



**HAL**  
open science

## Hierarchical modelling of archaeomagnetic data and curve estimation by moving average technique

Philippe Lanos, Maxime Le Goff, Mary Kovacheva, Elisabeth Schnepf

► **To cite this version:**

Philippe Lanos, Maxime Le Goff, Mary Kovacheva, Elisabeth Schnepf. Hierarchical modelling of archaeomagnetic data and curve estimation by moving average technique. *Geophysical Journal International*, 2005, 160 (2), pp.440-476. 10.1111/j.1365-246X.2005.02490.x . halshs-00396776

**HAL Id: halshs-00396776**

**<https://shs.hal.science/halshs-00396776>**

Submitted on 30 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hierarchical modelling of archaeomagnetic data and curve estimation by moving average technique

Philippe Lanos,<sup>1</sup> Maxime Le Goff,<sup>2</sup> Mary Kovacheva<sup>3</sup> and Elisabeth Schnepf<sup>4,\*</sup>

<sup>1</sup>Laboratoire d'Archéomagnétisme, UMR 6566 and UMR 6118, CNRS, Université de Rennes 1, campus de Beaulieu, 35042 Rennes, France.

E-mail: philippe.lanos@univ-rennes1.fr

<sup>2</sup>Laboratoire de Géomagnétisme et Paléomagnétisme, UMR 7577, CNRS, Institut de Physique du Globe de Paris, 94100 Saint-Maur, France

<sup>3</sup>Palaeomagnetic Laboratory, Geophysical Institute, 1113 Sofia, Bulgaria

<sup>4</sup>GGA-Institut, Arbeitsbereich Grubenhagen, OT Rotenkirchen, 37574 Einbeck, Germany

Accepted 2004 September 21. Received 2004 July 30; in original form 2002 August 5

## SUMMARY

A Bayesian hierarchical modelling is proposed for the different sources of scatter occurring in archaeomagnetism, which follows the natural hierarchical sampling process implemented by laboratories in field. A comparison is made with the stratified statistics commonly used up to now. The Bayesian statistics corrects the disturbance resulting from the variability in the number of specimens taken from each sample or site. There is no need to publish results at sample level if a descending hierarchy is verified. In this case, often verified by archaeomagnetic data, only results at site level are useful for geomagnetic reference curve building. Typically, a study with at least 20 samples will give an  $\alpha_{95i}$  5 per cent close to the optimal  $\alpha_{95i}$  for a fixed site number  $m_i$  and if errors are random with zero mean (no systematic errors). The precision on the curve itself is essentially controlled, through hierarchical elliptic statistics, by the number of reference points per window and by dating errors, rather than by the confidence angles  $\alpha_{95ij}$  at site level (if a descending hierarchy). The Bayesian elliptic distribution proposed reveals the influence of the window width. The moving average technique is well adapted to numerous and very well dated data evenly distributed along time. It is not a global functional approach, but a (linear) local one.

**Key words:** archaeomagnetism, geomagnetic secular variation, remanent magnetization, statistical methods.

## 1 INTRODUCTION

Archaeomagnetic dating utilizes the property of some materials to record information about the Earth's magnetic field (EMF). This fossilized magnetic information, remanent magnetism, is acquired, for example, by heating materials to high temperatures. Typically, archaeomagnetic data relate to the direction (inclination,  $I$ , and declination,  $D$ ) and intensity (denoted  $F$  for magnetic field) and so a single archaeomagnetic record is three-dimensional. These three components of the EMF vary according to geographical location (latitude and longitude) and through time.

It is now well established that if sufficient archaeological sites in a given area of a given archaeological period are available, then local secular variation curves can be built for each of the three EMF variables. If these curves are known accurately, it is possible to date in the same geographical region other sites in which magnetization has been measured.

The knowledge of the variations of the geomagnetic field in the archaeological past has increased in recent years by new data sets. A revised long Bulgarian data set, comprising both directional and intensity archaeomagnetic determinations (Kovacheva 1997), enabled the geomagnetic field in Bulgaria to be estimated over the last eight millennia, but with some gaps, e.g. at approximately 2000 BC and 3500 BC. Daly and Le Goff's (1996) analysis of palaeosecular directional variations was restricted to the last 2000 yr, for which most data were available. Recent work on the direction variations have been made in Germany (respectively, Schnepf *et al.* 2003, 2004), in Hungary (Márton 2003), in France over the first millennium BC (Gallet *et al.* 2002), in Belgium (Hus & Geeraerts 1998), in USA (Labelle & Eighmy 1997), and also on the intensity variations in France over the past 2000 yr (Chauvin *et al.* 2000; Genevey & Gallet 2002) and in the Eastern Mediterranean over the past 8000 yr (Genevey *et al.* 2003).

\*Now at: Montanuniversität Leoben, Paleomagnetic Laboratory Gams, Gams 45, A-8146 Frohnleiten, Austria.

The biggest disadvantage of archaeomagnetic data is that they are generally unevenly distributed in time and space. It is difficult to obtain a long sequence of numerous, well-dated data for a given region because it is reliant on the discovery of suitably dated archaeological remains. In addition, the determination of the past field directions is not always accompanied by its intensity and vice versa. Moreover, different sampling and measurement procedures have been adopted by different laboratories, resulting in different measurement errors. However, the fundamental errors relate to the reliability of the dates of the reference sites.

The main aim of this paper is to model the sampling process using a Bayesian hierarchical framework, in both the bivariate and univariate case, i.e. intensity, and to evaluate the statistical characteristics of the moving average technique. The second section of this paper consists of a detailed analysis of the different sources of scatter occurring at the different levels of the sampling and measurement process. In Section 3, the classical statistics used and their application within the stratification framework are considered. Then, a new statistical approach is proposed, based on the Bayesian hierarchical statistics, which is applied to directional as well as intensity data. The hierarchical statistics and parameter estimation are presented in Section 4 and the consequences for the sampling strategy are discussed. Plotting the curves, in the context of the moving average technique, and drawing confidence intervals are described in Section 5. Some numerical examples and applications to French and Bulgarian databases for the last 2000 yr are shown in Section 6.

## 2 SAMPLING METHODS AND ANALYSIS OF THE ERRORS ENCOUNTERED

### 2.1 Sampling in field

Generally, the ancient magnetic field is estimated at a given place and time from a number of contemporaneous archaeological baked clay sites (kilns, ovens . . . , called structures in archaeology). Different sampling strategies can be distinguished as follows.

- (i) Between 10 and 30 independently oriented samples *in situ* (so-called blocks) are taken. The sample size varies in general between 300 and 2000 cm<sup>3</sup>. The measurements are made in the large inductometers (Paris in France, Dourbes in Belgium).
- (ii) The samples can be cut into one or several cubic or cylindrical specimens that are measured in cryogenic, spinner or astatic magnetometers (Rennes in France, Sofia in Bulgaria).
- (iii) 6 to 12 small oriented samples (ca. 1–20 cm<sup>3</sup>) are taken, that are not subsampled (cut into specimens). Such small samples can also be obtained from drilling *in situ*, using the palaeomagnetic technique. This sampling strategy characterizes US, Canadian as well as UK methods (UK, Tarling 1983; USA, Eighmy 1990). In this case, specimen and sample are synonymous.

The quality of each strategy depends on the number but also on the size of the samples and/or specimens because of magnetic heterogeneities.

### 2.2 Hierarchical sampling levels and associated errors

The hierarchical sampling process is naturally implemented in archaeomagnetism as well as in palaeomagnetism. The common idea is to realize a sampling strategy that takes into account the heterogeneity of the magnetisation at the different levels for a given time. Thus, different sampling levels called specimens, samples and sites (Tarling 1983) can be identified. This equates to the concept of hierarchy in Bayesian statistics (Barnett 1982; Lindley 1990; Howson & Urbach 1993; Drosbeke *et al.* 2002). Adding the measurement, field and curve levels defines six hierarchical levels, which relate to the unknown geomagnetic curve. Each level is denoted by an index, that is: *m* for measurement, *l* for specimen, *k* for sample, *j* for site, *i* for field (at a given time), but there is no index for curve (Fig. 1 and Table 1).

#### (i) Measurement hierarchical level (*ijklm*)

A single value  $Y_{ijklm}$  of the studied magnetic parameter (inclination, declination, intensity), is determined at a given step of a demagnetising process or an intensity experiment.

#### (ii) Specimen hierarchical level (*ijkl*)

A specimen is subsampled from an independently oriented sample, i.e. by cutting or drilling. Consequently, specimens have a common system of field marks. For a particular specimen, characterized by the index *ijkl*, the measured value  $Y_{ijklm}$  is assumed to belong to a probability distribution with unknown mean  $Y_{ijkl}$  and unknown variance  $\sigma_{ijkl}^2$ . This variance results from the different sources of scatter as follows.

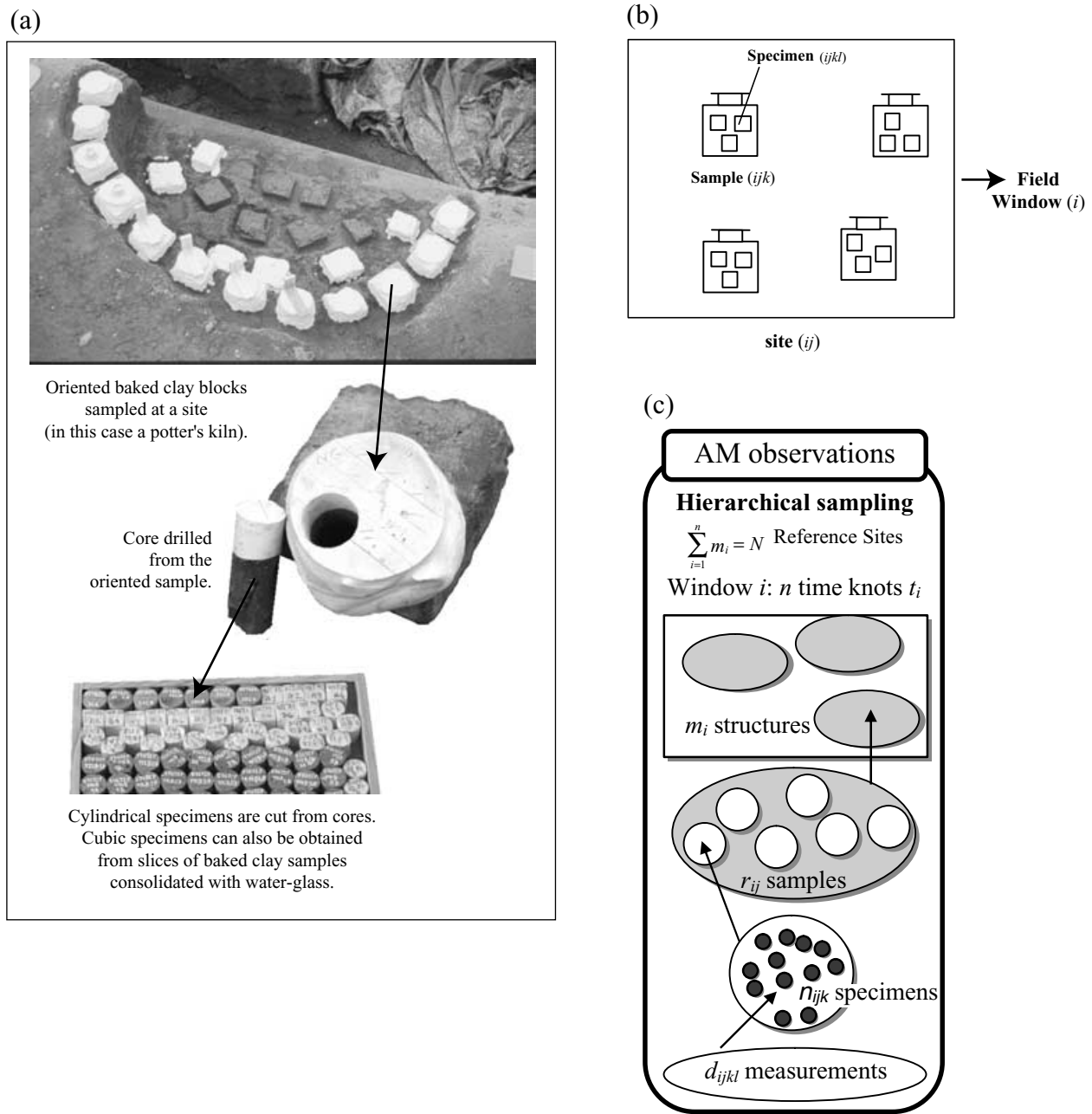
E1: errors in the measurement process (measurement noise, problems of positioning in magnetometers and demagnetisers, etc.).

E2: errors in physical and statistical interpretation (separation of the TRMs after demagnetisation, determination of the slope in Arai diagrams . . .).

These errors can be assumed randomly (Gaussian) distributed. No systematic error is expected (unless a calibration problem occurs). The realization of  $d_{ijkl}$  measurements (or determinations) allows an empirical mean  $\bar{Y}_{ijkl}$  and an empirical variance  $S_{ijkl}^2$  to be determined for the specimen.

#### (iii) Sample hierarchical level (*ijk*)

A sample is a (small or large) block of burnt clay taken from a site, or a fragment of brick or pottery in case of displaced materials (Lanos *et al.* 1999). It describes a material: a piece of clay taken following the precise conditions of orientation or a core drilled directly from the site



**Figure 1.** (a) Example of archaeomagnetic sampling in field (roman potter's Kiln, Angers, Saint-Laud station, France, M. Mortreau, INRAP, 2000), and examples of sample and specimens, (b) Sampling levels scheme used by archaeomagnetic laboratories, (c) hierarchical modelling of error sources at five levels: measurement ( $ijklm$ ), specimen ( $ijkl$ ), sample ( $ijk$ ), site or structure ( $ij$ ) and field (window  $i$ ).

(as usual with palaeomagnetic sampling). It possesses only one orientation mark generally obtained using magnetic and sun compasses. In the coring case, for very short cores, or in the English technique, specimen is equivalent to sample and the two levels are merged.

The unknown mean  $Y_{ijkl}$  at specimen level is assumed to belong to a probability distribution having an unknown mean value  $Y_{ijk}$  and an unknown variance  $\sigma_{ijk}^2$ . This variance between the specimens from the same sample results from different sources of scatter as follows.

E3: cutting errors (in the case of subsampling), transfer errors of reference field marks.

E4: errors of heterogeneities resulting from differences in magnetic minerals concentration or from mineralogical transformations.

E5: errors of magnetic anisotropy, demagnetising fields (Lanos 1987), or cooling rate (Chauvin *et al.* 2000).

If the specimen is the sample, then the source of scatter, E3, does not occur. These errors can be assumed randomly distributed, except for error E5, which is generally systematic. To avoid the transfer of this systematic error to an upper level, corrections have to be performed.

The observation of  $n_{ijk}$  specimens allows an empirical mean  $\bar{Y}_{ijk}$  and an empirical variance  $S_{ijk}^2$  to be determined for the sample.

**Table 1.** Variables used according to the sampling levels defined in the hierarchical modelling.

Sampling level	Sources of scatter	Unknown variates			Observations		Squared errors
		Mean	Variance	Number	Empirical mean	Empirical variance	
Measurement $ijklm$					$Y_{ijklm}$		
Specimen $ijkl$	E1, E2	$Y_{ijkl}$	$\sigma_{ijkl}^2$	$d_{ijkl}$ measurements	$\overline{Y}_{ijkl}$	$S_{ijkl}^2$	$e_{ijkl}^2 = \frac{\sigma_{ijkl}^2}{d_{ijkl}}$
Sample $ijk$	E3, E4, E5	$Y_{ijk}$	$\sigma_{ijk}^2$	$n_{ijk}$ specimens	$\overline{Y}_{ijk}$	$S_{ijk}^2$	$\frac{1}{U_{ijk}} = \left( \sum_{l=1}^{n_{ijk}} \frac{1}{\sigma_{ijk}^2 + e_{ijkl}^2} \right)^{-1}$
Site (structure) $ij$	E4, E5, E6, E7	$Y_{ij}$	$\sigma_{ij}^2$	$r_{ij}$ samples	$\overline{Y}_{ij}$	$S_{ij}^2$	$\frac{1}{V_{ij}} = \left( \sum_{k=1}^{r_{ij}} \frac{1}{\sigma_{ij}^2 + \frac{1}{U_{ijk}}} \right)^{-1}$
Field (window) $i$	E8, E9, E10, E11	$t_{ij}$ $Y_i = g(t_i)$	$\sigma_{ii}^2$ $\sigma_i^2$	$m_i$ sites weighted by $P_{ij}$	$\overline{Y}_i$	$S_i^2$	Bivariate case: $\Sigma_i$ (eq. 43) Univariate case (eq. A3.6): $\frac{1}{W_i} = \left( \sum_{j=1}^{m_i} \frac{P_{ij}}{\sigma_{ii}^2 g_i^2 + \sigma_i^2 + \frac{1}{V_{ij}}} \right)^{-1}$
Curve		$g(t)$		$n$ time knots			Interpolated envelope

(iv) Site (or structure) hierarchical level ( $ij$ )

A site (or structure: furnace, oven, burnt earth or wall) or also a homogeneous set of displaced material (tiles, bricks, pottery), is a ‘volume of material that can be considered to have been magnetized at the same time’ (Tarling 1983). A site  $ij$  possesses an unknown mean value  $Y_{ij}$ . The common characteristic is the geomagnetic field elements, at a given date  $t_{ij}$ , which are supposed to be locally homogeneous.

The unknown mean  $Y_{ijk}$  at sample level is assumed to belong to a probability distribution having the unknown mean value  $Y_{ij}$  and an unknown variance  $\sigma_{ijk}^2$ . This variance between the samples from one site results from different sources of scatter as follows.

- E4: errors of heterogeneities due to differences in magnetic minerals concentration or to mineralogical transformations.
- E5: errors of magnetic anisotropy, demagnetising fields or cooling rate.
- E6: sampling errors in the field, i.e. orientation errors in the horizontal plane and azimuth.
- E7: internal dislocations in a site, as a result of possible differential movements (subsidence, tree roots, seismic activity, solifluction, etc.)

These errors can be assumed randomly distributed, except for error E5 like in sample level. The observation of  $r_{ij}$  samples allows an empirical mean  $\overline{Y}_{ij}$  and an empirical variance  $S_{ij}^2$  to be determined for the site.

(v) Field (or window) hierarchical level ( $i$ )

This is the unknown value of the geomagnetic parameter of interest  $g(t_i)$  (replacing the expected notation  $Y_i$ ) and as a function of time  $t_i$ . The common characteristic here is the time  $t_i$  common for all the sites having the same date or belonging to the same time window with centre  $t_i$  [time expressed in the calendar (solar or sidereal) date system]. The ideal situation for estimating  $g(t_i)$  would be to take  $m_i$  dated sites at a precise time  $t_i$ . In practice, sites are taken that date to a given window. In other words, one takes field estimations belonging to different times  $t_{ij}$  in order to estimate the field at the centre of the time window. This temporal reduction is one of the specific characteristics of the moving average technique. This field level is also a locality or region level (Tarling 1983): sites are sought for the same time and as near as possible to the locality where the secular variation curve is to be constructed (using a geographical correction).

The unknown mean  $Y_{ij}$  at site level is assumed to belong to a probability distribution having an unknown mean value  $g(t_{ij})$  and an unknown variance  $\sigma_i^2$ . This variance between the sites for the same time  $t_{ij}$  results from different sources of scatter as follows.

- E8: tilt errors of entire site (sliding, tilting, earthquakes effects).
- E9: homogeneous magnetic perturbations at the scale of a site, owing to environmental effects (metallic masses, lightning, magnetic anomalies. . .).
- E10: divergence of the local magnetic field from the dipole model used to calculate the geographical correction (non-dipole component effect).

These errors cannot be assumed to be randomly distributed. If errors E8 or E9 can be random from site to site, it is not inevitably the case for errors E10.

The unknown time  $t_{ij}$  in the window is assumed to belong to a probability distribution having mean time  $t_i$  (the centre of the window) and unknown variance  $\sigma_{ii}^2$ . This variance results from the following.

E11: dating errors (from archaeology, history, or other chronometric methods), modelled through a weight  $P_{ij}$ , which measures the contribution of the site to the window.

E12: the window width; we use sites of different times  $t_{ij}$  along the curve.

The observation of  $m_i$  sites allows an empirical mean  $\bar{Y}_i$  and an empirical variance  $S_i^2$  to be determined for the field at time  $t_i$ .

(vi) Curve level

The final goal of the archaeomagnetic studies is to reconstruct temporal variation curves of  $I$ ,  $D$  and  $F$ . Moving average technique are most often used to construct reference curves from time-series. A useful method for constructing these is the moving (or sliding) window (Dirichlet type), which can overlap each other, depending on the time step  $h = t_{i+1} - t_i$  and the width  $L$  chosen for the window. The total number of different windows corresponding to knots  $t_i$  is  $n$ , the number of sites in each window is  $m_i$  and the number of different reference points in the database is  $N$ . The results can be presented, for example, with a time step of either 25 yr (Daly & Le Goff 1996; Kovacheva et al. 1998), or 100 yr (Kovacheva & Toshkov 1994; Marton 1996). Batt (1997) proposed a mixing of 25- and 50-yr steps for British data, as Sternberg (1989) did for American data. In the examples treated in this paper, 25-yr steps will be considered.

The different variables of the hierarchical modelling are summarized in Table 1.

### 2.3 Directional and intensity statistics

The problem is to assign a function to the probability distributions occurring at the different levels (Tarling 1983, p. 125). It is common to characterize the variance of archaeomagnetic (palaeomagnetic) directions using the Fisher (1953) distribution (Appendix A1) and the variance of intensity using the normal distribution. A recent justification of the use of Fisher statistics for archaeomagnetic data can be found in Love & Constable (2003), who represent the variance  $\sigma^2$  of palaeomagnetic vectors, at a particular site and of a particular polarity, by a probability density function in a Cartesian three-space of orthogonal magnetic field components consisting of a single (unimodal) non-zero mean symmetrical Gaussian function. In the geophysically relevant limit of small relative dispersion  $\sigma/F_\mu$  ( $F_\mu$  being the total mean field intensity), they demonstrate that the directional distribution of the off-axis angle approaches a Fisherian distribution and that the intensity distribution (of Rayleigh–Rician type) approaches a normal distribution. Consequently, this paper adopts these probability distributions. It has been shown that, if the relative dispersion  $\sigma/F_\mu < 25$  per cent, i.e. the Fisher concentration factor  $K$  is higher than 15 (which is almost always the case in archaeomagnetism), then the approximation is excellent.

## 3 FROM HIERARCHICAL SAMPLING TO HIERARCHICAL MODELLING

Assuming that the random variables are independent at each level, then increasing the number of observations in a particular level results in a more precise estimation of the mean at this level. Up to now, the natural statistical approach is to implement the stratified statistics (defined below) and to modify it heuristically by weightings in order to force it to become hierarchical. Before presenting the hierarchical modelling in detail, it is informative to overview the stratification approach.

