Physics of the Earth and Planetary Interiors 170 (2008) 1–14

Contents lists available at ScienceDirect

Physics of the Earth and Planetary Interiors

journal homepage: www.elsevier.com/locate/pepi

Historical variation of the geomagnetic axial dipole

Christopher C. Finlay

Institut für Geophysik, ETH Zürich, Zürich CH-8093, Switzerland

ARTICLE INFO

Article history:
Received 11 January 2008
Received in revised form 19 June 2008
Accepted 25 June 2008

Keywords:
Geomagnetism
Secular variation
Axial dipole
Historical field modelling
Archaeointensity

ABSTRACT

The geomagnetic axial dipole (hereinafter denoted g_1^0) is the largest component of our planet's magnetic field. Its magnitude determines the morphology of solar-terrestrial electrical current systems and it is the most fundamental diagnostic property of the core-generated geodynamo. Elucidating past and future variations of $g_1^0(t)$ is consequently of central importance in geomagnetism. Previous historical geomagnetic field models, such as *gufm1* of Jackson et al. [Jackson, A., Jonkers, A.R.T., Walker, M.R., 2000. Four centuries of geomagnetic secular variation from historical records. Philos. Trans. R. Soc. Lond. A 358, 957–990], used direct observations to constrain $g_1^0(t)$ only after 1840 A.D.; before this time a crude linear extrapolation of the post-1840 A.D. rate of change (15 nT/year) was employed. In this contribution I construct historical field models with $g_1^0(t)$ instead constrained from 1590 A.D. to 1840 A.D. by an archaeointensity dataset compiled by Korte et al. [Korte, M., Genevey, A., Constable, C.G., Frank, U., Schnepp, E., 2005. Continuous geomagnetic field models for the past 7 millennia. 1. A new global data compilation. Geochem. Geophys. Geosyst. 6, doi:10.1029/2004GC000800]. A range of possible linear models of the form $g_1^0(t) = g_1^0(1840) + \beta(t - 1840)$ are first explored; $\beta = 2.74 \pm 42.32$ nT/year is found to explain the archaeointensity dataset with maximum likelihood, consistent with the recent findings of Gubbins et al. [Gubbins, D., Jones, A.L., Finlay, C.C., 2006. Fall in Earth's magnetic field is erratic. Science 312, 900–902]. Relaxing the linear constraint in an effort to find more physically plausible models, I find it is necessary to artificially increase the weight given to the archaeointensity data in order to obtain acceptable models. Despite satisfactorily explaining both the historical and archaeointensity data, and possessing reasonable spatial and temporal complexity, such free evolution models perform worse than the simpler linearly constrained models when tested against the independent dataset of Gallet et al. [Gallet, Y., Genevey, A., Fluteau, F., 2005. Does Earth's magnetic field secular variation control centennial climate change? Earth Plan. Sci. Lett. 236, 159–173]. Bayesian model comparison techniques indicate that a model (*gufm1-g10c*) involving no change in $g_1^0(t)$ between 1590 A.D. and 1840 A.D. is most probable given the presently available data and current modelling techniques. I propose that this new, empirically derived, constraint on the evolution of the geomagnetic axial dipole be incorporated into the next generation of historical field models.

1. Introduction

Knowledge concerning the evolution of Earth's magnetic field can be conveniently represented in the form of time-dependent spherical harmonic field models. Such models are derived from large compilations of geomagnetic field observations during times when sufficient geographical coverage of data is available. They are the basis for testing proposed mechanisms of field evolution (secular variation) and they underlie ongoing attempts to understand and predict future field behaviour.

