic field for the last 195 million years, *PhD thesis,* Australian National University, Canberra.

Le Goff, M., 1990. Lissage et limites d'incertitude des courbes de migration polaire—pondération des données extension bivariate de la statistique de Fisher, *C. R. Acad. Sci. Paris,* **3111,** 1191–1198.

McElhinny, M. & McFadden, P., 1997. Paleosecular variation over the past 5 Myr based on a new generalized database, *Geophys. J. Int.,* **131,** 240–252.

McElhinny, M., McFadden, P. & Merrill, R., 1996. The time-averaged paleomagnetic field 0-5My the Earth, *J. geophys. Res.,* **101,** 25 007–25 027.

Merrill, R., McElhinny, M. & McFadden, P., 1996. *The Magnetic Field of the Earth,* Academic Press, London.

Press, C., Teukolsky, S., Vetterling, W. & Flannery, B., 1996. *Numerical Recipes in C,* 2nd edn, Cambridge University Press, Cambridge.

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.,* **15,** 1639–1642.

Schneider, S. & Kent, D., 1990. The time-averaged field, *Rev. Geophys.,* **28,** 71–96.

Wilson, R.L., 1970. Permanent aspects of the Earth's non-dipole magnetic field over upper Tertiary times, *Geophys. J. R. astr. Soc.,* **19,** 417–437.

Wilson, R.L., 1971. Dipole offset: the time-averaged paleomagnetic field over the past 25 million years, *Geophys. J. R. astr. Soc.,* **22,** 491–504.