### 3.1 Before hierarchy: the stratified statistics

#### 3.1.1 The stratification approach

For simplicity, and without loss of generality, the problem at site and field levels will be considered with  $Y_{ijk}$  as properties of a sample. In the stratification approach, all observations of  $Y_{ijk}$  are supposed to be independent and normally distributed as  $N(g_i, \frac{\sigma_i^2}{W_{ij}})$ . The factor  $W_{ij}$  is a weighting factor, which can take different forms according to the problem of interest. The joint probability density of the observations (or likelihood  $L$ ) is proportional to

$$L[Y_{ij1}, \dots, Y_{ij(r_{ij})}, \dots, Y_{i(m_i)(r_{ij})}] \propto \exp \left[ -\frac{1}{2\sigma_i^2} \sum_{j=1}^{m_i} \sum_{k=1}^{r_{ij}} W_{ij} (Y_{ijk} - g_i)^2 \right], \quad (1)$$

where  $r_{ij}$  is the number of samples and  $m_i$  the number of sites. Using the maximum likelihood estimation (MLE) technique, the unknown field parameter  $g_i$  can then be estimated by the stratified mean  $\hat{g}_i = \bar{Y}_i$ , defined by

$$\bar{Y}_i = \frac{1}{\sum_{j,k} W_{ij}} \sum_{j,k} W_{ij} Y_{ijk}. \quad (2)$$

This mean, at field level, can be expressed as a mean of the weighted means defined at site levels, the weights being proportional to the number of samples observed at each site level:

$$\bar{Y}_i = \frac{1}{\sum_j W_{ij} r_{ij}} \sum_{j=1}^{m_i} W_{ij} r_{ij} \bar{Y}_{ij}, \quad (3)$$

where

$$\bar{Y}_{ij} = \frac{1}{r_{ij}} \sum_{k=1}^{r_{ij}} Y_{ijk}. \quad (4)$$

The stratified variance  $S_i^2$  is defined by

$$S_i^2 = \frac{1}{\sum_{j,k} W_{ij}} \sum_{j,k} W_{ij} (Y_{ijk} - \bar{Y}_i)^2. \quad (5)$$

This field-level variance can be expressed as the sum of the weighted variance of the means defined at field level and the weighted mean of the variances observed at site level:

$$S_i^2 = \underbrace{\frac{1}{\sum_j W_{ij} r_{ij}} \sum_{j=1}^{m_i} W_{ij} r_{ij} (\bar{Y}_{ij} - \bar{Y}_i)^2}_{\text{weighted variance of means}} + \underbrace{\frac{1}{\sum_j W_{ij} r_{ij}} \sum_{j=1}^{m_i} W_{ij} r_{ij} S_{ij}^2}_{\text{weighted mean of variances}} \quad (6)$$

where

$$S_{ij}^2 = \frac{1}{r_{ij}} \sum_{k=1}^{r_{ij}} (Y_{ijk} - \bar{Y}_{ij})^2. \quad (7)$$

The unbiased quadratic error of the estimation of  $g_i$  will be given by

$$e_{gi}^2 = S_i^{2*} / \sum_j W_{ij} r_{ij}, \quad \text{with} \quad S_i^{2*} = \frac{\sum_j W_{ij} r_{ij}}{(\sum_j r_{ij}) - 1} S_i^2. \quad (8)$$

### 3.1.2 The Le Goff bivariate statistics (1992)

The bivariate approach, proposed by Le Goff (1990), Le Goff *et al.* (1992) and Daly & Le Goff (1996), consists of calculating a confidence ellipse around the mean direction from a set of site directions ( $ij$ ). Each site is characterized by a mean direction ( $I_{ij}$ ,  $D_{ij}$ ) and a Fisher concentration parameter  $K_{ij}$  obtained from  $r_{ij}$  samples. The mean directions of  $m_i$  sites are assumed distributed as an elliptic bivariate statistic around a mean direction ( $I_i$ ,  $D_i$ ), with concentration parameters  $K_x$  and  $K_y$  and orientation  $\Omega$ .

The concentration parameter  $K_{ij}$  can also be described via an inertia tensor  $E[T_{ijk}]$  (Appendix A2, eq. A2.2), where the notation  $E[\dots]$  designates the mathematical expectation. The tensor is weighted by a factor  $W_{Tij} = P_{ij}$ , function of the dating error E11 and given by eq. (28) below. A rotation  $R_{ij}$ , function of ( $\lambda = -I_{ij}$ ,  $\phi = D_{ij}$ ) (eq. A1.6) is applied. The Le Goff's global weighted inertia is defined by

$$\bar{T}_i^R = \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} R_{ij}^T (E[T_{ijk}]) R_{ij}. \quad (9)$$

According to Appendix A2, the Le Goff method is a stratification approach, but not exactly the same as in the previous section. Indeed, it uses the weight  $W_{Tij} = P_{ij}$  instead of  $W_{Tij} = P_{ij} r_{ij}$ , as in eqs (3) and (6), which puts the emphasis on the means of the sites offering many samples. Some long calculations show that an estimate of the two global concentration parameters  $K_x^R$  and  $K_y^R$  (eqs A2.11 to A2.14), which characterize the shape of the global dispersion ellipse, are very well approximated (within a few %) by:

$$\frac{1}{K_x^R} \approx \frac{1}{K_x} + \frac{1}{P_i} \sum_{j=1}^{m_i} P_{ij} \frac{1}{K_{ij}}, \quad (10)$$

$$\frac{1}{K_y^R} \approx \frac{1}{K_y} + \frac{1}{P_i} \sum_{j=1}^{m_i} P_{ij} \frac{1}{K_{ij}}, \quad (11)$$

where  $P_i = \sum_{j=1}^{m_i} P_{ij}$ .

An approximate weighted mean direction can be obtained from eq. (A2.16) and the concentration parameters  $K_x$  and  $K_y$  can be approximated by eq. (A2.17), using the change of coordinate system described in Appendix A1.

Eqs (10) and (11) are equivalent to variances and are equal to the sum of the weighted variance of the mean directions of the sites and of the weighted mean of the variances at site levels, as in eq. (6).

The elliptic confidence angles proposed by Le Goff *et al.* (1992) are:

$$\alpha_{95x} = t_{\beta F} / \sqrt{\left(\sum_j P_{ij} r_{ij}\right) K_x^R}, \quad (12)$$

$$\alpha_{95y} = t_{\beta F} / \sqrt{\left(\sum_j P_{ij} r_{ij}\right) K_y^R}. \quad (13)$$

The marginal errors of inclination and declination are given by:

$$e_{Y_{Ii}} = t_{\beta/2} \sqrt{\frac{1}{\left(\sum_j P_{ij} r_{ij}\right)} \left[ \frac{\cos^2 \Omega}{K_x^R} + \frac{\sin^2 \Omega}{K_y^R} \right]},$$

$$e_{Y_{Di}} = t_{\beta/2} \sqrt{\frac{1}{\left(\sum_j P_{ij} r_{ij}\right)} \left[ \frac{\cos^2 \Omega}{K_y^R} + \frac{\sin^2 \Omega}{K_x^R} \right]}, \quad (14)$$

where  $t_{\beta/2}$  is the Student coefficient.

These errors will be very small because of the contribution of all  $\sum_j r_{ij}$  samples, assumed to be independent, as in eq. (8). This observation constitutes the essential difference between the stratified approach and the hierarchical approach, although  $W_{Tij} = P_{ij}$ .

### 3.1.3 The classical weighted univariate statistics

The univariate statistics currently used in the palaeomagnetic and archaeomagnetic literature, performed on variates  $I$ ,  $D$  or  $F$  separately (here generically noted as  $X$ ), is also fundamentally based on a stratified approach. The estimate of the geomagnetic parameter  $g_i$  of interest is given by the formula

$$\hat{g}_i = \bar{X}_i = \frac{1}{W_i} \sum_{j=1}^{m_i} W_{ij} \bar{X}_{ij}, \quad \text{with} \quad W_i = \sum_{j=1}^{m_i} W_{ij} \quad (15)$$

where the weighting factor is

$$W_{ij} = P_{ij} \kappa_{ij}^* \quad (\text{Sternberg 1989}), \quad (16)$$

$$W_{ij} = \frac{P_{ij}}{\alpha_{95ij}^2} \quad (\text{Kovacheva \& Toshkov 1994; Kovacheva et al. 1998}), \quad (17)$$

$$W_{ij} = \frac{P_{ij}}{\alpha_{95ij}} \quad (\text{Batt 1997}), \quad (18)$$

and where  $\alpha_{95ij} = 2.45 / \sqrt{r_{ij} \kappa_{ij}^*}$  ( $\kappa_{ij}^*$  is defined in Appendix A1).

The factor  $P_{ij}$  is related to dating error E11 and is given by eq. (28). Assuming that the variates  $\bar{X}_{ij}$  have a Gaussian distribution, then the confidence interval for  $g(t_i)$  is given by the Student t-distribution:

$$p(\hat{g}_i - e_{\bar{X}_i} \leq g_i \leq \hat{g}_i + e_{\bar{X}_i}) = 1 - \beta, \quad (19)$$

where the unbiased weighted error is

$$e_{\bar{X}_i} = t_{\beta/2} \sqrt{\frac{S_i^{2*}}{m_i}}, \quad \text{with} \quad S_i^{2*} = \left( \frac{m_i}{m_i - 1} \right) \frac{1}{W_i} \sum_{j=1}^{m_i} W_{ij} (\bar{X}_{ij} - \hat{g}_i)^2. \quad (20)$$

This statistical approach considers only the last (field) level. Contrary to eqs (12) to (14), the estimation error decreases with  $m_i$ , not with  $\sum_j r_{ij}$ . Unfortunately, there is no theoretical justification for the form of the expression of the weight  $W_{ij}$  in eqs (15) to (18). Moreover, this estimation is very sensitive to the values of  $\kappa_{ij}^*$  or  $\alpha_{95ij}$  and a very high precision for a particular site ( $ij$ ) will strongly draw the estimate  $\hat{g}(t_i)$  towards the observation  $\bar{Y}_{ij}$  at this site. This statistics is also very sensitive to dating errors  $P_{ij}$ .

## 3.2 The hierarchical modelling

The previous section suggests that a hierarchical statistical model of the experimental errors would be more robust. This modelling can be implemented using the Bayesian statistics (Buck *et al.* 1991; Droesbeke *et al.* 2002), which relates posterior and predictive (marginal) densities to likelihood and prior densities. Here, hierarchical priors are considered and defined on hyperparameters  $Y_{ijkl}$ ,  $Y_{ijk}$ ,  $Y_{ij}$  and  $t_{ij}$  (Table 1).



The Fisher statistics (Appendix A1) can be approximated by a bivariate Gaussian statistics in the tangential plane normal to the polar mean direction of a given level in the hierarchical modelling. This very good approximation allows the analytic calculations to be easily performed in the hierarchical framework. For this purpose, the coordinate system needs to be changed using an appropriate rotation described in Appendix A1, to obtain directions near the geographical equatorial plane. In the new coordinate system, the variables  $Y_{\bullet}$  (the black point is for indices) are vectors with new components of inclination and declination ( $Y_{I\bullet}, Y_{D\bullet}$ ). Considering the hierarchical sampling in Section 2.2 and the relevant directional statistics in Section 2.3, the following probability densities can be defined.

### 3.2.1 Observation at measurement level

Each observation corresponds to an elementary measurement  $Y_{ijklm}$ .

### 3.2.2 Density of the observation at specimen level

The measurements  $Y_{ijklm}$  are assumed independent and normally distributed with an unknown mean value (hyperparameter)  $Y_{ijkl}$  and unknown intermeasurement variance  $\sigma_{ijkl}^2$  at specimen level:

$$p(Y_{ijklm}|Y_{ijkl}) = \frac{1}{2\pi\sigma_{ijkl}^2} \exp \left[ -\frac{1}{2\sigma_{ijkl}^2} (Y_{ijklm} - Y_{ijkl})^T (Y_{ijklm} - Y_{ijkl}) \right]. \quad (21)$$

The notation  $T$  is for the transposition.

### 3.2.3 Prior density at sample level

The unknown mean  $Y_{ijkl}$  of a specimen  $ijkl$  is assumed normally distributed with an unknown mean (hyperparameter)  $Y_{ijk}$  and unknown interspecimen variance  $\sigma_{ijk}^2$  at sample level:

$$p(Y_{ijkl}|Y_{ijk}) = \frac{1}{2\pi\sigma_{ijk}^2} \exp \left[ -\frac{1}{2\sigma_{ijk}^2} (Y_{ijkl} - Y_{ijk})^T (Y_{ijkl} - Y_{ijk}) \right]. \quad (22)$$

### 3.2.4 Prior density at site level

The unknown mean  $Y_{ijk}$  of a sample  $ijk$  is normally distributed with an unknown mean value (hyperparameter)  $Y_{ij}$  and variance  $\sigma_{ij}^2$  at site level:

$$p(Y_{ijk}|Y_{ij}) = \frac{1}{2\pi\sigma_{ij}^2} \exp \left[ -\frac{1}{2\sigma_{ij}^2} (Y_{ijk} - Y_{ij})^T (Y_{ijk} - Y_{ij}) \right]. \quad (23)$$

### 3.2.5 Prior densities at field level

**3.2.5.1 Prior density on  $Y_{ij}$ .** The unknown mean  $Y_{ij}$  of a site  $ij$  is normally distributed with an unknown mean vector  $Y_g(t_{ij})$  and variance  $\sigma_i^2$  at a field level:

$$p(Y_{ij}|Y_g(t_{ij})) = \frac{1}{2\pi\sigma_i^2} \exp \left[ -\frac{1}{2\sigma_i^2} (Y_{ij} - Y_g(t_{ij}))^T (Y_{ij} - Y_g(t_{ij})) \right]. \quad (24)$$

The unknown mean vector value  $Y_g(t_{ij})$  corresponds to the value of the geomagnetic field parameter at time  $t_{ij}$ . The aim of the work is to estimate the geomagnetic field at time  $t_i$ , the centre of the window  $i$ . This poses the problem in reducing the variables from  $Y_g(t_{ij})$  to  $Y_g(t_i)$ .

(a) A first rough approach is to assume that the variates  $Y_g(t_{ij})$  are normally distributed with respect to  $Y_g(t_i)$ , without any correlation between the components [ $Y_{gI}(t)$ ,  $Y_{gD}(t)$ ]. This case is most often not realistic because these components are expected to be dependent on time.

(b) In order to take into account the correlation between the two components with respect to time, the second approach consists of approximating  $Y_g$  to a linear function in the window, using the Taylor expansion of the first order:

$$Y_{gI}(t_{ij}) \approx Y_{gI}(t_i) + (t_{ij} - t_i)Y'_{gI}(t_i),$$

$$Y_{gD}(t_{ij}) \approx Y_{gD}(t_i) + (t_{ij} - t_i)Y'_{gD}(t_i), \quad (25)$$

where  $Y'_g(t_i)$  is the first derivative of curve  $Y_g$  at time  $t_i$ . This approximation is valid provided that the window width is never too large compared with the local variations of the curve. Then, the density (eq. 24) becomes

$$p(Y_{ij}|Y_g(t_i), t_{ij}) = \frac{1}{2\pi\sigma_i^2} \exp \left[ -\frac{1}{2\sigma_i^2} (Y_{ij} - Y_g(t_i) - (t_{ij} - t_i)Y'_g(t_i))^T (Y_{ij} - Y_g(t_i) - (t_{ij} - t_i)Y'_g(t_i)) \right]. \quad (26)$$

3.2.5.2 Prior density on time  $t_{ij}$ .

(a) If all the  $t_{ij}$  are equal to a known and precise time  $t_i$  ( $P_{ij} = 1$ ), then  $Y_{ij} \sim N(Y_g(t_i), \sigma_i^2)$ . The field  $Y_g(t_i)$  can be directly estimated as a mean value of the  $m_i$  studied sites corresponding to the time  $t_i$ . However, this case is very rare: it is difficult to obtain many reference sites having exactly the same date, within the same geographic zone.

(b) Most often, the times  $t_{ij}$  are not well known. In other words, they are somewhere in the interval  $[t_i - L/2, t_i + L/2]$ , without any more information. In this case, the unknown times  $t_{ij}$  are modelled as a normal prior density about the centre  $t_i$ , as suggested by Véges (1970) and applied by Marton (1996):

$$p(t_{ij}|Y_g(t_i)) = p(t_{ij}) = \frac{1}{\sigma_{ii}\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma_{ii}^2}(t_{ij} - t_i)^2\right), \quad (27)$$

where the time variance  $\sigma_{ii}^2$  depends on the window width. This says that the times  $t_{ij}$  are near the centre  $t_i$ , thus ensuring the validity of the Taylor approximation in eq. (25). This model takes into account the lack of information about where the times are inside the window (through variance  $\sigma_{ii}^2$ ).

3.2.5.3 Accounting the dating errors. The precision of the date attributed to the site by history, archaeology or chronometric methods, will be taken into account through the weight  $P_{ij}$ . To determine this weight, the idea of Sternberg (1989) is basically used here. It is calculated as a function of the contribution of the dating interval  $[t_{ij1}, t_{ij2}]$  to the window  $[t_i - L/2, t_i + L/2]$ . The weight is proportional to the overlap of the age range of site  $ij$  within the given window  $i$ :

$$P_{ij} = d([t_{ij1}, t_{ij2}] \cap [t_i - L/2, t_i + L/2]) / (t_{ij2} - t_{ij1}), \quad (28)$$

where the distance  $d([x, y]) = |x - y|$ . In the specific case when  $t_{ij} = t_i$ , then  $P_{ij} = 1$ .

When the distance  $h$  separating the times  $t_i$  is chosen exactly equal to the width of the window  $L$ , then  $\sum_{i=1}^n \sum_{j=1}^{m_i} P_{ij} = N$ , the total number of reference sites. This configuration has been chosen for example in Kovacheva & Toshkov (1994) and Márton (1996): window length and step size are 100 yr.

## 3.2.6 Joint density of the observations: hierarchical likelihood of order four

To pass from the measurement to the reference curve,  $d_{ijkl}$  measurements can be performed for one specimen to estimate the unknown mean value  $Y_{ijkl}$ , for fixed values of  $Y_{ijk}$ ,  $Y_{ij}$  and  $Y_g(t_{ij})$ , and  $n_{ijk}$  different specimens can be taken from a sample to estimate the unknown mean value  $Y_{ijk}$ , for fixed values of  $Y_{ij}$  and  $Y_g(t_{ij})$ . After that,  $r_{ij}$  different samples from a site can be taken to estimate the value  $Y_{ij}$ , for a fixed value of  $Y_g(t_{ij})$ . At the end, one can take  $m_i$  sites in a window  $i$  to estimate  $Y_g(t_i)$  (Fig. 1b). Thus, there can be  $d_{ijkl}$  measurements in specimen  $ijkl$ ,  $n_{ijk}$  specimens in sample  $ijk$ ,  $r_{ij}$  samples in a site  $ij$  and  $m_i$  sites with their dating ranges belonging (totally or partially) to the time window  $[t_i - L/2, t_i + L/2]$ .

Assuming that all the observations are independent at each of the levels described above, the hierarchical joint density  $L_i$ , or likelihood of order 4, of the observations  $Y_{ijklm}$  and the unknown variates (hyperparameters)  $Y_{ijkl}$ ,  $Y_{ijk}$ ,  $Y_{ij}$  and  $t_{ij}$ , can be written for fixed  $Y_g(t_i)$ . This likelihood relates the observation density in eq. (21) and the prior hierarchical specifications in eqs (22), (23), (26) and (27):

$$L_i = \prod_{j=1}^{m_i} \left[ \iint L_{ij} p(Y_{ij}|Y_g(t_i), t_{ij}) p(t_{ij}) dY_{ij} dt_{ij} \right]^{P_{ij}}, \quad (29)$$

where

$$L_{ij} = \prod_{k=1}^{r_{ij}} \int L_{ijk} p(Y_{ijk}|Y_{ij}) dY_{ijk}, \quad (29a)$$

$$L_{ijk} = \prod_{l=1}^{n_{ijk}} \int L_{ijkl} p(Y_{ijkl}|Y_{ijk}) dY_{ijkl}, \quad (29b)$$

$$L_{ijkl} = \prod_{m=1}^{d_{ijkl}} p(Y_{ijklm}|Y_{ijkl}). \quad (29c)$$

The field  $Y_g(t_i)$  is estimated by MLE, after elimination of the hyperparameters by integration and definition of appropriate empirical variables.

## 4 HIERARCHICAL STATISTICS

The following discussion deals with the direction of the magnetic field. Results for intensity (univariate case) are given in detail in Appendix A3.

4.1 Definition of the empirical variables and elimination of the hyperparameters

The aim of the calculation is to estimate the vector  $Y_g(t_i)$  starting from the observations (elementary measurements)  $Y_{ijklm}$ , by maximizing the joint density  $L_i$  in eq. (29).

(i) At a specimen level, the likelihood is given by eq. (29c).

The empirical means and variances are defined by

$$\begin{aligned} \overline{Y_{Iijkl}} &= \frac{1}{d_{ijkl}} \sum_{m=1}^{d_{ijkl}} Y_{Iijklm} & \overline{Y_{Dijkl}} &= \frac{1}{d_{ijkl}} \sum_{m=1}^{d_{ijkl}} Y_{Dijklm} \\ S^2_{Iijkl} &= \frac{1}{d_{ijkl}} \sum_{m=1}^{d_{ijkl}} (Y_{Iijklm} - \overline{Y_{Iijkl}})^2 & S^2_{Dijkl} &= \frac{1}{d_{ijkl}} \sum_{m=1}^{d_{ijkl}} (Y_{Dijklm} - \overline{Y_{Dijkl}})^2. \end{aligned} \tag{30}$$

The symbol bar above letters  $Y$  designates an empirical mean operation. These empirical determinations generally result from specific statistical treatments of the selected set of  $d_{ijkl}$  elementary measurements (principal component analysis, in the case of demagnetisation, and regression analysis applied to an Arai diagram in the case of intensity). It can also result from only one measurement, that is  $\overline{Y_{ijkl}} = Y_{ijklm}$ . Replacing in eq. (29c), the following equality can be written

$$\sum_{m=1}^{d_{ijk}} \frac{1}{\sigma^2_{ijk}} ((Y_{Iijklm} - Y_{Iijkl})^2 + (Y_{Dijklm} - Y_{Dijkl})^2) = \frac{d_{ijkl}}{\sigma^2_{ijk}} ((\overline{Y_{Iijkl}} - Y_{Iijkl})^2 + (\overline{Y_{Dijkl}} - Y_{Dijkl})^2 + S^2_{Iijkl} + S^2_{Dijkl}). \tag{31}$$

According to Cochran's theorem (Dudewicz & Mishra 1988), the empirical mean vector  $\overline{Y_{ijkl}}$  is normally distributed with a sampling squared error  $\frac{\sigma^2_{ijkl}}{d_{ijkl}}$ , and the variables  $\frac{d_{ijkl}S^2_{Iijkl}}{\sigma^2_{ijk}}$  and  $\frac{d_{ijkl}S^2_{Dijkl}}{\sigma^2_{ijk}}$  are  $\chi^2$  distributed with  $(d_{ijkl} - 1)$  degrees of freedom.