At present, the most widely used historical field model is *gufm1* (Jackson et al., 2000) which covers the interval from 1590 A.D. to 1990 A.D. and was constructed from more than 365,000 direct observations using a time-dependent, regularized, inversion technique (Bloxham and Jackson, 1992; Jackson et al., 2000). A major deficiency in the dataset used to construct *gufm1* is that prior to 1840 A.D. it contains no direct observations of field intensity. As has been pointed out by Jackson (2000), this implies that the absolute magnitude of field (and hence g_1^0) prior to 1840 A.D. can only be known to within a time-varying scale factor. To avoid this ambiguity Jackson et al. (2000) and earlier Bloxham and Jackson (1992) followed Barraclough (1974) and prescribed that $g_1^0(t) = g_1^0(1840) + \beta(t - 1840)$ with $\beta = 15$ nT/year before 1840 A.D. Barraclough (1974) justified this model on the grounds that it was the extrapolation of a straight line fit to g_1^0 values from single epoch field models spanning 1840–1970 A.D. when direct intensity measurements were available.

Clearly this assumption is a very crude device; $\mathrm{d}g_1^0(t)/\mathrm{d}t = \dot{g}_1^0(t)$ has varied considerably since 1840 A.D. (see, for example, Fig. 1 in

E-mail address: cfinlay@erdw.ethz.ch.

0031-9201/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.pepi.2008.06.029

Jackson, 2000) and there is no physical reason to suppose it should have been constant at earlier times. Furthermore, palaeomagnetic investigations of changes in field intensity on longer millennial to million year timescales (Hongre et al., 1998; Korte and Constable, 2005b; Valet, 2003) as well as theoretical investigations into the nature of the geodynamo process in the low Ekman number regime appropriate for Earth's core (Zhang, 2000; Zhang and Gubbins, 2000) suggest that the axial dipole field undergoes fluctuations across a range of timescales. Given the undesireable nature of the linear extrapolation approach, it is evident that new strategies are need in order to find better models for the evolution of the geomagnetic axial dipole during the historical era.

Indirect palaeointensity or archaeointensity measurements (see, for example, Tauxe and Yamazaki, 2007; Constable, 2007), which are sensitive to g_1^0, can in principle be used as an additional observational constraint to improve on Barraclough's proposal. Gubbins et al. (2006) recently proposed a simple strategy for combining directional information from existing historical field models and archaeointensity measurements. The key step in their method involved calculating the ratio of archaeointensity observations to the intensities predicted by a historical field model such as *gufm1*. By multiplying *gufm1*'s prediction for g_1^0 at the observation time by the ratio of the predicted to observed intensities, they obtained 315 new estimates for g_1^0 between 1590 A.D. and 1840 A.D. They then carried out a total least squares straight line fit to these estimates, using $g_1^0(1840)$ from *gufm1* as an anchor point, and found a best fitting slope of $\beta = 2.28\,\mathrm{nT/year} \pm 2.72\,\mathrm{nT/year}$. The stark contrast of this result compared to the average post-1840 A.D. rate of change of 15 nT/year illustrated that major changes in the mean rate of g_1^0 have occurred on centennial timescales; Gubbins et al. (2006) labelled this as 'erratic' behaviour.

A graphical summary of previously published models of $g_1^0(t)$ spanning the past 400 years, comprising both historical and archaeomagnetic field models is presented in Fig. 1. The differences between these models reflect differences in datasets and differences in the modelling procedures adopted. Notice that the most recent models (Korte and Constable, 2005a; Gubbins et al., 2006) show a less rapid decrease in the magnitude of the axial dipole moment in the 16th, 17th and 18th centuries compared with the earlier models of Hongre et al. (1998) and Jackson et al. (2000).

The recent studies by Gubbins et al. (2006) and Korte and Constable (2005a,b, 2006a) have demonstrated that the assumed $g_1^0(t)$ pre-1840 A.D. in historical field models such as *gufm1* needs to be reconsidered. In this study, I address this issue by inverting for new historical field models with $g_1^0(t)$, for $1590 \leq t \leq 1840$, constrained by the archaeointensity dataset of Korte et al. (2005). By comparing such models I determine those historical variations of the axial dipole field that are robust and should be interpreted with confidence. The latter point is important, for example, in evaluating the suggestion that historical changes in Earth's geomagnetic dipole field may have contributed towards past climatic variations (Gallet et al., 2005; Courtillot et al., 2007a,b).