(ii) At a sample level, the likelihood is defined by eq. (29b). The hyperparameter (vector)  $Y_{ijk}$ , being unknown, can be eliminated through an integration between  $-\infty$  and  $+\infty$ . The probability density function of  $\overline{Y_{ijkl}}$  with respect to the true sample mean  $Y_{ijk}$  is given by the following integration:

$$\begin{aligned} P(\overline{Y_{ijkl}} | Y_{ijk}) &= \int_{-\infty}^{+\infty} \frac{\sqrt{d_{ijkl}}}{2\pi\sigma_{ijk}\sigma_{ijkl}} \exp\left(-\frac{1}{2}\left(\frac{d_{ijkl}}{\sigma^2_{ijk}}(\overline{Y_{ijkl}} - Y_{ijkl})^T(\overline{Y_{ijkl}} - Y_{ijkl}) + \frac{1}{\sigma^2_{ijk}}(Y_{ijkl} - Y_{ijk})^T(Y_{ijkl} - Y_{ijk})\right)\right) dY_{ijkl} \\ &= \frac{1}{\sqrt{2\pi(\sigma^2_{ijk} + \frac{\sigma^2_{ijkl}}{d_{ijkl}})}} \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2_{ijk} + \frac{\sigma^2_{ijkl}}{d_{ijkl}}}(\overline{Y_{ijkl}} - Y_{ijk})^T(\overline{Y_{ijkl}} - Y_{ijk})\right)\right). \end{aligned} \tag{32}$$

The Bayesian hierarchical predictive variance (Droesbeke *et al.* 2002) here is equal to

$$\frac{1}{U_{ijkl}} = \sigma^2_{ijk} + \frac{\sigma^2_{ijkl}}{d_{ijkl}}.$$

The empirical means and variances are defined by

$$\begin{aligned} \overline{Y_{Iijk}} &= \frac{1}{U_{ijk}} \sum_{l=1}^{n_{ijk}} U_{ijkl} \overline{Y_{Iijkl}} & \overline{Y_{Dijk}} &= \frac{1}{U_{ijk}} \sum_{l=1}^{n_{ijk}} U_{ijkl} \overline{Y_{Dijkl}} \\ S^2_{Iijk} &= \frac{1}{U_{ijk}} \sum_{l=1}^{n_{ijk}} U_{ijkl} (\overline{Y_{Iijkl}} - \overline{Y_{Iijk}})^2 & S^2_{Dijk} &= \frac{1}{U_{ijk}} \sum_{l=1}^{n_{ijk}} U_{ijkl} (\overline{Y_{Dijkl}} - \overline{Y_{Dijk}})^2. \end{aligned} \tag{33}$$

Hence, according to Cochran's theorem, the empirical mean vector  $\overline{Y_{ijk}}$  is normally distributed with a squared error

$$\frac{1}{U_{ijk}} = \frac{1}{\sum_l U_{ijkl}} = \frac{1}{\sum_l \frac{1}{\sigma^2_{ijk} + \frac{\sigma^2_{ijkl}}{d_{ijkl}}}}, \tag{34}$$

and the empirical variables  $U_{ijk}S^2_{Iijk}$  and  $U_{ijk}S^2_{Dijk}$  are  $\chi^2$  distributed with  $(n_{ijk} - 1)$  degrees of freedom.

Assuming that the errors  $e^2_{ijkl} = \frac{\sigma^2_{ijkl}}{d_{ijkl}}$  are of the same order of magnitude from specimen to specimen, they can be replaced by

$$\overline{e^2_{ms_{ijk}}} = \frac{1}{n_{ijk}} \sum_{l=1}^{n_{ijk}} e^2_{ijkl}, \tag{35}$$

hence  $\frac{1}{U_{ijk}} = \frac{\sigma^2_{ijk} + \overline{\sigma^2_{ms_{ijk}}}}{n_{ijk}}$ .

(iii) At a site level, the likelihood is defined by eq. (29a). The hyperparameter (vector)  $Y_{ijk}$ , being unknown, can be eliminated through an integration between  $-\infty$  and  $+\infty$ . The Bayesian hierarchical predictive variance here is equal to  $\frac{1}{V_{ijk}} = \sigma^2_{ij} + \frac{1}{U_{ijk}}$ .

The empirical means and variances are defined by

$$\begin{aligned} \overline{Y_{Iij}} &= \frac{1}{V_{ij}} \sum_{k=1}^{r_{ij}} V_{ijk} \overline{Y_{Iijk}} & \overline{Y_{Dij}} &= \frac{1}{V_{ij}} \sum_{k=1}^{r_{ij}} V_{ijk} \overline{Y_{Dijk}} \\ S_{Iij}^2 &= \frac{1}{V_{ij}} \sum_{k=1}^{r_{ij}} V_{ijk} (\overline{Y_{Iijk}} - \overline{Y_{Iij}})^2 & S_{Dij}^2 &= \frac{1}{V_{ij}} \sum_{k=1}^{r_{ij}} V_{ijk} (\overline{Y_{Dijk}} - \overline{Y_{Dij}})^2, \end{aligned} \tag{36}$$

Hence, according to Cochran's theorem, the empirical mean vector  $\overline{Y_{ij}}$  is normally distributed with squared error

$$\frac{1}{V_{ij}} = \frac{1}{\sum_{k=1}^{r_{ij}} V_{ijk}} = \frac{1}{\sum_{k=1}^{r_{ij}} \frac{1}{\sigma_{ij}^2 + \sigma_{ijk}^2}} = \frac{1}{\sum_{k=1}^{r_{ij}} \frac{1}{\sigma_{ij}^2 + \frac{\sigma_{ijk}^2 + \sigma_{msijk}^2}{n_{ijk}}}}, \tag{37}$$

and the empirical variables  $V_{ij} S_{Iij}^2$  and  $V_{ij} S_{Dij}^2$  are  $\chi^2$  distributed with  $(r_{ij} - 1)$  degrees of freedom.

(iv) Finally, at a field level, the likelihood is defined by eq. (29). The hyperparameters (vector)  $Y_{ij}$  and  $t_{ij}$ , being unknown, can also be eliminated by an integration between  $-\infty$  and  $+\infty$ . The weighting factor  $P_{ij}$  measures the contribution of each site to the window and can be taken into account in the inverse squared errors defined below in eq. (40). This approach leads to the (more complicated) definition of the last set of empirical variables.

The empirical means are

$$\begin{aligned} \overline{Y_{Ii}} &= \frac{1}{W_{Ii} W_{Di} - W_{IDi}^2} \sum_{j=1}^{m_i} (W_{Di} W_{Iij} - W_{IDi} W_{IDij}) \overline{Y_{Iij}} + (W_{IDi} W_{Dij} - W_{Di} W_{IDij}) \overline{Y_{Dij}}, \\ \overline{Y_{Di}} &= \frac{1}{W_{Ii} W_{Di} - W_{IDi}^2} \sum_{j=1}^{m_i} (W_{IDi} W_{Iij} - W_{Ii} W_{IDij}) \overline{Y_{Iij}} + (W_{Ii} W_{Dij} - W_{IDi} W_{IDij}) \overline{Y_{Dij}}, \end{aligned} \tag{38}$$

and the empirical variances and covariance are

$$\begin{aligned} S_{Ii}^2 &= \frac{1}{W_{Ii}} \sum_{j=1}^{m_i} W_{Iij} (\overline{Y_{Iij}} - \overline{Y_{Ii}})^2, \\ S_{Di}^2 &= \frac{1}{W_{Di}} \sum_{j=1}^{m_i} W_{Dij} (\overline{Y_{Dij}} - \overline{Y_{Di}})^2, \\ S_{IDi} &= \frac{1}{W_{IDi}} \sum_{j=1}^{m_i} W_{IDij} (\overline{Y_{Iij}} - \overline{Y_{Ii}}) (\overline{Y_{Dij}} - \overline{Y_{Di}}), \end{aligned} \tag{39}$$

where the weights  $W_{Iij}$ ,  $W_{Dij}$  and  $W_{IDij}$  are defined by

$$\begin{aligned} W_{Ii} &= \sum_{j=1}^{m_i} W_{Iij} = \sum_{j=1}^{m_i} \frac{P_{ij}}{(1 - \rho_{ij}^2) B_{Iij}}, \\ W_{Di} &= \sum_{j=1}^{m_i} W_{Dij} = \sum_{j=1}^{m_i} \frac{P_{ij}}{(1 - \rho_{ij}^2) B_{Dij}}, \\ W_{IDi} &= \sum_{j=1}^{m_i} W_{IDij} = \sum_{j=1}^{m_i} \frac{P_{ij} \rho_{ij}}{(1 - \rho_{ij}^2) \sqrt{B_{Iij} B_{Dij}}}, \end{aligned} \tag{40}$$

and where the Bayesian hierarchical variances  $B_{Iij}$  and  $B_{Dij}$  are

$$B_{Iij} = \sigma_{ii}^2 g_I'^2 + \sigma_i^2 + \frac{1}{V_{ij}}, \quad B_{Dij} = \sigma_{ii}^2 g_D'^2 + \sigma_i^2 + \frac{1}{V_{ij}}.$$

The function  $g'$  is the derivative of  $g$ . The correlation factor  $\rho_{ij}$ , which depends on time distribution and slope of the curve in the window is given by

$$\rho_{ij} = \frac{\sigma_{ii}^2 g_I' g_D'}{\sqrt{B_{Iij} B_{Dij}}}. \tag{41}$$

#### 4.2 Probability density function of the direction

The normal joint density of vector  $\overline{Y}_i$ , conditional to  $Y_g(t_i)$ , is obtained by

$$p(\overline{Y}_i | Y_g(t_i)) = \frac{1}{2\pi} |\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\overline{Y}_i - Y_g(t_i))^T \Sigma_i^{-1} (\overline{Y}_i - Y_g(t_i))\right), \tag{42}$$

where the squared error matrix is defined by eqs (40):

$$\Sigma_i = \frac{1}{W_{Ii}W_{Di} - W_{IDi}^2} \begin{pmatrix} W_{Di} & W_{IDi} \\ W_{IDi} & W_{Ii} \end{pmatrix}. \quad (43)$$

The Bayesian hierarchical approach leads to the matrix  $\Sigma_i$ , which defines the squared error on the mean vector  $\bar{Y}_i$ , and is itself a combination of squared errors (also called sampling variances) and variances resulting from the sources of scatter, in a similar way to eq. (34) for sample levels or eq. (37) for site levels. Their expressions are analogous to a harmonic mean.

The probability density function of the empirical variances  $S_{Ii}^2$ ,  $S_{Di}^2$ ,  $S_{IDi}^2$ , defined in eqs (39), is a Wishart distribution  $W(m_i - 1, \Sigma_i)$  (Anderson 1984):

$$p(S_i) = \frac{|S_i|^{\frac{m_i-4}{2}}}{2^{m_i-1} \sqrt{\pi} \Gamma(\frac{m_i-2}{2}) \Gamma(\frac{m_i-1}{2}) |\Sigma_i|^{\frac{m_i-1}{2}}} \exp \left[ -\frac{1}{2} \text{trace}(\Sigma_i^{-1} S_i) \right], \quad (44)$$

where  $S_i$  is the empirical variance-covariance matrix

$$S_i = \begin{pmatrix} S_{Ii}^2 & S_{IDi} \\ S_{IDi} & S_{Di}^2 \end{pmatrix}. \quad (45)$$

### 4.3 Hierarchical elliptic statistics

From eq. (42), the probability density function in the tangential plane to the sphere is

$$p(\bar{Y}_i | Y_g(t_i)) = \frac{\sqrt{W_{Ii}W_{Di} - W_{IDi}^2}}{2\pi} \exp \left[ -\frac{1}{2} [W_{Ii}(\bar{Y}_{Ii} - Y_{gI})^2 - 2W_{IDi}(\bar{Y}_{Ii} - Y_{gI})(\bar{Y}_{Di} - Y_{gD}) + W_{Di}(\bar{Y}_{Di} - Y_{gD})^2] \right]. \quad (46)$$

The equivalent function on the sphere is defined by a change of variable:

$$\bar{Y}_{Ii} - Y_{gI} = \delta \cos \varphi,$$

$$\bar{Y}_{Di} - Y_{gD} = \delta \sin \varphi,$$

where  $\delta$  is approximately the off-axis error angle from the mean direction  $(\bar{Y}_{Ii}, \bar{Y}_{Di})$  and  $\varphi$  is the azimuth direction. The Jacobian becomes  $\delta$  and

$$p(\delta, \varphi) = \frac{\sqrt{W_{Ii}W_{Di} - W_{IDi}^2}}{2\pi} \exp \left[ -\frac{1}{2} [W_{Ii} \cos^2 \varphi - W_{IDi} \sin(2\varphi) + W_{Di} \sin^2 \varphi] \delta^2 \right] \times \delta. \quad (47)$$

This is the density in the eigenvector coordinate system after diagonalization. The eigenvalues are

$$K_{ex} = \frac{1}{\cos(2\Omega)} [W_{Ii} \cos^2 \Omega - W_{Di} \sin^2 \Omega], \quad (48)$$

$$K_{ey} = \frac{1}{\cos(2\Omega)} [W_{Di} \cos^2 \Omega - W_{Ii} \sin^2 \Omega],$$

where  $\Omega$  is the angle between the direction of the principal axis of the ellipse in the tangential plane and the geographical meridian passing through the mean direction:

if  $W_{IDi} = 0$ , then

$$\Omega = 0;$$

else

if  $W_{Di} - W_{Ii} = 0$ , then  $\Omega = \pi/4$ ,

if  $W_{Di} - W_{Ii} > 0$ , then  $\Omega = 0.5 \arctan \left[ \frac{2W_{IDi}}{W_{Di} - W_{Ii}} \right]$ ,

if  $W_{Di} - W_{Ii} < 0$ , then  $\Omega = 0.5 \arctan \left[ \frac{2W_{IDi}}{W_{Di} - W_{Ii}} \right] + \frac{W_{IDi}}{|W_{IDi}|} \frac{\pi}{2}$ . (49)

After diagonalization, and making the approximation  $\cos \delta \approx 1 - \frac{\delta^2}{2}$  and  $\sin \delta \approx \delta$ , then the probability density function of the error angle  $\delta$  and the azimuth  $\varphi$  is defined as

$$p(\delta, \varphi) \approx \frac{\sqrt{K_{ex}K_{ey}}}{2\pi} \exp \left[ [K_{ex} \cos^2(\varphi - \Omega) + K_{ey} \sin^2(\varphi - \Omega)] (\cos \delta - 1) \right] \times \sin \delta. \quad (50)$$

It is possible to define a probability density function for the distribution of an off-axis angle  $\theta$  by introducing the two Bayesian elliptical concentration factors:

$$K_x = K_{ex}/m_i, \quad K_y = K_{ey}/m_i, \tag{51}$$

$$p(\theta, \varphi) = \frac{\sqrt{K_x K_y}}{4\pi \operatorname{sh} [K_x \cos^2(\varphi - \Omega) + K_y \sin^2(\varphi - \Omega)]} \exp [[K_x \cos^2(\varphi - \Omega) + K_y \sin^2(\varphi - \Omega)] \cos \theta] \times \sin \theta. \tag{52}$$

This result is similar to that proposed by Le Goff (1990) and Le Goff *et al.* (1992). The new contribution here is that the concentration factors are clearly related to the sampling errors. The elliptical distribution observed reveals the influence of the window width. The sites attributed to the window, and so contributing to the calculation of  $Y_g(t_i)$ , in fact belong to different parts of the curve inside the window. Consequently, the elliptical nature of the distribution describes the bias effect resulting from the reduction of the data to time  $t_i$ : the principal axis of the ellipse will be tangential to a great circle in the considered window  $i$ . As such, this statistics represents a regression on the sphere and the elliptic tendency of the directions ( $I_{ij}, D_{ij}$ ) will increase when the data are widely distributed in time, provided that the linear approximation in eq. (25) holds. This implicit approximation of  $Y_g$  by straight lines in each window results in an important smoothing effect. The wider the window, the larger the smoothing effect on  $Y_g$ . Conversely, if the width of the window is reduced, this approximation becomes more reasonable. At the same time, the number of sites  $m_i$  falling in the window will diminish, resulting in less precision of the estimation of  $Y_g(t_i)$  (also mentioned by Sternberg 1989; Batt 1997). Thus, the question is the choice of the ideal window width. In the present approach, there is no theoretical argument that permits the determination of an optimal window width.

**4.4 Hierarchical Fisher statistics**

In the context of Fisherian statistics, an equivalent Bayesian concentration factor can be defined as

$$\frac{1}{K_B} = \frac{1}{2} \left( \frac{1}{K_x} + \frac{1}{K_y} \right) = \frac{m_i}{2} \left( \frac{W_{Ii} + W_{Di}}{W_{Ii} W_{Di} - W_{Di}^2} \right). \tag{53}$$

Replacing  $K_x$  and  $K_y$  by  $K_B$ , the probability density function of  $\theta$  and  $\varphi$  becomes

$$p(\theta, \varphi) \approx \frac{K_B}{4\pi \sinh K_B} \exp (K_B \cos \theta) \times \sin \theta, \tag{54}$$

which defines the polar hierarchical Fisher statistics.

In the specific case of sites having the same date  $t_i$ , e.g.  $P_{ij} = 1$  and  $\sigma_{ii}^2 = 0$ , then  $W_{Di} = 0$  and the Bayesian concentration factor  $K_B$  in (53) will be explicitly expressed as  $K_{B0}$ :

$$\frac{1}{K_{B0}} = \frac{m_i}{2} \left( \frac{1}{W_{Ii}} + \frac{1}{W_{Di}} \right) = \frac{m_i}{\sum_{j=1}^{m_i} \frac{1}{\sigma_{ij}^2 + \frac{1}{\sum_{k=1}^{I_{ij}} \frac{1}{\sigma_{ijk}^2 + \frac{\epsilon_{ms_{ijk}}^2}{n_{ijk}}}}}}, \tag{55}$$

which defines how to combine different Bayesian Fisherian subpopulations.

Considering the very good approximation of the Fisher statistics to a bivariate Gaussian one, provided that concentration factor is large enough (Appendix A1), the variances  $\sigma_i^2$ ,  $\sigma_{ij}^2$  and  $\sigma_{ijk}^2$  at the different levels can be replaced by the Fisher concentration factors  $1/K_i$ ,  $1/K_{ij}$  and  $1/K_{ijk}$ , respectively.

**4.5 Estimation of the directional parameters**

The unknown variances  $\sigma_i^2$ ,  $\sigma_{ij}^2$ ,  $\sigma_{ijk}^2$  and  $\sigma_{ijkl}^2$ , and the function  $Y_g(t_i)$  at time  $t_i$  will be determined in the context of the theory of parameter estimation by maximum likelihood, called MLE (Dudewicz & Mishra 1988; Tassi 1992), and adapted to the hierarchical statistics. In the following, the notation  $\hat{\cdot}$  is for an estimate and  $\ast$  is that corrected for bias. In particular,  $S_{ijkl}^{2\ast} = \frac{d_{ijkl}}{d_{ijkl}-1} S_{ijkl}^2$ ,  $S_{ijk}^{2\ast} = \frac{n_{ijk}}{n_{ijk}-1} S_{ijk}^2$ ,  $S_{ij}^{2\ast} = \frac{r_{ij}}{r_{ij}-1} S_{ij}^2$  and  $S_i^{2\ast} = \frac{m_i}{m_i-1} S_i^2$ . Moreover, the unbiased empirical concentration factors  $\kappa_{ijkl}^*$ ,  $\kappa_{ijk}^*$ ,  $\kappa_{ij}^*$  and  $\kappa_i^*$  are used as defined in eq. (A1.5).

**4.5.1 The general solution**

(i) At a specimen level, an unbiased estimation of the intermeasurement variance  $\sigma_{ijkl}^2$  will be given by

$$\hat{\sigma}_{ijkl}^2 = \frac{S_{ijkl}^{2\ast} + S_{Dijkl}^{2\ast}}{2} = \frac{1}{\kappa_{ijkl}^*} \tag{56}$$

and the mean squared error at the specimen level can be calculated:

$$\overline{e^2_{ms_{ijk}}} = \frac{1}{n_{ijk}} \sum_{l=1}^{n_{ijk}} \frac{\sigma^2_{ijkl}}{d_{ijkl}}$$

(ii) At a sample level, an unbiased estimation of the interspecimen variance  $\sigma^2_{ijk}$  will be given by:  $\widehat{\sigma^2_{ijk} + e^2_{ms_{ijk}}} = \frac{S^{2*}_{Iijk} + S^{2*}_{Dijk}}{2}$ , hence the Bayesian solution:

$$\hat{\sigma}^2_{ijk} = \frac{\hat{1}}{K_{ijk}} = \max \left( 0, \frac{S^{2*}_{Iijk} + S^{2*}_{Dijk}}{2} - \overline{e^2_{ms_{ijk}}} \right) \approx \max \left( 0, \frac{1}{\kappa^*_{ijk}} - \overline{e^2_{ms_{ijk}}} \right). \tag{57}$$

(iii) At site level, the intersample variance  $\sigma^2_{ij}$  is deduced from the estimation of the squared error:

$$\frac{\hat{1}}{V_{ij}} = \frac{\hat{1}}{\sum_{k=1}^{r_{ij}} \frac{1}{\sigma^2_{ij} + \frac{\sigma^2_{Iijk} + e^2_{ms_{ijk}}}{n_{ijk}}}} = \frac{1}{r_{ij}} \left( \frac{S^{2*}_{Iij} + S^{2*}_{Dij}}{2} \right) \approx \frac{1}{r_{ij} \kappa^*_{ij}}. \tag{58}$$

The estimate  $\hat{\sigma}^2_{ij}$  of  $\sigma^2_{ij}$  is obtained by the bisection method (Press *et al.* 1997). If there is no positive solution, then the intersample variance estimate is set equal to zero. This variance estimated, zero or not, is used to recalculate the squared error  $\frac{\hat{1}}{V_{ij}}$  in eq. (58) and the mean values  $\overline{Y_{ij}}$  and  $\overline{Y_{Dij}}$  in eq. (36).

(iv) At field level, following Anderson (1984), the maximum likelihood unbiased estimation of the geomagnetic field  $Y_g(t_i)$  is given by the empirical means,

$$\hat{Y}_{gI}(t_i) = \overline{Y_{Ii}}, \quad \hat{Y}_{gD}(t_i) = \overline{Y_{Di}}, \tag{59}$$

and the terms of the squared error matrix  $\Sigma_i$  in eq. (43) is estimated from the empirical matrix  $S_i$  in eq. (45):

$$\hat{\Sigma}_i = \frac{S_i}{m_i - 1} = \frac{S_i^*}{m_i},$$

that is,

$$\frac{\hat{W}_{Di}}{W_{Ii} W_{Di} - W^2_{IDi}} = \frac{S^{2*}_{Di}}{m_i}, \quad \frac{\hat{W}_{Ii}}{W_{Ii} W_{Di} - W^2_{IDi}} = \frac{S^{2*}_{Di}}{m_i}, \quad \frac{\hat{W}_{IDi}}{W_{Ii} W_{Di} - W^2_{IDi}} = \frac{S^*_{IDi}}{m_i}. \tag{60}$$

Eq. (60) forms a non-linear system of three equations of which the resolution in the Bayesian context allows the variables  $x_1 = \sigma_{ii} g'_I$ ,  $x_2 = \sigma_{ii} g'_D$  and  $x_3 = \sigma_i^2$  to be estimated.

Let be  $x_1 = x_2 = x_3 = 0$  and  $W_i = \sum_{j=1}^{m_i} P_{ij} V_{ij}$ , where  $V_{ij}$  is given by eq. (58).

(a) If  $(\frac{1}{W_i} - \frac{S^{2*}_{Ii}}{m_i} \geq 0)$  and  $(\frac{1}{W_i} - \frac{S^{2*}_{Di}}{m_i} \geq 0)$  then the solution  $x_1 = x_2 = x_3 = 0$  is confirmed. These solutions correspond to Fig. 2 H3c ( $\sigma_i^2 = 0, \sigma^2_{ii} g'^2_I = 0, \sigma^2_{ii} g'^2_D = 0$ ).

(b) If  $(\frac{1}{W_i} - \frac{S^{2*}_{Ii}}{m_i} \geq 0)$  and  $(\frac{1}{W_i} - \frac{S^{2*}_{Di}}{m_i} < 0)$  then  $x_1 = x_3 = 0$  and  $x_2$  is the solution of  $(\sum_{j=1}^{m_i} \frac{1}{\sigma^2_{ii} g'^2_D + \frac{1}{V_{ij}}})^{-1} - \frac{S^{2*}_{Di}}{m_i} = 0$ , solved by the bisection method. These solutions correspond to Fig. 2 H3b ( $\sigma_i^2 = 0, \sigma^2_{ii} g'^2_I = 0, \sigma^2_{ii} g'^2_D \neq 0$ ).

(c) If  $(\frac{1}{W_i} - \frac{S^{2*}_{Ii}}{m_i} < 0)$  and  $(\frac{1}{W_i} - \frac{S^{2*}_{Di}}{m_i} \geq 0)$  then  $x_2 = x_3 = 0$  and  $x_1$  is solution of  $(\sum_{j=1}^{m_i} \frac{1}{\sigma^2_{ii} g'^2_I + \frac{1}{V_{ij}}})^{-1} - \frac{S^{2*}_{Ii}}{m_i} = 0$ , solved by the bisection method. These solutions correspond to Fig. 2 H3a ( $\sigma_i^2 = 0, \sigma^2_{ii} g'^2_I \neq 0, \sigma^2_{ii} g'^2_D = 0$ ).

(d) If  $(\frac{1}{W_i} - \frac{S^{2*}_{Ii}}{m_i} < 0)$  and  $(\frac{1}{W_i} - \frac{S^{2*}_{Di}}{m_i} < 0)$  then  $\sigma^2_{ii} g'^2_I + \sigma_i^2 \neq 0$  and  $\sigma^2_{ii} g'^2_D + \sigma_i^2 \neq 0$ . The variable  $x_3$  is first assumed equal to zero and there are always two solutions  $x_1$  and  $x_2$  obtained from the system of two equations:

$$\frac{W_{Di}}{W_{Ii} W_{Di} - W^2_{IDi}} = \frac{S^{2*}_{Ii}}{m_i}, \quad \frac{W_{Ii}}{W_{Ii} W_{Di} - W^2_{IDi}} = \frac{S^{2*}_{Di}}{m_i}$$

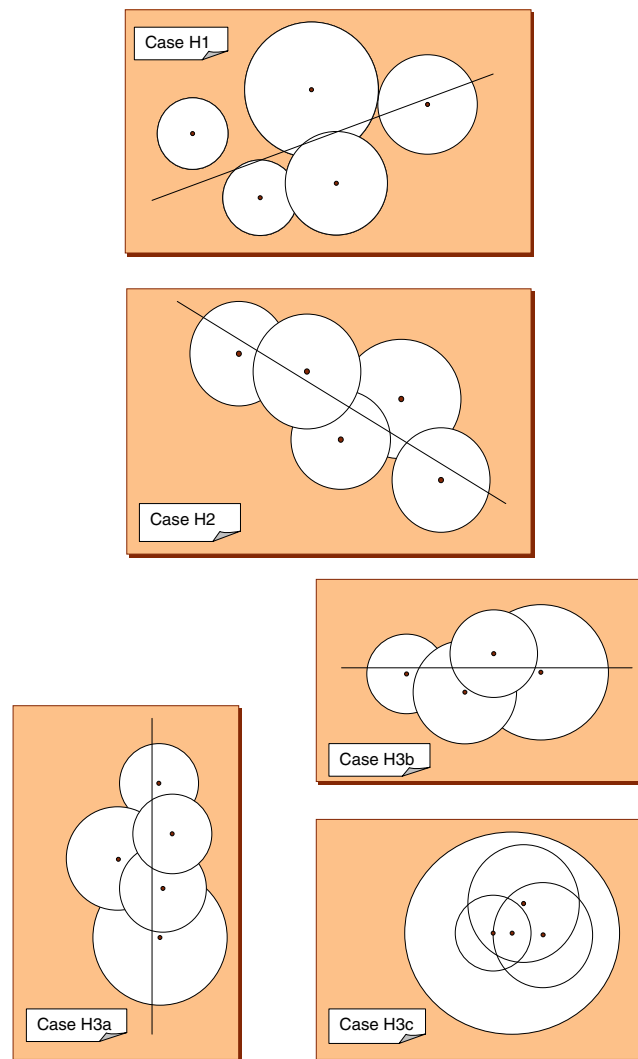
solved by the Newton–Raphson method with backtrack searching (Press *et al.* 1997).

(d1) If the difference  $(\frac{W_{IDi}}{W_{Ii} W_{Di} - W^2_{IDi}})^2 - (\frac{S^*_{IDi}}{m_i})^2 \leq 0$ , then  $x_3 = 0$  and  $x_1$  and  $x_2$  are kept. These solutions correspond to Fig. 2 H2 ( $\sigma_i^2 = 0, \sigma^2_{ii} g'^2_I \neq 0, \sigma^2_{ii} g'^2_D \neq 0$ ).

(d2) If the difference is positive, it means that  $|x_1|$  and  $|x_2|$  are too high. This implies that  $x_3$  is strictly positive. The three variables are then deduced by the resolution of the three equations in eq. (60) using the Newton–Raphson method with backtrack searching. These solutions correspond to Fig. 2 H1 ( $\sigma_i^2 \neq 0, \sigma^2_{ii} g'^2_I \neq 0, \sigma^2_{ii} g'^2_D \neq 0$ ).

Finally, knowing the three variances, the variance–covariance matrix  $\hat{\Sigma}_i$  in eq. (43), the mean vector  $\overline{Y}_i$  in eq. (38) and the Bayesian concentration factor in eq. (53) can be calculated.

The field direction estimate  $\hat{g}_I(t_i)$  and  $\hat{g}_D(t_i)$  in the geographic coordinate system is deduced from  $\hat{Y}_{gI}(t_i) = \overline{Y_{Ii}}$  and  $\hat{Y}_{gD}(t_i) = \overline{Y_{Di}}$ , using the inverse of rotation matrix in eq. (A1.6).



**Figure 2.** Hierarchical variances at field level (boxes represent the tangential plane to the sphere for the time window considered, with inclination upwards and declination perpendicular): each circle represents the mean direction determined for a given site with its confidence cone  $\alpha_{95ij}$ , the principal axis of the ellipse is shown by a line. (H1) all the variances are different from zero (see example in Table 5 and Fig. 4a), (H2) only intersite variance  $\hat{\sigma}_{ij}^2$  is zero, (H3a) intersite variance  $\hat{\sigma}_{ij}^2$  as well as declination curve variance  $\sigma_{Di}^2 g_D^2$  are zero but inclination curve variance  $\sigma_{Di}^2 g_I^2$  is not zero, (H3b) intersite variance as well as inclination curve variance are zero but declination curve variance is not zero, (H3c) all the intersite and curve variances are zero: the dispersion between the directions of sites is explained only by the high  $\alpha_{95ij}$  observed on each site (see example in Table 6 and Fig. 4b).

4.5.2 Particular solution when the data are concentrated at time  $t_i$

In the specific case where all the directions are of the same precise date  $t_i$ , that is  $P_{ij} = 1$  and  $\sigma_{ii}^2 = 0$ , the correlation becomes  $\rho_{ij} = 0$  ( $W_{Dij} = 0$ ). Some drastic simplifications occur in the exact solution, leading to

$$\hat{Y}_{gI}(t_i) = \overline{Y}_{Ii} = \frac{1}{W_i} \sum_{j=1}^{m_i} W_{ij} \overline{Y}_{Iij} \quad \text{and} \quad \hat{Y}_{gD}(t_i) = \overline{Y}_{Di} = \frac{1}{W_i} \sum_{j=1}^{m_i} W_{ij} \overline{Y}_{Dij}, \tag{61}$$

where  $W_i = \sum_{j=1}^{m_i} W_{ij}$ .

The weighting factors  $W_{Iij}$  and  $W_{Dij}$  in eq. (40) become equal to  $W_{ij}$  so that

$$\frac{1}{W_i} = \frac{1}{\sum_{j=1}^{m_i} W_{ij}} = \frac{1}{\sum_{j=1}^{m_i} \frac{1}{\sigma_i^2 + \frac{1}{V_{ij}}}} = \frac{1}{m_i K_{B0}} \tag{62}$$

(cf. eq. 55). From eq. (60),

$$\frac{\hat{1}}{W_i} = \frac{1}{m_i} \left( \frac{S_{Ii}^{2*} + S_{Di}^{2*}}{2} \right) \approx \frac{1}{m_i \kappa_i^*}. \tag{63}$$

Hence the exact Bayesian solution for  $\sigma_i^2$  is obtained with the bisection method.



**Table 2.** Values of the different confidence coefficients  $t_\beta$  (Fisher, bivariate normal, elliptic, Student) having the probability  $\beta = 0.05$  to be exceeded for increasing numbers  $m_i$ .

$m_i$	$t_{\beta F}$ Fisher	$t_{\beta N}$ WIDi = 0	$t_{\beta H}$ ellipse	$t_{\beta/2}$ Student
2	6.16	24.50	$\infty$	12.71
3	3.73	6.40	28.25	4.30
4	3.21	4.36	7.55	3.18
5	2.99	3.67	5.05	2.78
6	2.86	3.36	4.17	2.57
7	2.79	3.17	3.73	2.45
8	2.73	3.04	3.46	2.37
9	2.70	2.95	3.29	2.31
10	2.67	2.88	3.17	2.26
30	2.51	2.57	2.63	2.05
100	2.47	2.49	2.50	1.99
$\infty$	2.45	2.45	2.45	1.96

The Bayesian geomagnetic field estimate simplifies to

$$\hat{Y}_{gI}(t_i) = \overline{Y}_{Ii} = \frac{1}{W_i} \sum_{j=1}^{m_i} \frac{1}{\sigma_i^2 + \frac{1}{V_{ij}}} \overline{Y}_{Iij} \quad \text{and} \quad \hat{Y}_{gD}(t_i) = \overline{Y}_{Di} = \frac{1}{W_i} \sum_{j=1}^{m_i} \frac{1}{\sigma_i^2 + \frac{1}{V_{ij}}} \overline{Y}_{Dij}. \tag{64}$$

Thus, the Bayesian mean appears to be tempered by the intersite variance  $\sigma_i^2$ . The Bayesian approach brings out a solution to the problem of weighting the archaeomagnetic or palaeomagnetic data, and in particular allows for solving some contradictions, outlined in Section 3.1.3, concerning the intuitive classical weighting techniques. The relationships between the mean observations, the variances resulting from the sources of scatter and the sampling errors are clearly established. Consequently, some important information can be obtained on the sampling numbers needed (Section 4.6).

#### 4.5.3 Descending hierarchy

If the estimates  $\hat{\sigma}_{ijk}^2$  and  $\hat{\sigma}_{ij}^2$  are strictly positive in eqs (57) and (58), then there is a descending hierarchy. That is, squared errors at a given level are less than the variance (resulting from sources of scatter) at an upper level. Consequently, at site level, we have

$$\frac{\hat{1}}{V_{ij}} = \frac{1}{r_{ij} \kappa_{ij}^*} = \frac{\alpha_{95ij}^2}{2.45^2}. \tag{65}$$

Such a situation can be achieved using enough measurements and/or specimens in the analyses. This seems to be often verified in practice. This important situation justifies the results at sample and lower levels not being published.

#### 4.5.4 Comparison with stratification approach

The hierarchical mean calculation can be compared with the stratification approach described in Section 3.1 in two cases.

(a) If the descending hierarchy is not verified at a field level, that is dispersion between sites is small in comparison with intrasite dispersions, then the estimations are  $\sigma_i^2 = 0$ ,  $\sigma_{ii}^2 g_I^2 = 0$  and  $\sigma_{ii}^2 g_D^2 = 0$  (Fig. 2 H3C) and eq. (59) becomes of the form

$$\hat{Y}_{gI}(t_i) = \overline{Y}_{Ii} = \frac{1}{W_i} \sum_{j=1}^{m_i} \frac{P_{ij}}{\alpha_{95ij}^2} \overline{Y}_{Iij} = \frac{1}{W_i} \sum_{j=1}^{m_i} P_{ij} r_{ij} \kappa_{ij}^* \overline{Y}_{Iij}.$$

This is equivalent to Kovacheva's eqs (15) and (17) in Section 3.1.3. Consequently, this statistics is well adapted to bad data.

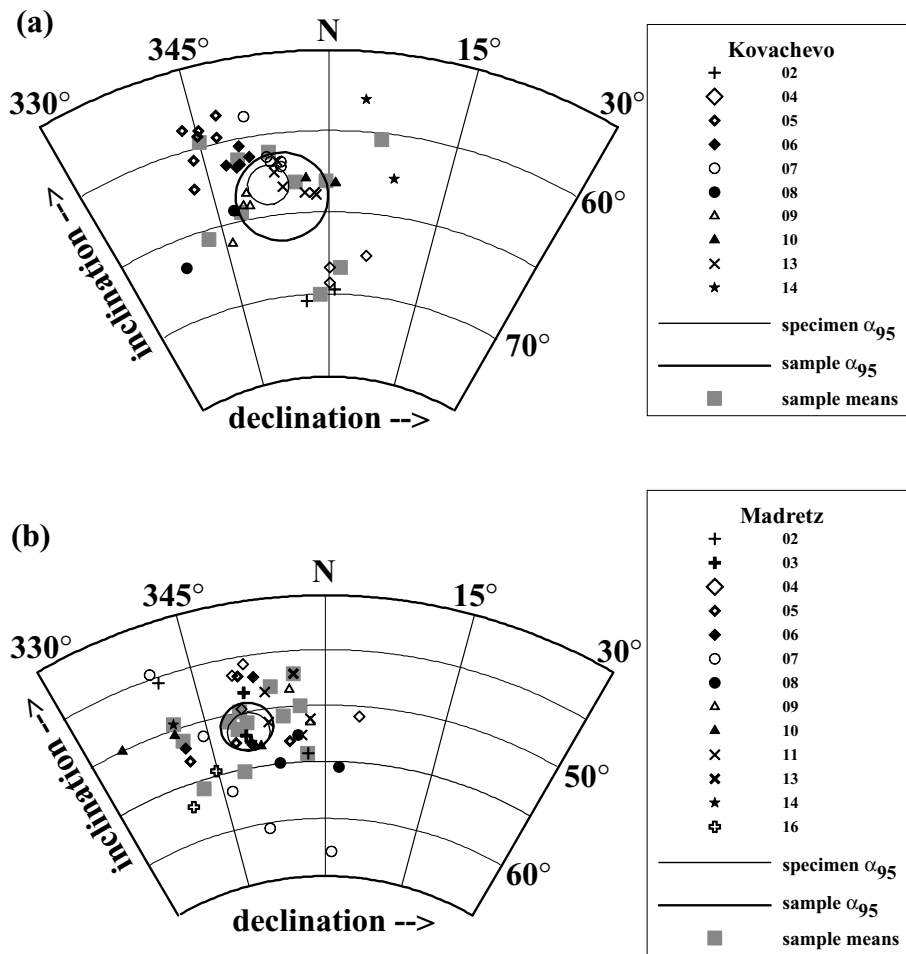
(b) If all the  $\alpha_{95ij}^2$  are equal or zero, then eq. (59) becomes always of the form

$$\hat{Y}_{gI}(t_i) = \overline{Y}_{Ii} = \frac{1}{W_i} \sum_{j=1}^{m_i} P_{ij} \overline{Y}_{Iij}.$$

This is equivalent to the Le Goff statistics in eq. (A2.16). This last statistics is well adapted to homogeneous data (bad or not).

### 4.6 Estimation of the sampling numbers

The main idea that appears here is that the precision on geomagnetic secular variation curves is essentially controlled by the number  $m_i$  of sites and, more generally, the errors in archaeomagnetism (as well as in palaeomagnetism) are dominated by errors in age (Tarling & Dobson 1995) through the weighting  $P_{ij}$  and variance  $\sigma_{ii}^2$ .



**Figure 3.** Hierarchy versus stratification at sample and specimen levels: the data from Tables 3 and 4 are plotted in stereographic projection. Different symbols indicate the specimens from each independently oriented sample. Comparison is shown between the  $\alpha_{95ij}$  calculated over the total number of specimens measured (given with different symbols for each independently oriented sample: thin circle) in the stratification context, and over the number of the samples in the hierarchical context (after grouping of the specimens for each sample: thick circle).

4.6.1 The number of samples and specimens needed in archaeomagnetic studies

The effects of the magnitudes of  $n_{ijk}$  and  $r_{ij}$  are evaluated in the simple case when the variances are of the same order at each level, that is  $\sigma_{ijk}^2 = \sigma_{ij}^2 = \sigma_i^2 = \sigma^2 = \frac{1}{K}$  and when the errors  $e_{ms_{ijk}}^2$  are small. In order to make the discussion easier,  $\sigma_{ii}^2 g_I^2$  and  $\sigma_{ii}^2 g_D^2$  are approximated in eq. (40) by

$$\sigma_{Ci}^2 = \frac{1}{K_{Ci}} = \frac{\sigma_{ii}^2 g_I^2 + \sigma_{ii}^2 g_D^2}{2} \quad \text{and} \quad P_{ij} \text{ by } \bar{P}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} P_{ij},$$

and taking the same number of specimens  $n_{ijk} = n_i$  from each sample and the same number of samples  $r_{ij} = r_i$  from each site. Then, the Bayesian concentration factor from eq. (53) becomes

$$\frac{1}{K_B} \approx \frac{1}{\bar{P}_i} \left\{ \frac{1}{K_{Ci}} + \frac{1}{K} \left( 1 + \frac{1}{r_i} + \frac{1}{r_i n_i} \right) \right\}. \tag{66}$$

This clearly relates the Bayesian concentration factor  $K_B$  to the sampling numbers and to the dating errors. If  $n_i$  and  $r_j$  tend to infinity (measuring the entire site!) and if all the dating ranges tend to zero, then the maximum value  $K$  for the Bayesian concentration factor will be attained.

The number of samples  $r_i$  such that the confidence angle  $\alpha_{95i} = t_{\beta F}(m_i) / \sqrt{m_i K_B}$  is near the optimum confidence cone  $\alpha_{95i(\text{optim.})} = t_{\beta F}(m_i) / \sqrt{m_i K}$ , at percentage  $p = \frac{\alpha_{95i} - \alpha_{95i(\text{optim.})}}{\alpha_{95i(\text{optim.})}}$ , is given by

$$r_i = \frac{1}{[\bar{P}_i(1+p)^2 - (1+D_i)]} \left( 1 + \frac{1}{n_i} \right), \tag{67}$$

where the factor  $D_i = \frac{K}{K_{Ci}}$  is introduced.

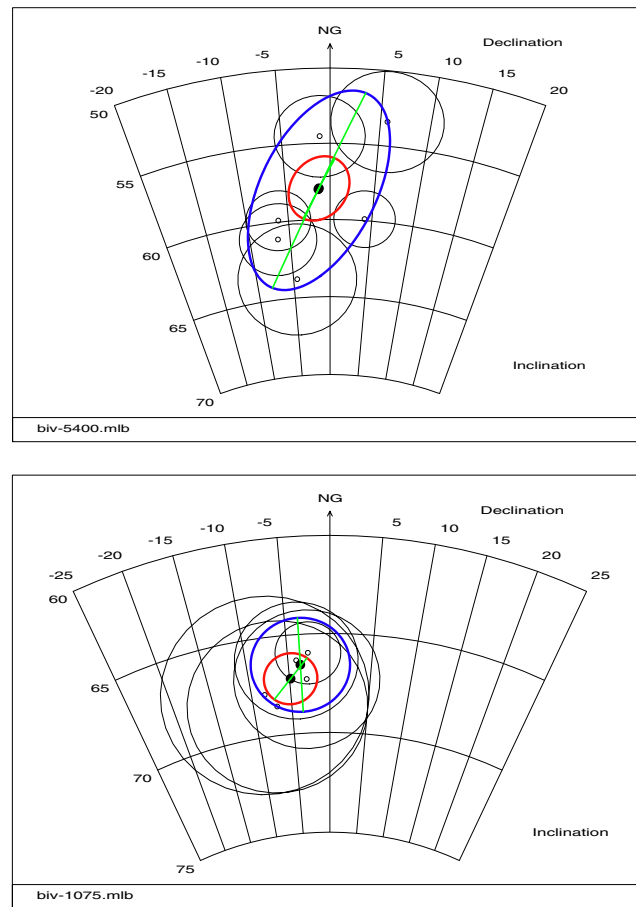
**Table 3.** Kovachevo site: measured declinations and inclinations listed for specimens and the corresponding Fisherian sample mean. At the bottom, a comparison between the results of Fisher statistics and bivariate Gaussian statistics with stratification and with hierarchy is given. In the latter case, the calculation was performed using a rotation (Appendix A1) in order to approximate Fisher statistics by Gaussian bivariate statistics. The results confirm the very good Gaussian bivariate approximation.

Kovachevo site (Bulgaria)					
Field number	40 specimens		Field number	10 samples	
	$\overline{D}_{ijkl}$ (°)	$\overline{I}_{ijkl}$ (°)		$\overline{D}_{ijk}$ (°)	$\overline{I}_{ijk}$ (°)
Ko2a	0.9	69.7	Ko2	358.5	70
Ko2b	356.1	70.4			
Ko4a	5.7	67.6	Ko4	1.9	68.4
Ko4b	0.1	69.3			
Ko4c	0.1	68.4			
Ko5a	342.9	62.4	Ko5	345	59.7
Ko5b	347.5	58.3			
Ko5c	347.1	59.7			
Ko5d	345	59.3			
Ko5e	343.8	60.7			
Ko5f	345.3	59			
Ko5g	343.5	58.7			
Ko6a	350.4	61.2	Ko6	349	61.3
Ko6b	348.7	61.7			
Ko6c	347.5	61.5			
Ko6d	349.4	60.5			
Ko6e	349.1	61.6			
Ko7a	354.1	62.1	Ko7	352.8	61.1
Ko7b	354.1	61.8			
Ko7c	352.4	61.4			
Ko7d	350.5	58.7			
Ko7e	352.8	61.7			
Ko8a	338.4	66.8	Ko8	342.8	65.6
Ko8b	347.2	64.3			
Ko9a	347.4	64.4	Ko9	348.1	64.5
Ko9b	346	66.2			
Ko9c	349.4	64.2			
Ko9d	348.6	64.1			
Ko9e	349.2	63.4			
Ko10a	0.9	63.2	Ko10	359.6	63.1
Ko10b	0.9	63.2			
Ko10c	357.1	62.9			
Ko13a	353.7	61.7	Ko13	355.7	63.1
Ko13b	353.1	62.4			
Ko13c	354	63.3			
Ko13d	358.1	63.8			
Ko13e	356.8	63.8			
Ko13f	358.4	63.9			
Ko14a	8.3	62.7	Ko14	6.2	60.4
Ko14b	4.1	58			

Stratification		Hierarchy
	Fisher statistics	
$\overline{I}_{ij} = 63.11$	$\overline{D}_{ij} = -7.79$	$\overline{I}_{ij} = 63.93$
$\kappa_{ij}^* = 344.69$	$\alpha_{95ij} = 1.22$	$\overline{D}_{ij} = -6.19$
	Gaussian statistics (using rotation: $\lambda = -63.106, \phi = -7.786$ )	$\alpha_{95ij} = 2.88$
$\overline{I}_{ij} = 63.11$	$\overline{D}_{ij} = -7.78$	$\overline{I}_{ij} = 63.93$
$\frac{S_{Iij}^{2*} + S_{Dij}^{2*}}{2} = 344.16$	$\alpha_{95ij} = 1.22$	$\overline{D}_{ij} = -6.19$
		$\frac{S_{Iij}^{2*} + S_{Dij}^{2*}}{2} = 278.72$
		$\alpha_{95ij} = 2.89$

Downloaded from https://academic.oup.com/gji/article-abstract/160/2/440/658812 by guest on 30 July 2020



**Figure 4.** Examples of hierarchy versus stratification at field level. (a) Exact elliptic bivariate statistics (bold circle) and Le Goff statistics (grey circle) from Table 5. This corresponds to case H1 in Fig. 2. (b) Exact elliptic bivariate statistics (bold circle) and Le Goff statistics (grey circle) from Table 6. This corresponds to case H3c in Fig. 2. The hierarchical confidence ellipse is always larger than that with stratification.

In the case of  $n_i = 1$  and dates at  $t_i$  (no dating error), that is  $P_{ij} = 1$  and  $\sigma_{ii}^2 = 0$ ,  $D_i = 0$  and  $\bar{P}_i = 1$ , then  $r_i = 10$  samples are needed for obtaining the precision  $p = 10$  per cent and  $r_i = 20$  samples for obtaining the precision  $p = 5$  per cent (of course, the number of samples needed will increase with the dating factor  $\bar{P}_i$ ). This analysis gives the range of the number of samples needed in palaeomagnetic and archaeomagnetic studies, provided that the variances at the different levels resulting from the sources of scatter are of the same order (this assumption is in fact very wide). This assessment is higher than that of 6 to 12 samples proposed by Tarling (1983, p. 140).

Moreover, considering sampling strategy where the entire, large sample is measured, that is  $n_i$  is like infinity (measurement strategy in Paris for instance), and sampling strategy with one specimen measured per sample, i.e.  $n_i = 1$  (measurement strategy in Rennes), then at least twice as many samples need to be taken with the second strategy than with the first one in order to obtain the same estimation of  $1/K_B$ .

#### 4.6.2 Post-selection of sites having small $\alpha_{95ij}$

Considering the case when samples are entirely measured, that is  $n_{ijk} = \infty$ , and when the sites are dated at the same time  $t_i$  ( $P_{ij} = 1$  and  $\sigma_{ii}^2 = 0$ ), then, from eq. (55), the Bayesian estimation error on the mean direction of the magnetic field is

$$\alpha_{95i} = \frac{t_{\beta F}}{\sqrt{m_i K_{B0}}} = \frac{t_{\beta F}}{\sqrt{\sum_{j=1}^{m_i} \frac{1}{k_i + r_{ij} k_{ij}}}}. \tag{68}$$

This error will always decrease or remain constant with the increase of the number of sites  $m_i$ , whatever the squared errors  $\frac{1}{r_{ij} k_{ij}} \propto \alpha_{95ij}^2$  at a site level are. This means that a post-selection of the weaker  $\alpha_{95ij}^2$  in a database never improves the precision of the geomagnetic estimate. Provided that the errors resulting from the sources of scatter are random, there is no justification in eliminating bad  $\alpha_{95ij}^2$  from a database. The Bayesian calculation automatically does this. However, if a very bad  $\alpha_{95ij}^2$  is confirmed to be associated with systematic error, then the variance  $\frac{1}{k_i}$  will become high and the Bayesian error cone  $\alpha_{95i}^2$  will increase. Consequently, in this specific case of systematic error, the datum has to be corrected or excluded from the database.

Table 4. Madretz site (see caption of Table 3).

Madretz site (Bulgaria)					
Field number	37 specimens		Field number	13 samples	
	$\overline{D}_{ijkl}$ (°)	$\overline{I}_{ijkl}$ (°)		$\overline{D}_{ijk}$ (°)	$\overline{I}_{ijk}$ (°)
M2a	341.1	45.5	M2	349.4	49.9
M2b	357.7	54.3			
M3a	350.2	52.7	M3	350.1	51
M3b	349.7	52.1			
M3c	350.3	48.3			
M4a	349.7	49.7	M4	353.5	48.1
M4b	349.3	46.6			
M4c	4.3	50.9			
M4d	350.8	45.7			
M5a	348.2	52.6	M5	348.8	51.4
M5b	355.3	53.1			
M5c	341.8	53.1			
M5d	349.9	46.7			
M6a	341.7	51.9	M6	347.9	50.6
M6b	350.3	53			
M6c	351.7	47			
M7a	350.9	60.5	M7	348.6	55.2
M7b	346.3	56.7			
M7c	1.1	62.9			
M7d	344.3	51.3			
M7e	340.5	44.6			
M8a	353.7	54.9	M8	357.6	54.3
M8b	1.9	55.5			
M8c	356.4	52.6			
M9a	358.1	51.5	M9	356.9	50
M9b	355.7	48.5			
M10a	334.2	50	M10	341.7	51.2
M10b	340.9	50.5			
M10c	351.4	53.2			
M11a	352.8	51.2	M11	354.7	50.8
M11b	352.8	48.5			
M11c	357	52.6			
M11d	358.1	51.2			
M13	356.3	47.1	M13	356.3	47.1
M14	341.2	49.5	M14	341.2	49.5
M16a	340.1	57	M16	342.4	55.8
M16b	344.7	54.6			

Stratification

Hierarchy

Fisher statistics

$$\overline{I}_{ij} = 51.76$$

$$\kappa_{ij}^* = 189.80$$

$$\overline{D}_{ij} = -9.86$$

$$\alpha_{95ij} = 1.71$$

$$\overline{I}_{ij} = 51.39$$

$$\kappa_{ij}^* = 341.19$$

$$\overline{D}_{ij} = -10.09$$

$$\alpha_{95ij} = 2.24$$

Gaussian statistics (using rotation:  $\lambda = -51.757, \phi = -9.855$ )

$$\overline{I}_{ij} = 51.76$$

$$\frac{S_{Iij}^{2*} + S_{Dij}^{2*}}{2} = 189.32$$

$$\overline{D}_{ij} = -9.86$$

$$\alpha_{95ij} = 1.71$$

$$\overline{I}_{ij} = 51.39$$

$$\frac{S_{Iij}^{2*} + S_{Dij}^{2*}}{2} = 340.54$$

$$\overline{D}_{ij} = -10.09$$

$$\alpha_{95ij} = 2.25$$

**Table 5.** Six reference points taken from the Bulgarian database (Kovacheva 1997), from window  $ti = 5400 \text{ BC} \pm 50 \text{ yr}$ , without geographical correction. The field direction estimates  $\hat{g}_I(t_i)$  and  $\hat{g}_D(t_i)$  in the geographic coordinate system were calculated from the polar estimates  $\hat{Y}_{gI}(t_i) = \overline{Y}_{Ii}$  and  $\hat{Y}_{gD}(t_i) = \overline{Y}_{Di}$ , using the inverse of the rotation matrix in eq. (A1.6), with a rotation  $(\lambda, \phi) = (-57.966, -1.243)$ . Errors  $e_{\hat{g}_I}$  and  $e_{\hat{g}_D}$  were calculated from eqs (90) and (91). (a)  $t_{ij1}$  and  $t_{ij2}$ , date interval;  $r_{ij}$ , number of samples;  $\overline{I}_{ij}$ , empirical mean inclination;  $\overline{D}_{ij}$ , empirical mean declination;  $\kappa_{ij}^*$ , unbiased Bayesian concentration factor (eq. A1.5);  $P_{ij}$ , weight for dating error (eq. 28);  $\alpha_{95ij}$ , Fisher confidence cone at 95 per cent;  $\overline{Y}_{Iij}$ ,  $\overline{Y}_{Dij}$ , polar inclination and declination after polar rotation. (b)  $S_{Ii}^{2*}$ ,  $S_{Di}^{2*}$ ,  $S_{ID}^{2*}$ , empirical variances and covariance;  $\hat{\sigma}_i^2$ , intersite variance estimate;  $\hat{\sigma}_{ii} g'_I$ ,  $\hat{\sigma}_{ii} g'_D$ , curve variance estimates for inclination and declination;  $K_x$ ,  $K_y$ , elliptic concentration factors along the maximal and minimal axes;  $\Omega$ , orientation angle of the principal axis of the error ellipse with respect to the meridian passing through the mean direction;  $\alpha_{95x}$ ,  $\alpha_{95y}$ , elliptic confidence cones;  $K_B$ , equivalent Fisher concentration factor (eq. 53). (c) Directional results in the geographical coordinate system,  $\overline{I}_i = \hat{g}_I$ ,  $\overline{D}_i = \hat{g}_D$ ; error estimates of inclination and declination,  $e_{\hat{g}_I}$ ,  $e_{\hat{g}_D}$ .

(a) Raw data

No.	$t_{ij1}$	$t_{ij2}$	$r_{ij}$	$\overline{I}_{ij}$	$\overline{D}_{ij}$	$\kappa_{ij}^*$	$P_{ij}$	$\alpha_{95ij}$	$\overline{Y}_{Iij}$	$\overline{Y}_{Dij}$
19	-5420	-5220	17	53.40	5.70	113	0.350	3.20	-4.353	4.145
13	-5470	-5370	11	54.53	-1.06	285	0.800	2.50	-3.436	0.106
5	-5490	-5390	24	61.12	-6.50	160	0.600	2.26	3.253	-2.540
10	-5520	-5420	29	59.90	4.13	209	0.300	1.80	2.041	2.693
262	-5530	-5430	28	59.90	-6.20	194	0.200	1.90	5.869	-1.445
261	-5590	-5410	10	63.80	-4.50	180	0.222	3.30	2.025	-2.485

Polar rotation:  $\lambda = -57.966$ ,  $\phi = -1.243$

(b) Bivariate results in polar coordinate system ( $Y_I, Y_D$ )

	$S_{Ii}^{2*}$	$S_{Di}^{2*}$	$S_{ID}^{2*}$	$\hat{\sigma}_i^2$	$\hat{\sigma}_{ii} g'_I$	$\hat{\sigma}_{ii} g'_D$	$K_x$	$K_y$	$\Omega$	$\alpha_{95x}$	$\alpha_{95y}$	K or $K_B$
<b>LeGoff Statistics</b>							<b>98.800</b>	<b>158.949</b>	<b>153.001</b>	<b>2.090</b>	<b>1.648</b>	<b>121.856</b>
Approxim. Le Goff	13.004	5.570	-5.181				98.296	159.012	152.830	2.095	1.647	121.491
<b>Bayesian bivariate</b>	<b>14.505</b>	<b>6.421</b>	<b>-4.761</b>	<b>0.756</b>	<b>-0.947</b>	<b>2.047</b>	<b>196.472</b>	<b>778.340</b>	<b>155.165</b>	<b>6.953</b>	<b>3.493</b>	<b>313.746</b>

(c) Bivariate and univariate results in geographical coordinate system ( $I, D$ )

	$\overline{I}_i = \hat{g}_I$	$\overline{D}_i = \hat{g}_D$	err $\hat{g}_I$	err $\hat{g}_D$
<b>LeGoff Statistics</b>	<b>57.964</b>	<b>-1.252</b>	<b>1.605</b>	<b>2.636</b>
Approximate Le Goff	57.966	-1.242		
<b>Bayesian bivariate</b>	<b>58.022</b>	<b>-1.317</b>	<b>3.998</b>	<b>5.022</b>
Kovacheva's calculation	58.368	-1.687	3.671	5.172
Sternberg's calculation	57.508	-1.588	3.998	4.488
Batt's calculation	58.118	-1.651	3.914	5.128

#### 4.6.3 Case of the common specimen = sample method

In sampling techniques used in the US, Canada, UK, Belgium or Paris, specimens and samples are identical. This means that there is no subsampling. From a statistical point of view, the analysis passes from the specimen level to the field level by  $\sigma_{ijk}^2 = 0$  and  $n_{ijk} = 1$  (the specimen is the sample). The squared error at site level in eq. (37) becomes

$$\frac{1}{V_{ij}} = \frac{1}{\sum_{k=1}^{r_{ij}} \frac{1}{\sigma_{ij}^2 + e_{ms_{ijk}}^2}} \approx \frac{\sigma_{ij}^2 + \overline{e_{ms_{ij}}^2}}{r_{ij}}, \tag{69}$$

where the mean squared error  $\overline{e_{ms_{ijk}}^2}$  in eq. (35) becomes  $\overline{e_{ms_{ij}}^2} = \frac{1}{r_{ij}} \sum_{k=1}^{r_{ij}} \frac{\sigma_{ijk(t=1)}^2}{d_{ijk(t=1)}}$ .

**Table 6.** Five synthetic points well grouped with low  $\kappa_{ij}^*$ . This example is taken from Bucur's (1994) database, at  $t_i = -75 \pm 120$  yr, with geographical correction (VGP correction), but with all  $P_{ij} = 1$  and  $\kappa_{ij}^*$  divided by 10 for didactic reasons. See caption of Table 5.

(a) Raw data

No.	$t_{i1}$	$t_{i2}$	$r_{ij}$	$\bar{I}_{ij}$	$\bar{D}_{ij}$	$\kappa_{ij}^*$	$P_{ij}$	$\alpha_{95ij}$	$\bar{Y}_{ij}$	$\bar{Y}_{Dij}$
358	-120	-80	45	65.96	-2.58	182.5	1.000	1.55	-1.233	0.932
307	-100	-80	11	67.28	-2.87	169.6	1.000	3.25	0.082	0.772
308	-100	-50	10	67.88	-8.33	93.1	1.000	4.60	0.706	-1.303
306	-80	-40	14	66.31	-4.01	179.7	1.000	2.80	-0.898	0.345
4	-40	-40	8	68.54	-6.91	162.9	1.000	3.88	1.342	-0.747

Polar rotation:  $\lambda = -67.210$   $\phi = -4.869$

(b) Bivariate results in polar coordinate system ( $Y_I, Y_D$ )

	$S_{Ii}^{2*}$	$S_{Di}^{2*}$	$S_{ID}^*$	$\hat{\sigma}_i^2$	$\hat{\sigma}_{ti} g'_I$	$\hat{\sigma}_{ti} g'_D$	$K_x$	$K_y$	$\Omega$	$\alpha_{95x}$	$\alpha_{95y}$	K or $K_B$
<b>LeGoff Statistics</b>							<b>138.565</b>	<b>146.725</b>	<b>138.515</b>	<b>1.271</b>	<b>1.235</b>	<b>142.528</b>
<i>Approxim. Le Goff</i>	0.927	0.768	-0.664				138.394	146.659	138.410	1.272	1.236	142.407
<b>Bayesian bivariate</b>	<b>0.913</b>	<b>0.571</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>2965.62</b>	<b>2965.62</b>	<b>0.0</b>	<b>2.375</b>	<b>2.375</b>	<b>2965.62</b>

(c) Bivariate and univariate results in geographical coordinate system ( $I, D$ )

	$\bar{I}_i = \hat{g}_I$	$\bar{D}_i = \hat{g}_D$	$err \hat{g}_I$	$err \hat{g}_D$
<b>LeGoff Statistics</b>	<b>67.207</b>	<b>-4.858</b>	<b>1.004</b>	<b>2.584</b>
<i>Approximate Le Goff</i>	67.210	-4.869		
<b>Bayesian bivariate</b>	<b>66.542</b>	<b>-3.551</b>	<b>1.306</b>	<b>3.281</b>
<i>Kovacheva's calculation</i>	66.533	-3.601	1.185	2.420
<i>Sternberg's calculation</i>	67.084	-4.543	1.353	2.916
<i>Batt's calculation</i>	66.863	-4.240	1.329	2.910

## 5 FROM HIERARCHICAL STATISTICS TO MOVING AVERAGE TECHNIQUE

A knowledge of the estimation precision of the curves is essential for dating purposes (Lanos *et al.* 1999) as well as for the time varying spherical harmonic analyses in geomagnetism (Hongre *et al.* 1998). In this section, the confidence (or error) ellipses, cones and intervals are determined on the geomagnetic parameters of interest (direction, or marginal parameter: inclination or declination) and the plotting technique of the curves is described. The case of intensity is treated in Appendix A3.

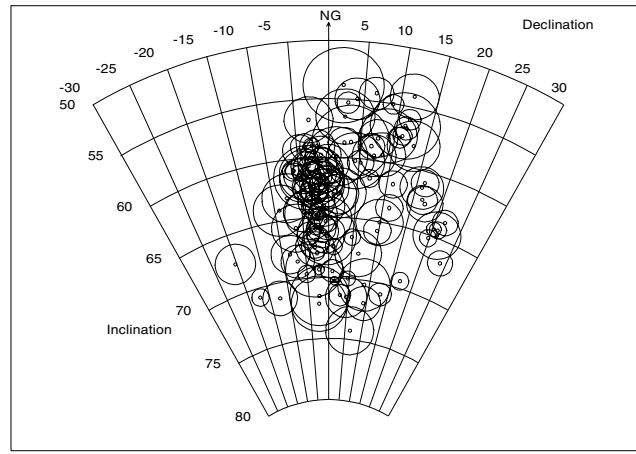
### 5.1 Confidence interval for the mean direction

From eqs (42) and (60), it is possible to define a confidence ellipse for the vector  $Yg(t_i)$  by considering the quantity:

$$T^2 = m_i (\bar{Y}_i - Y_g(t_i))^T S_i^{*-1} (\bar{Y}_i - Y_g(t_i)). \tag{70}$$

$T^2$  is Hotelling distributed (Anderson 1984) and, in the specific case of two dimensions (here inclination  $\bar{Y}_{Ii}$  and declination  $\bar{Y}_{Di}$ ), then  $\frac{m_i - 2}{2(m_i - 1)} T^2$  is  $F(2, m_i - 2)$  (Fisher) distributed, that is

$$p(T^2) = \frac{(m_i - 2)}{2(m_i - 1) \left(1 + \frac{1}{m_i - 1} T^2\right)^{m_i/2}}. \tag{71}$$



**Figure 5.** Bucur’s database (1994): 119 well dated reference points (labelled PC in Table 1, Bucur 1994). The raw mean directions (inclination, declination) are plotted with  $\alpha_{95ij}$  confidence cones, between 200 BC and AD 1900.

The cumulative distribution is given by

$$p(0 \leq T^2 \leq t_{\beta H}^2) = 1 - \left(1 + \frac{1}{m_i - 1} t_{\beta H}^2\right)^{-\frac{m_i - 2}{2}}. \tag{72}$$

If  $\beta$  is the probability of being outside the confidence region, then it is possible to determine  $t_{\beta H}$  such that  $p(0 \leq T^2 \leq t_{\beta H}^2) = 1 - \beta$ :

$$t_{\beta H}^2 = (m_i - 1) [\beta^{-2/(m_i - 2)} - 1]. \tag{73}$$

These results permit a confidence ellipse to be defined in the tangential plane to the sphere by making the following change of variables:

$$\overline{Y}_{Ii} - g_I = \delta \cos \varphi,$$

$$\overline{Y}_{Di} - g_D = \delta \sin \varphi.$$

Letting

$$b_\varphi = \hat{W}_{Ii} \cos^2 \varphi - 2 \hat{W}_{IDi} \sin \varphi \cos \varphi + \hat{W}_{Di} \sin^2 \varphi, \tag{74}$$

then  $T^2 = \delta^2 b_\varphi$ , where  $\delta$  characterizes the size of the error ellipse and  $b_\varphi$  characterizes the shape of the ellipse. The off-set error angle  $\delta_\beta$  of the ellipse at  $(1 - \beta)$  confidence level will be given by

$$\delta_\beta = t_{\beta H} \frac{1}{\sqrt{b_\varphi}}, \tag{75}$$

where

$$t_{\beta H} = \sqrt{(m_i - 1) [\beta^{-2/(m_i - 2)} - 1]} \underset{m_i \rightarrow \infty}{\approx} \sqrt{-2 \ln(\beta)}. \tag{76}$$

$\delta_\beta$  depends on the angle  $\varphi$  which describes the ellipse.

From eqs (48) and (51), the confidence ellipse centred on the mean direction will be characterized by the angles

$$\alpha_{95ix} = t_{\beta H} / \sqrt{m_i K_x} = t_{\beta H} / \sqrt{\frac{1}{\cos(2\Omega)} [W_{Ii} \cos^2 \Omega - W_{Di} \sin^2 \Omega]}, \tag{77}$$

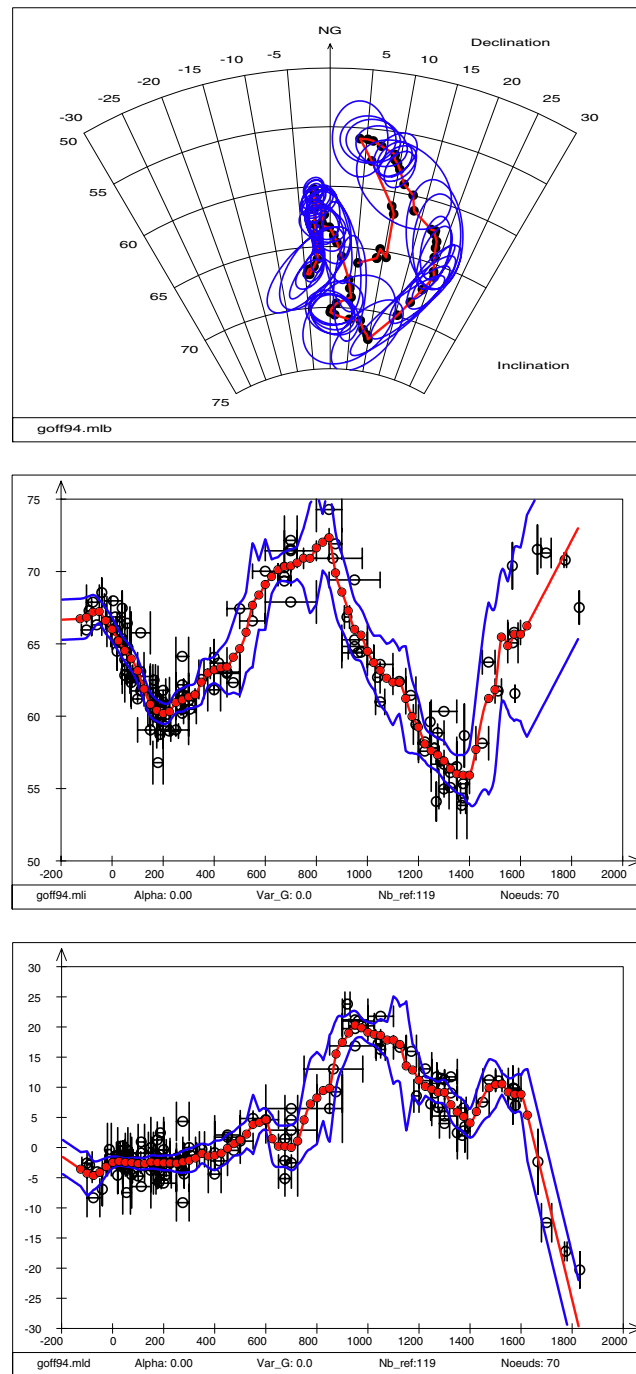
$$\alpha_{95iy} = t_{\beta H} / \sqrt{m_i K_y} = t_{\beta H} / \sqrt{\frac{1}{\cos(2\Omega)} [W_{Di} \cos^2 \Omega - W_{Ii} \sin^2 \Omega]}, \tag{78}$$

defined according to the principal axes  $(x, y)$  of the error ellipse.

### 5.1.1 Particular case: the data are concentrated at time $t_i$

In the specific case where all the directions are of the same precise date  $t_i$ , that is  $P_{ij} = 1$  and  $\sigma_{ii}^2 = 0$ , then  $W_{Di} = 0$ . Hence the variate  $T^2$  is distributed as the sum of two Student–Fisher variates squared. The probability density function is



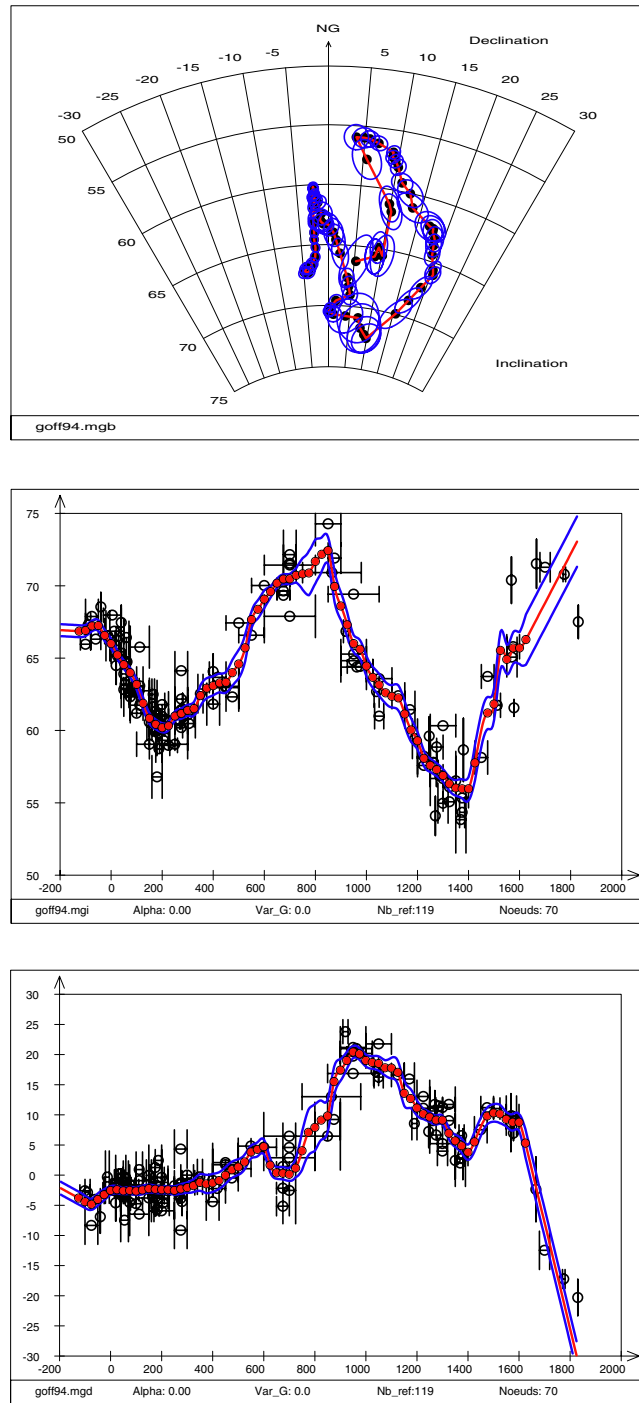


**Figure 6.** Bucur's database, hierarchical bivariate statistics: overlapping windows of width 100 yr plotted at 25-yr steps with at least three reference points per window. (a) Stereographic diagram (inclination declination) with ellipse at the 95 per cent confidence level, (b) marginal inclination curve with error band at 95 per cent, (c) marginal declination curve with error band at 95 per cent.

$$p(T^2) = 2 \frac{\Gamma^2\left(\frac{m_i}{2}\right)}{\Gamma^2\left(\frac{m_i-1}{2}\right)} [2(m_i - 1)]^{m_i-1} \sum_{n=0}^{\infty} \frac{\left(\frac{m_i}{2}\right)_n \left(\frac{1}{2}\right)_n}{(1)_n n!} \cdot \frac{T^{2n}}{[2(m_i - 1) + T^2]^{m_i+2n}}, \quad (79)$$

where  $(x)_0 = 1$  and  $(x)_n = \Gamma(x + 1)/\Gamma(x)$ ,  $\Gamma(x)$  being the gamma function.

It is possible to define a  $(1 - \beta)$  confidence size  $\delta_\beta$  by integrating with respect to  $T^2$ . This size will be equivalent to  $\alpha_{95}$  (Fisher statistics) when  $m_i$  is large enough. Finally, the  $t_{\beta N}$  coefficient, factor of the error  $1/\sqrt{b_\varphi}$  can be calculated. This gives the size of the confidence cone about the mean direction, at a 95 per cent confidence level, with respect to the number  $m_i$  of sites. For comparison, Table 2 gives the coefficients for the three bivariate cases, which are the Fisher statistics ( $t_{\beta F}$ ), the bivariate normal distribution ( $t_{\beta N}$ ) and the Hotelling distribution ( $t_{\beta H}$ ), and for Student t-distribution ( $t_{\beta/2}$ ).



**Figure 7.** Bucur's database, stratified bivariate statistics (Le Goff's method): overlapping windows of width 100 yr plotted at 25-yr steps with at least three reference points per window. (a) Stereographic diagram (inclination declination) with ellipse at 95 per cent confidence level, (b) and (c) marginal inclination and declination curves, with corresponding confidence bands.

Recall that  $t_{\beta F}$  for Fisher (1953) statistics is given by

$$t_{\beta F} = \sqrt{2(m_i - 1) [\beta^{-1/(m_i-1)} - 1]}_{m_i \rightarrow \infty} \approx \sqrt{-2 \ln(\beta)}, \quad (80)$$

$$b_\varphi = \frac{1}{\sqrt{m_i K_{B0}}}, \quad (81)$$

where  $K_{B0}$  is defined in eq. (55). The Fisher confidence cone is thus given by

$$\alpha_{95i} = t_{\beta F} / \sqrt{m_i K_{B0}}. \quad (82)$$

### 5.1.2 Marginal confidence intervals

The marginal distribution of  $\overline{Y_{Ii}}$  and  $\overline{Y_{Di}}$  can be expressed explicitly as

$$p(\overline{Y_{Ii}}/Y_{gI}(t_i)) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{W_{Ii}W_{Di} - W_{IDi}^2}{W_{Di}}} e^{-\frac{1}{2} \left( \frac{W_{Ii}W_{Di} - W_{IDi}^2}{W_{Di}} \right) (\overline{Y_{Ii}} - Y_{gI}(t_i))^2}, \quad (83)$$

$$p(\overline{Y_{Di}}/Y_{gD}(t_i)) = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{W_{Ii}W_{Di} - W_{IDi}^2}{W_{Ii}}} e^{-\frac{1}{2} \left( \frac{W_{Ii}W_{Di} - W_{IDi}^2}{W_{Ii}} \right) (\overline{Y_{Di}} - Y_{gD}(t_i))^2}. \quad (84)$$

The confidence intervals for inclination and declination can be defined from the marginal probability density functions. For inclination for example, the ratio  $\frac{\overline{Y_{Ii}} - Y_{gI}(t_i)}{\sqrt{\frac{W_{Di}}{W_{Ii}W_{Di} - W_{IDi}^2}}}$  estimated by  $\frac{\overline{Y_{Ii}} - \hat{Y}_{gI}(t_i)}{\sqrt{\frac{S_{Ii}}{m_i - 1}}}$ , named the Student ratio, follows the Student t-distribution with  $m_i - 1$  degrees of freedom. We have

$$p(\hat{Y}_{gI}(t_i) - e_{\overline{Y_{Ii}}} \leq Y_{gI}(t_i) \leq \hat{Y}_{gI}(t_i) + e_{\overline{Y_{Ii}}}) = 1 - \beta, \quad (85)$$

where

$$e_{\overline{Y_{Ii}}} = t_{\beta/2} \sqrt{\frac{W_{Di}}{W_{Ii}W_{Di} - W_{IDi}^2}}. \quad (86)$$

The Student coefficients  $t_{\beta/2}$  are given in Table 2. When  $m_i$  becomes large ( $\geq 30$ ), this ratio tends to the normal distribution. Then  $t_{\beta/2} = 1.96$  can be taken to define a confidence interval of 95 per cent for  $Y_{gI}(t_i)$ . However, when  $m_i$  becomes small, typically less than 10,  $t_{\beta/2}$  must be taken from the table of Student distribution. The marginal errors on inclination and declination can be clearly related to the elliptic parameters  $K_x$ ,  $K_y$  and  $\Omega$  as

$$e_{\overline{Y_{Ii}}} = t_{\beta/2} \sqrt{\frac{W_{Di}}{W_{Ii}W_{Di} - W_{IDi}^2}} = t_{\beta/2} \sqrt{\frac{1}{m_i} \left[ \frac{\cos^2 \Omega}{K_x} + \frac{\sin^2 \Omega}{K_y} \right]}, \quad (87)$$

$$e_{\overline{Y_{Di}}} = t_{\beta/2} \sqrt{\frac{W_{Ii}}{W_{Ii}W_{Di} - W_{IDi}^2}} = t_{\beta/2} \sqrt{\frac{1}{m_i} \left[ \frac{\cos^2 \Omega}{K_y} + \frac{\sin^2 \Omega}{K_x} \right]}. \quad (88)$$

More specifically, if  $W_{IDi} = 0$ , then the error expressions become

$$e_{\overline{Y_{Ii0}}} = e_{\overline{Y_{Di0}}} = e_{\overline{Y_{i0}}} = t_{\beta/2} \sqrt{\frac{1}{W_i}}, \quad (89)$$

which defines error of the means given by eqs (61) and (62).

Returning to the geographical coordinate system, and in the limits of remarks on concentration factors in Appendix A1.1, the marginal errors of inclination and declination can be defined as

$$e_{\hat{g}_I} = e_{\overline{Y_{i0}}}, \quad (90)$$

$$e_{\hat{g}_D} = e_{\overline{Y_{i0}}} / \cos \hat{g}_I. \quad (91)$$

## 5.2 Plotting the curves: interpolation, fitting and smoothing

For plotting the curve on a stereographic diagram for inclination–declination, an interpolation by piecewise great circles between regularly distributed time knots  $t_i$  is made and confidence ellipses are drawn. The successive ellipses draw a band on the sphere representative of the

confidence of the mean direction path. For the marginal or univariate cases, an interpolation by a natural cubic spline with continuous first derivative is applied to the estimated points  $g(t_i)$  and also to the confidence envelope. The continuous second derivative constraint (that is an interpolating natural cubic spline) is not used here because, when the number of sites  $m_i$  is not high, it can lead to unrealistic wiggled interpolation curves.

The moving average technique comes back to a linear regression, which is drastically dependent on width of the window. However, there is no obvious argument to choose the optimal width, or to vary it. The width of 100 yr used for the different examples below is only heuristic! Moreover, using large windows poses some serious problems. The moving average operation constitutes a linear filter: it suppresses components presenting some certain forms and lets others pass. If a signal has a periodic component of period  $T$ , then a moving average with window width equal to  $T$  will filter this periodic component. If  $c(t)$  is a linear time function, this will go through the moving average filter without change. However, if  $c(t)$  is any function, then the gravity centres produced by the moving average will lie in the concavity of the curve. If there are some reversals of trend and if the curve is dissymmetric, then the reversals of the filtered series will be advanced or delayed. However, if the signal possesses some random fluctuations  $\varepsilon(t)$ , then a moving average operator will divide the variance of  $\varepsilon(t)$  by  $T$ , hence an attenuation, a smoothing of fluctuations will occur. This is one reason for the use of this technique. However, variates become highly correlated after that and this correlation can generate some periodic movements absent in the original series, especially when moving average operation is repeated.

It is emphasized that the moving average treatment does not constitute a global functional estimation of  $Y_g$ , although it is bivariate by considering inclination and declination simultaneously. If the physical connection between the geomagnetic elements through time is taken into account, an additional hypothesis about the global nature of  $Y_g$  should be introduced into the statistical approach. This question of the preliminary choice of the function  $Y_g$  of type (e.g.) penalized spline has been discussed elsewhere (Jupp & Kent 1987; Tsunakawa 1992; Green & Silverman 1994; Lanos 2001, 2004). Here, the moving average technique assures only a point estimate of the field at each time  $t_i$ , not a functional one.

The moving average needs numerous and time-even data. Otherwise, in order to compensate this default, it is tempting to enlarge the width of the window in order to fill the gaps, thus increasing previous drawbacks. A solution would be to match the width of the window to the density of points over time. This is possible, but must be in the context of the functional approach mentioned above.

## 6 APPLICATIONS

The following examples aim to show the implementation of the hierarchical calculations on true data and to compare the results with those given by the stratified statistics.

### 6.1 Examples at sample and site levels

Fig. 3 shows the results of the Fisher statistics and its Gaussian approximation applied to two sites studied in Bulgaria, in the contexts of stratification and hierarchy, respectively.

The Kovatchevo site (Fig. 3a and Table 3) has inclination and declination for 10 samples derived from 40 specimens coming from the remains of an ancient fire. The two circles of confidence  $\alpha_{95ij}$  were calculated using stratification and hierarchy procedures. As expected, mean declinations and inclinations do not agree because the number of specimens per sample is not the same. The average result based on specimens is drawn towards the direction of the samples with greater number of specimens measured. Here, the empirical Fisher concentration factor  $\kappa_{ij}^*$  is larger in the case of stratification than in the case of hierarchy ( $\kappa_{ij}^*$  stratification = 345;  $\kappa_{ij}^*$  hierarchy = 281).

The Madretz site (Fig. 3b and Table 4) has inclination and declination for 13 samples derived from 37 specimens. As in the previous example, the mean inclination and declination are closer to the sample direction with the greater number of measured specimens in the case of stratification. However here,  $\kappa_{ij}^*$  remains smaller in the case of stratification ( $\kappa_{ij}^*$  stratification = 190,  $\kappa_{ij}^*$  hierarchy = 341).

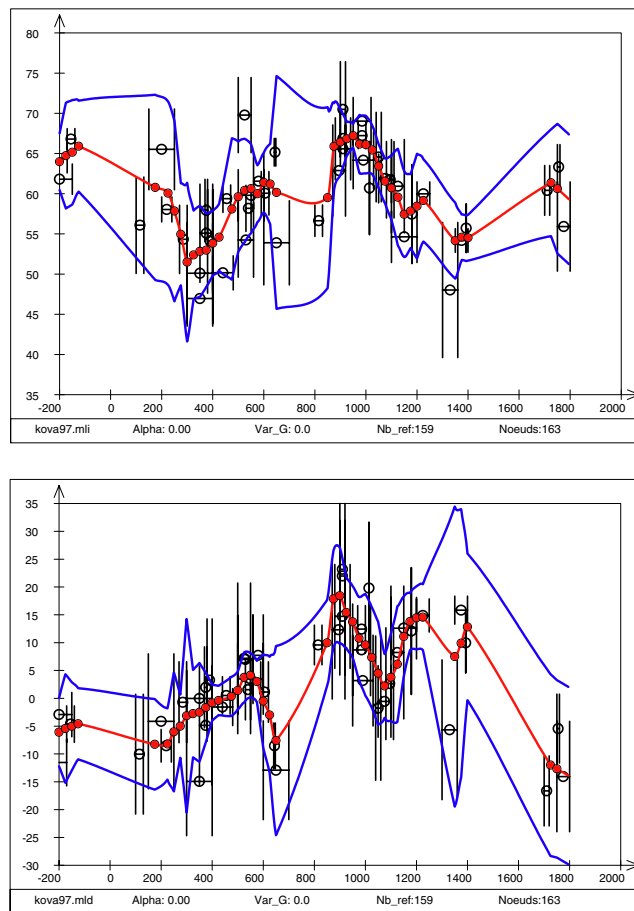
The hierarchy model corrects the distortion coming from the variability in the numbers of specimen from sample to sample. If hierarchical modelling had not been made, the samples having many specimens would have induced a bias in the mean value.

### 6.2 Examples at a field level

The following two examples aim to show the effect of hierarchy modelling onto error ellipses obtained at field level.

The first example (Fig. 4a) comprises six reference sites taken from the Bulgarian database (Kovacheva 1997), with overlapping window  $t_i = 5400 \text{ BC} \pm 50 \text{ yr}$ . Site inclination values are taken without geographical correction (the complete statistical results for the different statistics discussed in the paper are given in Table 5: cf. Sections 2 and 3).

The second example (Fig. 4b) comprises five synthetic sites well grouped with low  $\kappa_{ij}^*$ . This example is taken from Bucur's (1994) database, at  $t_i = -75 \pm 120 \text{ yr}$ , with geographical correction (virtual geomagnetic pole correction), but with all  $P_{ij} = 1$  and  $\kappa_{ij}^*$  divided by 10 for didactic reasons. The statistical results are given in Table 6.



**Figure 8.** Kovacheva's database (1997), from 200 BC to AD 1900. Curves obtained with overlapping windows of width 100 yr plotted at 25-yr steps: (a) marginal inclination curve, (b) marginal declination curve, (c) univariate inclination curve obtained with *in situ* sites and sets of bricks, (d) univariate intensity curve. All the curves are with 95 per cent confidence bands.

Mean estimates are not so different, within approximately  $1^\circ$ . The essential difference lies in the error estimates on the parameters. Larger errors obtained with the Bayesian statistics are easily explained by the fact that it was the number of sites that was used in the calculation of the confidence ellipses, instead of the number of samples.

### 6.3 Examples at a curve level

#### 6.3.1 What is necessary in archaeomagnetic databases?

In the published databases, the necessary parameters are conventionally given at a site level. This means that the measurement squared errors  $e_{ijkl}^2$  and interspecimen variances  $\sigma_{ijk}^2$  are not given. Thus, it is implicitly supposed that squared errors between specimens are lower than intersamples variances, that is, the descending hierarchy scheme described in Section 4.5.3 was verified. Usually the Fisherian statistics were used to obtain the site direction  $(\overline{I}_{ij}, \overline{D}_{ij})$ . The confidence angle  $\alpha_{95ij}$  and the empirical concentration parameter  $\kappa_{ij}^*$  are given. In order to compute the statistics, we need the following.

- (i) For the site of reference  $ij$ .

$t_{ij1}, t_{ij2}$ : time interval attributed to the point of reference (afterwards the centred time  $t_{ij}$  is defined).

$r_{ij}$ : the number of samples contributing to the statistics. The number of samples shown in Kovacheva (1997) is the number of independently oriented samples and not the total number of specimens measured. Often this is not the case, which can artificially diminish the calculated  $\alpha_{95ij}$ , as was shown in Section 6.1 (Figs 3a and b).

- (ii) For the direction (inclination and declination expressed in degrees).

$\overline{Y}_{ij}, \overline{Y}_{Dij}$ : values of the inclination and the declination, after geographical correction via the VGP model in order to transfer all the measurements to the same reference site.

$\kappa_{ij}^*$ : concentration factor, related to the confidence angle (in degrees):  $\alpha_{95ij} = 140 / \sqrt{r_{ij}\kappa_{ij}^*}$

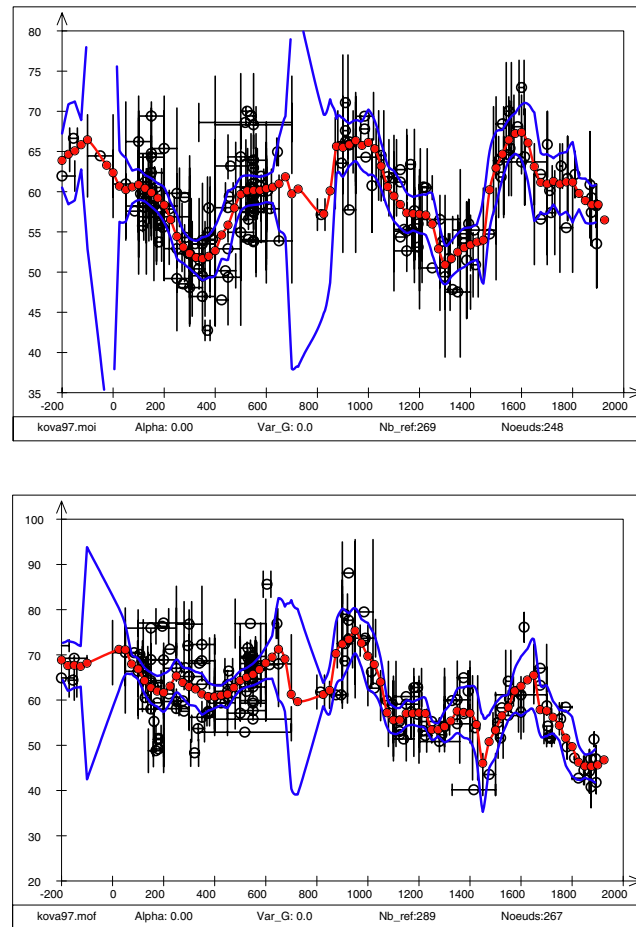


Figure 8. (Continued.)

Many sites in the AD period (Burlatskaya 1986; Kovacheva 1997) are only represented by average inclination  $\overline{Y}_{ij}$  results obtained from bricks. In Kovacheva (1997), these results are associated with the standard deviation  $S_{lij}$  of the population (without correction for bias) and not with the standard error of the mean, which is  $S_{lij}/\sqrt{r_{ij}}$ . Then, the equivalent concentration factor can be defined as  $\kappa_{ij}^* = \frac{r_{ij}-1}{r_{ij}S_{lij}^2}$ , provided that inclinations are not near the pole.

(iii) For the intensity (expressed in  $\mu T$ ).

$\overline{F}_{ij}$ : value of the intensity  $F$  after latitudinal correction.

$S_{Fij}$ : value of the observed dispersion, not corrected for bias (1 standard deviation).

### 6.3.2 The calculated curves for France

An application of the algorithms described in Sections 4 and 5 is shown for the French and Bulgarian data. The 119 French data points, labelled PC in Bucur (1994), are plotted on a stereogram in Fig. 5. The curve obtained using the Bayesian hierarchical statistics (with a minimum of three reference points per window with 100 yr width, at every 25-yr step, with error date weighting according to eq. (28) is given in Fig. 6(a) with ellipses at the 95 per cent confidence level. The corresponding marginal inclination and declination curves are shown in Figs 6(b) and (c). Curves obtained with Le Goff's stratification statistics are plotted in Figs 7(a), (b) and (c). The mean curves are globally similar (as outlined in Section 6.2), but the 95 per cent confidence ellipses are very different for the two treatments.

### 6.3.3 The calculated curves for Bulgaria

The marginal curves for Bulgaria (Kovacheva 1997), in the Bayesian hierarchical bivariate case with window width of 100 yr, are given in Fig. 8(a) for inclination and in Fig. 8(b) for declination, between 200 BC and AD 1900, for comparison with France (only inclination and declination data from kilns and hearths were used). The inclination curve (Fig. 8c) has been obtained in the Bayesian univariate case, considering only inclination values from *in situ* sites and sets of bricks (Kovacheva 1997). The intensity curve, also computed in the Bayesian univariate case (Appendix A3), is shown in Fig. 8(d). A 100-yr window was chosen in order to get enough points  $m_i$  in each window, otherwise

error bands on the Bayesian hierarchical curves would become very large. However, this enlargement of width leads to undesirable strong smoothing effects.

Whatever the statistical approach, stratified or hierarchical, the compromise between fitting and smoothing remains difficult to realize with the moving average technique when the number of reference points is small.

## 7 GENERAL CONCLUSION

A re-examination of sampling procedures led to the introduction of a Bayesian hierarchical model for archaeomagnetic experimental errors, using the natural hierarchical sampling process commonly implemented by laboratories in the field. The common idea in archaeomagnetism, as in palaeomagnetism, is to realize a hierarchical sampling that is representative of the heterogeneity of the magnetization in the site and that permits a check on the presence of possible systematic errors. The main conclusions are as follows.

(i) Calculations in the hierarchical context avoid the disturbance resulting from variability in the number of specimens taken from each sample and in the number of samples taken from each site. The weighting factors used for calculating the mean geomagnetic field direction (or intensity) are rather complicated in the case of dating errors, but they are fundamentally based on a combination of a variance resulting from sources of scatter at field level and a squared error resulting from sampling errors at the lower levels.

(ii) The fact that the results at sample and lower levels are not published in the literature implicitly implies that the average of the interspecimen squared errors are always less than or equal to the intersample variance in a site. This is the hypothesis of descending hierarchy, which often seems to be verified in practice.

(iii) Typically, a study of at least 20 samples will give an  $\alpha_{95i}$  5 per cent close to the optimal  $\alpha_{95i}$  for a fixed site number  $m_i$  and if errors at the different levels are random with zero mean (no systematic errors).

(iv) The precision of the curve itself is essentially controlled by the number  $m_i$  of reference points per window and by dating errors  $P_{ij}$ , more than by  $\alpha_{95ij}$ , if the descending hierarchy is verified.

(v) Error bands of curves never increase with bad data, that is with sites having large  $\alpha_{95ij}$ , in the context of random errors. Accordingly, it is recommended that published databases should always contain all the observed data without any prior selection. From a statistical point of view, it is much more preferable to do only posterior selections, for example when systematic errors are detected.

(vi) The Bayesian elliptic distribution proposed here is clearly related to the window width. As a result of dating uncertainties, sites attributed to a window, and so contributing to the field estimation, in fact belong to different time parts of the curve. Consequently, the principal axis of the ellipse will be, theoretically, tangential to a great circle, thus defining a local regression on the sphere.

Finally, the moving average technique is well adapted to numerous and very well dated data evenly distributed over time, i.e. without gaps. In this ideal case, the window width can be as small as possible and the present statistical analysis becomes operational. This encourages the acquisition of much more data at a site level in order to increase significantly the size of the archaeomagnetic data sets, that is, at least 10 sites for every step (e.g. 10 or 20 yr, depending on the time resolution needed). In practice, this is certainly very optimistic because archaeomagnetic data remain generally unevenly distributed in time, with large dating errors. On this basis, an other approach has been developed (Schnepp *et al.* 2003; Lanos 2004) that puts some prior knowledge on the global nature of the curve to be estimated. This more sophisticated approach allows the window width to be automatically adapted to the density of points along time, making the points movable inside the dating ranges, and can take into account some qualitative information such as any stratigraphic constraints provided by archaeology.

## ACKNOWLEDGMENTS

This work was carried out in the framework of the joint project DRI no. 6118 between the CNRS (UMR 6566 Civilisations Atlantiques et Archéosciences, Rennes, UMR 6118 Géosciences-Rennes, and Institut de Physique du Globe de Paris, France) and the Bulgarian Academy of Sciences (Geophysical Institute, Sofia, Bulgaria). ES was supported by the Deutsche Forschungsgemeinschaft, grants Schn 366/4-1 to 4-5. The valuable remarks and corrections made by the two referees Professor D. Tarling and Dr P. Márton, by Professor Jean Deshayes (Institut de Mathématiques de Rennes, France), Dr C. Buck (Department of Probability and Statistics, University of Sheffield, UK) and Dr A. Chauvin (Géosciences-Rennes, France) are highly appreciated. We also thank Ph. Dufresne for his valuable help in figure drawing.

## REFERENCES

- Aitken, M. & Hawley, M., 1967. Archaeomagnetic measurements in Britain IV, *Archaeometry*, **10**, 129–135.
- Anderson, T.W., 1984. *An introduction to multivariate statistical analysis*, 2nd edn, J. Wiley and Sons, New York, p. 675.
- Barnett, V., 1982. *Comparative statistical inference*, 2nd edn, Wiley, New York.
- Batt, C.M., 1997. The British archaeomagnetic calibration curve: an objective treatment, *Archaeometry*, **39**(1), 153–168.
- Buck, C.E., Kenworthy, J.B., Litton, C.D. & Smith, A.F.M., 1991. Combining archaeological and radiocarbon information: a Bayesian approach to calibration, *Antiquity*, **65**, 808–821.
- Bucur, I., 1994. The direction of the terrestrial magnetic field in France during the last 21 centuries. Recent progress, *Phys. Earth planet. Int.*, **87**, 95–109.
- Burlatskaya, S.P., 1986. *Archaeomagnetic determinations of geomagnetic field elements, world data*, Soviet Geophysical Committee of the USSR Academy of Sciences, Moscow.
- Chauvin, A., Garcia, Y., Lanos, Ph. & Laubheimer, F., 2000. Palaeointensity of the geomagnetic field recovered on archaeomagnetic sites from France, *Phys. Earth planet. Int.*, **120**, 111–136.

- Clark, R.M., Tarling, D.H. & Noel, M., 1988. Developments in archaeomagnetic dating in Britain, *J. Archaeol. Sci.*, **15**, 645–667.
- Daly, L. & Le Goff, M., 1996. An updated and homogeneous world secular variation data base. 1. Smoothing of the archaeomagnetic results, *Phys. Earth planet. Int.*, **93**, 159–190.
- Droesbeke, J.-J., Fine, J. & Saporta, G. (eds), 2002. *Méthodes Bayésiennes en statistique* Technip, Paris, p. 418.
- Dudewicz, E.J. & Mishra, S.N., 1988. *Modern Mathematical statistics*, J. Wiley and Sons, New York, p. 838.
- Eighmy, J.L., 1990. Archaeomagnetic dating: practical problems for the archaeologist, in *Archaeomagnetic dating*, 33–64, eds Eighmy, J.L. & Sternberg, R.S., The Univ. of Arizona Press, Tucson.
- Fisher, R.A., 1953. Dispersion on a Sphere, *Proc. R. Soc., Lond., A*, **217**, 295–305.
- Fisher, N.I., Lewis, T. & Embleton, B.J.J., 1987. *Statistical analysis of spherical data*, Cambridge University Press, Cambridge, p. 329.
- Gallet, Y., Genevey, A. & Le Goff, M., 2002. Three millennia of directional variation of the Earth's magnetic field in western Europe as revealed by archaeological artefacts, *Physics Earth planet. Int.*, **131**, 81–89.
- Genevey, A. & Gallet, Y., 2002. Intensity of the geomagnetic field in Western Europe over the past 2000 years: new data from ancient French pottery, *J. geophys. Res.*, **107**(B11), 2285, doi: 10.1029/2001JB000701.
- Genevey, A., Gallet, Y. & Margueron, J.-C., 2003. Eight thousand years of geomagnetic field intensity variations in the Eastern Mediterranean, *J. geophys. Res.*, **108**(B5), 2228, doi: 10.1029/2001JB001612.
- Green, P.J. & Silverman, B.W., 1994. *Nonparametric regression and generalized linear models, a roughness penalty approach*, Chapman & Hall, London, p. 182.
- Hongre, L., Hulot, G. & Khokhlov, A., 1998. An analysis of the geomagnetic field over the past 2000 years, *Phys. Earth planet. Int.*, **106**(3–4), 311–335.
- Howson, C. & Urbach, P., 1993. *Scientific reasoning: the Bayesian approach*, 2nd edn, Open Court, Chicago, Illinois.
- Hus, J. & Geeraerts, R., 1998. The direction of geomagnetic field in Belgium since roman time and the reliability of archaeomagnetic dating, *Phys. Chem. Earth*, **23**(9–10), 997–1007.
- Jupp, P.E. & Kent, J.T., 1987. Fitting smooth paths to spherical data, *Applied Statist.*, **36**(1), 34–46.
- Kovacheva, M., 1997. Archaeomagnetic database from Bulgaria: the last 8000 years, *Phys. Earth planet. Int.*, **102**, 145–151.
- Kovacheva, M. & Toshkov, A., 1994. Geomagnetic field variations as determined from Bulgarian archaeomagnetic data. Part I: The last 2000 years AD, *Surv. Geophys.*, **15**, 673–701.
- Kovacheva, M., Jordanova, N. & Karloukovski, V., 1998. Geomagnetic field variations as determined from Bulgarian archaeomagnetic data. Part II: The last 8000 years, *Surv. Geophys.*, **19**, 431–460.
- Labelle, J.M. & Eighmy, J.L., 1997. Additional archaeomagnetic data on the south-west USA master geomagnetic pole curve, *Archaeometry*, **39**(2), 431–439.
- Lanos, Ph., 1987. The effects of demagnetising fields on thermoremanent magnetisation acquired by parallel-sided baked clay blocks, *Geophys. J. R. astr. Soc.*, **91**, 985–1012.
- Lanos, Ph., 2001. L'approche bayésienne en chronométrie: application à l'archéomagnétisme, in *Datation*, XXIe rencontres internationales d'archéologie et d'histoire d'Antibes, eds Barrandon, J.-N., Guibert, P., Michel, V., éditions APDCA, Antibes, France, pp. 113–139.
- Lanos, Ph., 2004. Bayesian inference of calibration curves : application to archaeomagnetism, in *Tools for constructing chronologies: crossing disciplinary boundaries*, Vol. 177, pp. 43–82, eds Buck, C. & Millard, A., Springer-Verlag, London.
- Lanos, Ph., Kovacheva, M. & Chauvin, A., 1999. Archaeomagnetism, methodology and applications: implementation and practice of the archaeomagnetic method in France and Bulgaria, *European J. Archaeology*, **2**(3), 365–392.
- Le Goff, M., 1990. Lissage et limites d'incertitude des courbes de migration polaire: pondération des données et extension bivariate de la statistique de Fisher, *C. R. Acad. Sci. Paris*, **311**(II), 1991–1998.
- Le Goff, M., Henry, B. & Daly, L., 1992. Practical method for drawing a VGP path, *Phys. Earth. planet. Int.*, **70**, 201–204.
- Lindley, D.V., 1990. The 1988 Wald memorial lectures: the present position in Bayesian statistics, *Statistical Science*, **5**(1), 44–89.
- Love, J.J. & Constable, C.G., 2003. Gaussian statistics for paleomagnetic vectors, *Geophys. J. Int.*, **152**, 515–565.
- McFadden, P.L. & Reid, A.B., 1982. Analysis of palaeomagnetic inclination data, *Geophys. J. R. astr. Soc.*, **69**, 307–319.
- Márton, P., 1996. Archaeomagnetic directions: the Hungarian calibration curve, in *Palaeomagnetism and Tectonics of the Mediterranean Region*, Vol. 105, pp. 385–399, eds Morris, A. & Tarling, D., Geological Society, London.
- Márton, P., 2003. Recent achievements in archaeomagnetism in Hungary, *Geophys. J. Int.*, **153**, 675–690.
- Press, H.P., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1997. *Numerical Recipes in C, the art of scientific computing*, 2nd edn, Cambridge University Press, Cambridge, p. 994.
- Schnepp, E., Pucher, R., Goeddicke, C., Manzano, A., Müller, U. & Lanos, Ph., 2003. Paleomagnetic directions and Thermoluminescence dating from a bread oven-floor sequence in Lübeck (Germany): a record of 450 years of geomagnetic secular variation, *J. geophys. Res.*, **108**, 53–66.
- Schnepp, E., Pucher, R., Reinders, J., Hambach, U., Soffel, H.C. & Hedley, I., 2004. A German catalogue of archaeomagnetic data, *Geophys. J. Int.*, **157**, 64–78.
- Sternberg, R., 1989. Secular variation of the Archaeomagnetic Direction in the American Southwest, AD750–1425, *J. geophys. Res.*, **94**, 527–546.
- Tarling, D., 1983. *Paleomagnetism*, Chapman and Hall, London and New York, p. 379.
- Tarling, D. & Dobson, M., 1995. Archaeomagnetism: An Error Assessment of burnt Material Observations in the British Directional Database, *J. Geomag. Geoelectr.*, **47**, 5–18.
- Tassi, Ph., 1992. *Méthodes statistiques*, Collection Economie et Statistiques avancées, 2nd edn, Economica, Paris, p. 474.
- Tsunakawa, H., 1992. Bayesian approach to smoothing palaeomagnetic data using ABIC, *Geophys. J. Int.*, **108**, 801–811.
- Véges, I., 1970. Map plotting with weighted average on the surface of a circular disc, *Pure appl. Geophys.*, **78**, 5–17.
- Watson, G.S., 1983. *Statistics on spheres*, John Wiley & Sons, New-York, p. 238.
- Westphal, M. & Gurevitch, E., 1996. Le problème de l'analyse statistique des mesures d'inclinaison paléomagnétiques lorsqu'on ne dispose pas de la déclinaison. Cas particulier des inclinaisons élevées., *C. R. Acad. Sci. Paris*, **322**(IIa), 261–268.

## APPENDIX A 1

### A1.1 Fisher statistics

The formulae are presented at field level ( $i$ ). Let the mean direction ( $I_{ij}$ ,  $D_{ij}$ ) observed at site level be Fisher distributed (Fisher 1953; Fisher et al. 1987) around the mean direction ( $I_i$ ,  $D_i$ ), then the probability density function of  $I_{ij}$  and  $D_{ij}$  will be given by

$$p(I_{ij}, D_{ij}) = \frac{K_i}{4\pi \sinh K_i} \left[ e^{K_i [\cos I_{ij} \cos I_i \cos(D_{ij} - D_i) + \sin I_{ij} \sin I_i]} \right] \cos I_{ij}, \quad (\text{A1.1})$$



where  $K_i$  is the concentration factor. The marginal probability density functions of  $I_{ij}$  and  $D_{ij}$  are approximately Gaussian provided that  $K_i$  is large enough and that the mean direction is not too near the geographical pole (Aitken & Hawley 1967; McFadden & Reid 1982; Westphal & Gurevitch 1996). Typically, for  $K_i > 50$ , the mean inclination  $I_i$  has to be  $< 70^\circ$ . Thus, it is possible to define the two marginal variances  $\sigma_{I_i}^2 = \frac{1}{K_i}$  and  $\sigma_{D_i}^2 = \frac{1}{K_i \cos^2 I_i}$  and the approximate density:

$$p(I_{ij}, D_{ij}) \approx \frac{1}{2\pi \sigma_{I_i} \sigma_{D_i}} e^{-\frac{1}{2} \left[ \left( \frac{I_{ij} - I_i}{\sigma_{I_i}} \right)^2 + \left( \frac{D_{ij} - D_i}{\sigma_{D_i}} \right)^2 \right]} \quad (\text{A1.2})$$

### A1.1.1 Estimation

The mean direction  $(I_i, D_i)$  is estimated from observation of  $m_i$  sites with directions  $(I_{ij}, D_{ij})$ :

$$\hat{I}_i = \arcsin(\bar{z}_i / \bar{R}_i), \quad \hat{D}_i = \arctan(\bar{y}_i / \bar{x}_i), \quad (\text{A1.3})$$

where the mean components are

$$\bar{x}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} x_{ij} \quad \bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij} \quad \bar{z}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} z_{ij}$$

$$x_{ij} = \cos I_{ij} \cos D_{ij} \quad y_{ij} = \cos I_{ij} \sin D_{ij} \quad z_{ij} = \sin I_{ij}$$

and where the resultant mean length  $\bar{R}_i$  is given by  $\bar{R}_i = \sqrt{\bar{x}_i^2 + \bar{y}_i^2 + \bar{z}_i^2}$ .

The Fisher statistics defines a confidence cone (or error cone) about the estimated mean direction  $(\hat{I}_i, \hat{D}_i)$ , at  $(1 - \beta) = 95$  per cent level, in radians, as

$$\alpha_{95i} = t_{\beta F}(m_i) / \sqrt{m_i K_i}, \quad (\text{A1.4})$$

where  $t_{\beta F}(x) = \sqrt{2(x-1)[\beta^{-1/(x-1)} - 1]}$  for  $x \geq 2$  (Fisher 1953).  $t_{\beta F} \approx 2.45$  when  $x$  tends to infinity (in fact, this value is nearly achieved for  $x > 10$ ; see Table 2). The unbiased (hence symbol \*) estimation  $\kappa_i^*$  of the concentration factor  $K_i$  is given by

$$\kappa_i^* = \left( \frac{m_i - 1}{m_i} \right) \frac{1}{1 - \bar{R}_i}. \quad (\text{A1.5})$$

### A1.2 Gaussian bivariate statistics

The Fisher statistics can be approximated by a bivariate Gaussian statistics in the tangential plane normal to the polar mean direction of a given level in the hierarchical modelling (here the field level is considered). For this purpose, the coordinate system needs to be changed using an appropriate rotation ( $\lambda = -I_R$ ,  $\phi = D_R$ ) where  $I_R$  is as near as possible to  $I_i$ , to obtain directions  $(Y_{ij}, Y_{Dij})$  near the geographical equatorial plane.

Using the rotation (Fisher *et al.* 1987, p. 32),

$$R = \begin{pmatrix} \cos \lambda \cos \phi & \cos \lambda \sin \phi & -\sin \lambda \\ -\sin \phi & \cos \phi & 0 \\ \sin \lambda \cos \phi & \sin \lambda \sin \phi & \cos \lambda \end{pmatrix}, \quad (\text{A1.6})$$

the new coordinates are

$$\begin{bmatrix} x_{ij}^E \\ y_{ij}^E \\ z_{ij}^E \end{bmatrix} = R \begin{bmatrix} x_{ij} = \cos I_{ij} \cos D_{ij} \\ y_{ij} = \cos I_{ij} \sin D_{ij} \\ z_{ij} = \sin I_{ij} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x_i^E \\ y_i^E \\ z_i^E \end{bmatrix} = R \begin{bmatrix} x_i = \cos I_i \cos D_i \\ y_i = \cos I_i \sin D_i \\ z_i = \sin I_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The new site direction is

$$Y_{ij} = \arcsin(z_{ij}^E), \quad Y_{Dij} = \arctan(y_{ij}^E / x_{ij}^E).$$

In the same way, the polar mean direction becomes  $(Y_{Ii}, Y_{Di})$ .

This rotation to the equatorial plane leads to very small differences  $(Y_{Iij} - Y_{Ii})$  and  $(Y_{Dij} - Y_{Di})$  provided that the concentration factor  $K_i$  is large enough, more than 50 (half-angle of dispersion cone less than  $20^\circ$ ) and direction  $(I_R, D_R)$  is near to  $(I_i, D_i)$ . In practice, the mean direction  $(I_R, D_R)$  is deduced from the Fisher statistics applied to site directions  $(I_{ij}, D_{ij})$ , weighted by a dating factor  $P_{ij}$  (eq. 28).

Consequently, eq. (A1.1) leads to the probability density function

$$p(Y_{Iij}, Y_{Dij}) = \frac{K_i}{4\pi \sinh K_i} \left[ e^{K_i [\cos Y_{Iij} \cos Y_{Ii} \cos(Y_{Dij} - Y_{Di}) + \sin Y_{Iij} \sin Y_{Ii}]} \right] \cos Y_{Iij}. \quad (\text{A1.7})$$

Letting the polar angles,

$$\theta_{ij} = \arccos[\cos Y_{Iij} \cos Y_{Li} \cos(Y_{Dij} - Y_{Di}) + \sin Y_{Iij} \sin Y_{Li}], \tag{A1.8}$$

$$\varphi_{ij} = \arccos[\cos Y_{Iij} \sin(Y_{Di} - Y_{Dij}) / |\sin \theta_{ij}|] - \frac{\pi}{2}, \tag{A1.9}$$

with the condition

$$\text{if } (\tan Y_{Iij} - \tan Y_{Li} \cos(Y_{Dij} - Y_{Di}) \leq 0) \text{ then } \varphi_{ij} \text{ becomes } ((Y_{Dij} - Y_{Di}) / |Y_{Dij} - Y_{Di}|)\pi - \varphi_{ij}. \tag{A1.10}$$

The probability density function of  $(\theta_{ij}, \varphi_{ij})$  is given by

$$p(\theta_{ij}, \varphi_{ij}) = \frac{K_i}{4\pi \sinh K_i} [e^{K_i \cos \theta_{ij}}] \sin \theta_{ij}. \tag{A1.11}$$

If  $(Y_{Iij} - Y_{Li})$  and  $(Y_{Dij} - Y_{Di})$  are small, then the approximations of  $(\theta_{ij}, \varphi_{ij})$  are

$$\theta_{ij} \approx (Y_{Iij} - Y_{Li})^2 + (Y_{Dij} - Y_{Di})^2, \quad \tan \varphi_{ij} \approx \frac{(Y_{Dij} - Y_{Di})}{(Y_{Iij} - Y_{Li})},$$

that is,

$$Y_{Iij} - Y_{Li} \approx \theta_{ij} \cos \varphi_{ij}, \quad Y_{Dij} - Y_{Di} \approx \theta_{ij} \sin \varphi_{ij}. \tag{A1.12}$$

Accordingly, the probability density function of  $(Y_{Iij}, Y_{Dij})$  is approximately

$$p(Y_{Iij}, Y_{Dij}) \approx \frac{1}{2\pi \sigma_i^2} e^{-\frac{1}{2} \left[ \left( \frac{Y_{Iij} - Y_{Li}}{\sigma_i} \right)^2 + \left( \frac{Y_{Dij} - Y_{Di}}{\sigma_i} \right)^2 \right]} \quad \text{with} \quad \sigma_i^2 = \frac{1}{K_i}, \tag{A1.13}$$

which defines a Gaussian bivariate statistics (in the tangential plane to the sphere perpendicular to the polar mean direction, in the new coordinate system). The new variables  $(Y_{Iij}, Y_{Dij})$  can be considered as independent Gaussian random variables. The Bayesian calculations in the paper are made using a rotation  $R$  with  $(\lambda = -I_R, \phi = D_R) = (-\hat{I}_i, \hat{D}_i)$  (eq. A1.3). At the end, we go back to the geographical coordinate system using the inverse of the matrix (A1.6).

## APPENDIX A2

### A2.1 The bivariate Le Goff (1992) statistics

The bivariate approach proposed by Le Goff (1990), Le Goff *et al.* (1992) and Daly & Le Goff (1996), consists of calculating a confidence ellipse around the mean direction using the inertia tensor concept.

#### A2.1.1 Fisherian statistics

Each sample is characterized by an inclination  $I_{ijk}$  and a declination  $D_{ijk}$ , assumed to be Fisher distributed around the mean direction  $(I_{ij}, D_{ij})$  and with concentration parameter  $K_{ij}$ . The associated inertia tensor  $T_{ijk}$  is defined by

$$T_{ijk} = \begin{pmatrix} 1 - x_{ijk}^2 & -x_{ijk}y_{ijk} & -x_{ijk}z_{ijk} \\ -x_{ijk}y_{ijk} & 1 - y_{ijk}^2 & -y_{ijk}z_{ijk} \\ -x_{ijk}z_{ijk} & -y_{ijk}z_{ijk} & 1 - z_{ijk}^2 \end{pmatrix}, \tag{A2.1}$$

where

$$x_{ijk} = \cos \theta_{ijk} \cos \varphi_{ijk}, \quad y_{ijk} = \cos \theta_{ijk} \sin \varphi_{ijk}, \quad z_{ijk} = \sin \theta_{ijk}.$$

Variables  $\theta_{ijk}, \varphi_{ijk}$  are defined by eqs (A1.8) to (A1.10), replacing indices  $(ij)$  by  $(ijk)$ , indices  $(i)$  by  $(ij)$ ,  $Y_{I\bullet}$  by  $I_{\bullet}$  and  $Y_{D\bullet}$  by  $D_{\bullet}$ .

The expected inertia tensor  $E[T_{ijk}]$ , which is diagonal, is very well approximated by

$$E[T_{ijk}] \approx \begin{pmatrix} 1 - \frac{1}{K_{ij}}(1 - \frac{1}{K_{ij}}) & 0 & 0 \\ 0 & 1 - \frac{1}{K_{ij}}(1 - \frac{1}{K_{ij}}) & 0 \\ 0 & 0 & \frac{2}{K_{ij}}(1 - \frac{1}{K_{ij}}) \end{pmatrix} \approx \begin{pmatrix} \frac{K_{ij}}{1+K_{ij}} & 0 & 0 \\ 0 & \frac{K_{ij}}{1+K_{ij}} & 0 \\ 0 & 0 & \frac{2}{1+K_{ij}} \end{pmatrix}, \tag{A2.2}$$

provided that  $K_{ij}$  is large enough ( $>50$ ). The trace of the inertia tensor is equal to 2.

The empirical mean tensor for  $r_{ij}$  samples is defined by

$$\overline{T_{ij}} = \frac{1}{r_{ij}} \sum_{k=1}^{r_{ij}} T_{ijk} \tag{A2.3}$$

and can be diagonalized as  $\overline{T}_{ij}^d$ . Using the moment method estimation (MME; Watson 1983), the expected inertia tensor  $E[T_{ijk}]$  is estimated by  $\overline{T}_{ij}^d$ . Then, the unknown concentration parameter  $K_{ij}$  (assumed large) of the distribution can be given by

$$\hat{K}_{ij} = (\overline{T}_{ij\max}^d + \overline{T}_{ij\text{int}}^d) / \overline{T}_{ij\min}^d = \frac{r_{ij} + \sum_k \cos^2 \theta_{ijk}}{r_{ij} - \sum_k \cos^2 \theta_{ijk}}. \quad (\text{A2.4})$$

### A2.1.2 Bivariate statistics

Each site is characterized by an inclination  $I_{ij}$  and a declination  $D_{ij}$ , assumed to be distributed as the bivariate elliptic statistics defined by Le Goff *et al.* (1992), eq. (52) in this paper, around the mean direction  $(I_i, D_i)$ , and with concentration parameters  $K_x$  and  $K_y$  and orientation  $\Omega$ . The associated inertia tensor  $T_{ij}$  is defined by

$$T_{ij} = \begin{pmatrix} 1 - x_{ij}^2 & -x_{ij}y_{ij} & -x_{ij}z_{ij} \\ -x_{ij}y_{ij} & 1 - y_{ij}^2 & -y_{ij}z_{ij} \\ -x_{ij}z_{ij} & -y_{ij}z_{ij} & 1 - z_{ij}^2 \end{pmatrix}, \quad (\text{A2.5})$$

where

$$x_{ij} = \cos \theta_{ij} \cos \varphi_{ij}, \quad y_{ij} = \cos \theta_{ij} \sin \varphi_{ij}, \quad z_{ij} = \sin \theta_{ij}.$$

Variables  $\theta_{ij}$ ,  $\varphi_{ij}$  are defined by eqs (A1.8) to (A1.10), replacing  $Y_{I\bullet}$  by  $I_\bullet$  and  $Y_{D\bullet}$  by  $D_\bullet$ .

It can be shown that the expected inertia tensor  $E[T_{ij}]$  is very well approximated by

$$E[T_{ij}] \approx \begin{pmatrix} 1 - \left(\frac{\cos^2 \Omega}{K_x} + \frac{\sin^2 \Omega}{K_y}\right) & -\left(\frac{1}{K_x} - \frac{1}{K_y}\right) \frac{\sin 2\Omega}{2} & 0 \\ -\left(\frac{1}{K_x} - \frac{1}{K_y}\right) \frac{\sin 2\Omega}{2} & 1 - \left(\frac{\sin^2 \Omega}{K_x} + \frac{\cos^2 \Omega}{K_y}\right) & 0 \\ 0 & 0 & \frac{1}{K_x} + \frac{1}{K_y} \end{pmatrix}. \quad (\text{A2.6})$$

Using the MME, the empirical mean tensor for  $m_i$  sites,

$$\overline{T}_i = \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} T_{ij}, \quad (\text{A2.7})$$

can be related to the unknown concentration parameters  $K_x$  and  $K_y$  of the expected inertia. If one poses  $K_x < K_y$ , the diagonalization of the expected inertia tensor and of the empirical mean tensor gives

$$E[T_{ij}]^d \approx \begin{pmatrix} 1 - \frac{1}{K_x} & 0 & 0 \\ 0 & 1 - \frac{1}{K_y} & 0 \\ 0 & 0 & \frac{1}{K_x} + \frac{1}{K_y} \end{pmatrix} \hat{=} \overline{T}_i^d = \begin{pmatrix} \overline{T}_{i\text{int}}^d & 0 & 0 \\ 0 & \overline{T}_{i\max}^d & 0 \\ 0 & 0 & \overline{T}_{i\min}^d \end{pmatrix}, \quad (\text{A2.8})$$

where the symbol  $\hat{=}$  means estimated by. Hence, the estimation of  $K_x$  and  $K_y$ :

$$\begin{aligned} \hat{K}_x &= 2 / (\overline{T}_{i\max}^d - \overline{T}_{i\text{int}}^d + \overline{T}_{i\min}^d), \\ \hat{K}_y &= 2 / (\overline{T}_{i\text{int}}^d - \overline{T}_{i\max}^d + \overline{T}_{i\min}^d). \end{aligned} \quad (\text{A2.9})$$

The diagonal terms  $\overline{T}_{i\max}^d$ ,  $\overline{T}_{i\text{int}}^d$  and  $\overline{T}_{i\min}^d$  are the maximal, intermediary and minimal inertia terms expressed in the eigenvector basis. The two estimates in eq. (A2.9) give results very closed to those obtained with Le Goff's formulae (1992, p. 202).

### A2.1.3 Mixture of Fisherian and bivariate statistics

Le Goff *et al.* (1992) define a global tensor  $\overline{T}_i^R$  that mixes directions  $(ijk)$  from different Fisherian populations  $(ij)$ , which are distributed as an elliptic bivariate statistics around a mean direction  $(I_i, D_i)$ . This tensor is obtained using a rotation  $R_{ij}$ , function of  $\lambda = \theta_{ij}$  and  $\phi = \varphi_{ij}$  (eq. A1.6) and applied to each expected tensor  $E[T_{ijk}]$  weighted by a coefficient  $W_{Tij}$  (in practice the dating error  $P_{ij}$ ):

$$\overline{T}_i^R = \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} R_{ij}^T (E[T_{ijk}]) R_{ij}. \quad (\text{A2.10})$$

The mean direction  $(I_i, D_i)$  of the magnetic field within time window  $t_i$  is determined using a diagonalization procedure applied to  $\overline{T}_i^R$ . It corresponds to the minimal inertia axis. Considering that the mean directions of the sites  $(ij)$  are elliptic bivariate distributions, the expected global tensor  $E_{\theta_{ij}, \varphi_{ij}}[\overline{T}_i^R]$  is expressed by

$$E[\overline{T}_i^R] = E[T_{ij}] - \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} \frac{1}{K_{ij}} I + \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} \frac{3}{K_{ij}} E[V_{ij}], \quad (\text{A2.11})$$

where  $I$  is the identity matrix and  $V_{ij} = I - T_{ij}$ . If  $K_{ij}$ ,  $K_x$  and  $K_y$  are large, the diagonalized global tensor can be very well approximated by

$$E[\overline{T}_i^R]^d \approx E[T_{ij}]^d - \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} \frac{1}{K_{ij}} I, \quad (\text{A2.12})$$

that is,

$$E[\overline{T}_i^R]^d \approx \begin{pmatrix} 1 - \frac{1}{K_x} - \frac{1}{W_{Ti}} \sum W_{Tij} \frac{1}{K_{ij}} & 0 & 0 \\ 0 & 1 - \frac{1}{K_x} - \frac{1}{W_{Ti}} \sum W_{Tij} \frac{1}{K_{ij}} & 0 \\ 0 & 0 & \frac{1}{K_x} + \frac{1}{K_x} - \frac{1}{W_{Ti}} \sum W_{Tij} \frac{1}{K_{ij}} \end{pmatrix}. \quad (\text{A2.13})$$

The trace of  $E[\overline{T}_i^R]^d$  is equal to  $2 - (\frac{1}{W_{Ti}} \sum W_{Tij} \frac{3}{K_{ij}})$ , not far from 2 provided that  $K_{ij}$  is large. Consequently, an estimate of the two global concentration parameters  $K_x^R$  and  $K_y^R$  are given by the following very good approximations (within a few %):

$$\begin{aligned} \frac{1}{K_x^R} &\approx \frac{1}{K_x} + \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} \frac{1}{K_{ij}}, \\ \frac{1}{K_y^R} &\approx \frac{1}{K_y} + \frac{1}{W_{Ti}} \sum_{j=1}^{m_i} W_{Tij} \frac{1}{K_{ij}}. \end{aligned} \quad (\text{A2.14})$$

Eq. (A2.14) is equivalent to variances and equal to the sum of the weighted variance of the mean directions of the sites and of the mean of the weighted intersample variances. These expressions operate as a weighted combination of intersite and intrasite variances, which corresponds to the stratification statistical approach.

The expected global tensor  $E[\overline{T}_i^R]^d$  can be estimated by the diagonalized empirical global tensor  $\overline{T}_i^{Rd}$ , following the same procedure as in eqs (A2.8) and (A2.9). The global mean direction corresponds to the minimal inertia axis.

The orientation  $\Omega$  of the ellipse is obtained from Le Goff's formula (1992, p. 203). It can also be obtained from eqs (A1.8) to (A1.10) simplified as

$$\Omega = \arccos [\cos I_{\text{int}} \sin(D_{\text{min}} - D_{\text{int}})] - \frac{\pi}{2},$$

with the condition

$$\text{if } I_{\text{int}} \leq 0 \text{ then } \Omega \text{ becomes } ((D_{\text{int}} - D_{\text{min}})/|D_{\text{int}} - D_{\text{min}}|)\pi - \Omega, \quad (\text{A2.15})$$

where  $(I_{\text{min}}, D_{\text{min}})$  is the direction of minimal inertia axis and  $(I_{\text{int}}, D_{\text{int}})$  the direction of intermediary inertia axis (e.g.  $K_x < K_y$ ).

#### A2.1.4 Approximate estimation of Le Goff's parameters

Using the rotation described in Appendix A1.2, that is  $(\lambda = -I_R, \phi = D_R) = (-\hat{I}_i, \hat{D}_i)$ , the inclination  $I_{ij}$  and declination  $D_{ij}$  obtained for a site can be transformed to  $Y_{ij}$  and  $Y_{Dij}$ . Assuming that weightings  $P_{ij}$  and concentration parameters  $K_{ij}$  are respectively of the same order, the minimal inertia axis in the new coordinate system can be very well approximated by the weighted arithmetic mean:

$$Y_{I \text{ min}} \approx Y_{Ii} = \frac{1}{P_i} \sum_{j=1}^{m_i} P_{ij} Y_{Iij}, \quad Y_{D \text{ min}} \approx Y_{Di} = \frac{1}{P_i} \sum_{j=1}^{m_i} P_{ij} Y_{Dij}. \quad (\text{A2.16})$$

Applying the inverse rotation (eq. A1.6), the mean direction is obtained in the geographical coordinate system. This approximate direction is very close to the mean direction given by the minimal inertia axis of  $\overline{T}_i^R$  (see numerical examples in Tables 5 and 6, Section 6.2).

Using the fact that the inertia tensor is related to the variance tensor, the two concentration parameters  $K_x$  and  $K_y$  can be given by the following very good approximations (within a few %):

$$\begin{aligned} \frac{1}{K_x} &\approx \frac{1}{\cos 2\Omega_a} [S_{Ii}^2 \cos^2 \Omega_a - S_{Di}^2 \sin^2 \Omega_a], \\ \frac{1}{K_y} &\approx \frac{1}{\cos 2\Omega_a} [S_{Di}^2 \cos^2 \Omega_a - S_{Ii}^2 \sin^2 \Omega_a], \end{aligned} \quad (\text{A2.17})$$

where

$$\Omega_a = 0.5 \arctan \left[ \frac{2 S_{IDi}}{S_{Ii}^2 - S_{Di}^2} \right] \quad (\text{A2.18})$$

and

$$\begin{aligned}
 S_{Ii}^2 &= \frac{1}{P_i} \sum_{j=1}^{m_i} P_{ij} (Y_{Iij} - Y_{Ii})^2, \\
 S_{Di}^2 &= \frac{1}{P_i} \sum_{j=1}^{m_i} P_{ij} (Y_{Dij} - Y_{Di})^2, \\
 S_{IDi} &= \frac{1}{P_i} \sum_{j=1}^{m_i} P_{ij} (Y_{Iij} - Y_{Ii})(Y_{Dij} - Y_{Di}), \\
 P_i &= \sum_{j=1}^{m_i} P_{ij}.
 \end{aligned} \tag{A2.19}$$

On the other part, the approximate Fisher concentration factor is estimated from eq. (A1.5) by  $\hat{K}_{ij} = \kappa_{ij}^*$ .

Eqs (A2.16) to (A2.19) lead to results very close to those obtained directly via the inertia tensor (Tables 5 and 6). Of course, if weightings  $P_{ij}$  and concentration parameters  $K_{ij}$  are very different from site to site, the full tensorial approach is needed.

### APPENDIX A 3

#### A3.1 Probability density function of the intensity

Following the same approach as in Section 4, the distribution of the empirical mean intensity  $\bar{F}_i$  at field level is a hierarchical Gaussian statistics around the unknown intensity  $g_i = g_{Fi}$  of the magnetic field, univariate, analogue to eq. (42):

$$p(\bar{F}_i | g_{Fi}) = \frac{\sqrt{m_i}}{\sigma_{BFi} \sqrt{2\pi}} \exp \left[ -\frac{1}{2} m_i \left( \frac{\bar{F}_i - g_{Fi}}{\sigma_{BFi}} \right)^2 \right], \tag{A3.1}$$

where  $\sigma_{BFi}^2$  is the Bayesian variance, which characterizes the dispersion of the empirical mean intensity around the true (unknown) mean intensity. This variance appears as the inverse of a harmonic mean of variances  $\sigma_{Fi}^2$ ,  $\sigma_{Fij}^2$  and  $\sigma_{Fijk}^2$ , which are defined at field, site and sample levels, respectively:

$$\sigma_{BFi}^2 = \frac{m_i}{W_{Fi}} = \frac{m_i}{\sum_{j=1}^{m_i} \frac{P_{ij}}{\sigma_{Fi}^2 \sigma_{Fij}^2 + \sigma_{Fi}^2} + \sum_{k=1}^{r_{ij}} \frac{1}{\frac{\sigma_{Fijk}^2 + e_{ms_{Fijk}}^2}{\sigma_{Fij}^2 + \frac{e_{ms_{Fijk}}^2}{n_{ijk}}}}}. \tag{A3.2}$$

#### A3.2 Estimation of the intensity parameters

(i) At specimen level, an unbiased estimation of the intermeasurement variance  $\sigma_{Fijkl}^2$  will be given by

$$\hat{\sigma}_{Fijkl}^2 = S_{Fijkl}^{2*}, \tag{A3.3}$$

where

$$S_{Fijkl}^{2*} = \frac{1}{d_{ijkl} - 1} \sum_{m=1}^{d_{ijkl}} (F_{ijklm} - \bar{F}_{ijkl})^2 \quad \text{and} \quad \bar{F}_{ijkl} = \frac{1}{d_{ijkl}} \sum_{m=1}^{d_{ijkl}} F_{ijklm},$$

and a mean squared error at specimen level can be calculated:

$$\overline{e_{ms_{ijk}}^2} = \frac{1}{n_{ijk}} \sum_{l=1}^{n_{ijk}} \frac{\sigma_{Fijkl}^2}{d_{ijkl}}.$$

(ii) At sample level, an unbiased estimation of the interspecimen variance  $\sigma_{Fijk}^2$  will be given by

$$\overbrace{\sigma_{Fijk}^2 + e_{ms_{ijk}}^2}^{\wedge} = S_{Fijk}^{2*}, \text{ hence the Bayesian solution: } \hat{\sigma}_{Fijk}^2 = \max(0, S_{Fijk}^{2*} - \overline{e_{ms_{ijk}}^2}), \tag{A3.4}$$

where

$$S_{Fijk}^{2*} = \frac{1}{n_{ijk} - 1} \sum_{l=1}^{n_{ijk}} (\bar{F}_{ijkl} - \bar{F}_{ijk})^2 \quad \text{and} \quad \bar{F}_{ijk} = \frac{1}{n_{ijk}} \sum_{l=1}^{n_{ijk}} \bar{F}_{ijkl}.$$

(iii) At site level, the intersample variance  $\sigma_{Fij}^2$  is deduced from

$$\frac{\hat{1}}{V_{Fij}} = \frac{\hat{1}}{\sum_{k=1}^{r_{ij}} V_{Fijk}} = \frac{\hat{1}}{\sum_{k=1}^{r_{ij}} \frac{1}{\frac{\hat{\sigma}_{Fijk}^2 + e_{ms_{ijk}}^2}{\sigma_{Fij}^2 + \frac{e_{ms_{ijk}}^2}{n_{ijk}}}}} = \frac{S_{Fij}^{2*}}{r_{ij}}, \tag{A3.5}$$

where

$$S_{F_{ij}}^{2*} = \left( \frac{r_{ij}}{r_{ij} - 1} \right) \frac{1}{V_{F_{ij}}} \sum_{k=1}^{r_{ij}} V_{F_{ijk}} (\overline{F_{ijk}} - \overline{F_{ij}})^2 \quad \text{and} \quad \overline{F_{ij}} = \frac{1}{V_{F_{ij}}} \sum_{k=1}^{r_{ij}} V_{F_{ijk}} \overline{F_{ijk}}.$$

(iv) At field level, the intersite total variance ( $\sigma_{F_i}^2 + \sigma_{t_i}^2 g_{F_i}^2$ ) is deduced from

$$\frac{\hat{1}}{W_{F_i}} = \frac{\hat{1}}{\sum_{j=1}^{m_i} W_{F_{ij}}} = \frac{\hat{1}}{\sum_{j=1}^{m_i} \frac{P_{ij}}{\sigma_{t_i}^2 g_{F_i}^2 + \sigma_{F_i}^2 + \frac{1}{V_{F_{ij}}}}} = \frac{S_{F_i}^{2*}}{m_i}, \tag{A3.6}$$

where

$$S_{F_i}^{2*} = \left( \frac{m_i}{m_i - 1} \right) \frac{1}{W_{F_i}} \sum_{j=1}^{m_i} W_{F_{ij}} (\overline{F_{ij}} - \overline{F_i})^2 \quad \text{and} \quad \hat{g}_F(t_i) = \overline{F_i} = \frac{1}{W_{F_i}} \sum_{j=1}^{m_i} W_{F_{ij}} \overline{F_{ij}}.$$

The Bayesian resolution of eqs (A3.5) and (A3.6) is carried out by the bisection method. If all the sites are of the same time  $t_i$ , then  $\sigma_{t_i}^2 = 0$  and the intersite variance  $\sigma_{F_i}^2$  can be estimated.

### A3.3 Confidence interval for the intensity

The confidence interval for intensity is determined using the Student t-distribution of  $\frac{\overline{F_i} - \hat{g}_F(t_i)}{\frac{S_{F_i}}{\sqrt{m_i - 1}}}$ . We have

$$p(\hat{g}_F(t_i) - e_{Y_{F_i}} \leq g_F(t_i) \leq \hat{g}_F(t_i) + e_{Y_{F_i}}) = 1 - \beta, \tag{A3.7}$$

where

$$e_{Y_{F_i}} = t_{\beta/2} \sqrt{\frac{\hat{1}}{W_{F_i}}} = t_{\beta/2} \sqrt{\frac{\hat{\sigma}_{B_{F_i}}^2}{m_i}}. \tag{A3.8}$$

The Student coefficients  $t_{\beta/2}$  are given in Table 2.