In Section 2, I outline the observations used in the construction of historical field models and discuss in detail the additional archaeointensity dataset compiled by Korte et al. (2005) used in this study. The troublesome issue of error budget allocation for the archaeointensity dataset is discussed before a summary of the field modelling methodology is then given in Section 3. A suite of new historical field models determined from the inversion of both the Korte et al. (2005) archaeointensity data spanning 1590 A.D. to 1840 A.D. and the historical and modern data used previously by Jackson et al. (2000) are presented in Section 4. I report results from both models in which the evolution of $g_1^0(t)$ is forced to take a linear form and also models without any explicit restriction on the form of $g_1^0(t)$ but involving a range of possible weights for the archaeointensity data. The fit of these models to historical datasets as well as to the independent dataset of Gallet et al. (2005) are presented and results of Bayesian model comparison calculations are reported. The interpretation of the results including comparisons with previous models, implications for the physical mechanisms producing axial dipole field change, compatibility with proposed tilted dipole mechanisms for archaeomagnetic jerks, as well as perspectives for future data collection strategies and new historical field models are set out in Section 5. A summary of the findings in Section 6 concludes the paper.

2. Data and error estimates

The inversions for historical field models reported here are based primarily on the direct observations previously used to con-

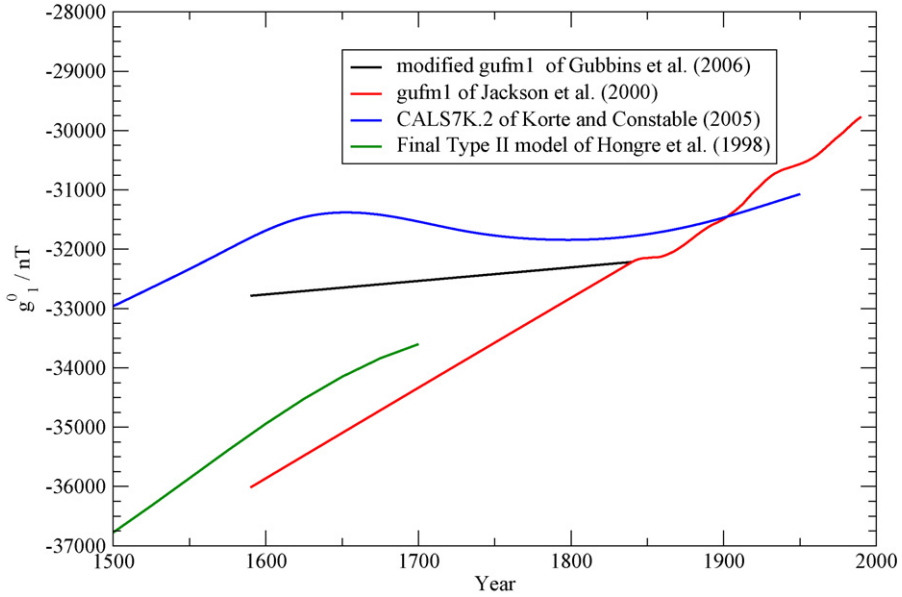

Fig. 1. Previous models for $g_1^0(t)$ showing the 15 nT/year rate of change employed by Jackson et al. (2000), the correction suggested by Gubbins et al. (2006) with a much smaller linear rate of change pre-1840 A.D., and the predictions of the archaeomagnetic field models of Korte and Constable (2005a) and Hongre et al. (1998).

struct the model *gufm1* by Jackson et al. (2000). These include the pre-20th century sources described in the review paper of Jonkers et al. (2003), together with first-differences of observatory annual means post-1840 A.D., 20th century survey data and datasets from the POGO, DE2 and MAGSAT satellites. Further details of these datasets and how error estimates were assigned to them can be found in the series of papers by Bloxham et al. (1989), Bloxham and Jackson (1989, 1992) and Jackson et al. (2000).

The solitary additional dataset used in the inversions presented here is a subset of the archaeomagnetic intensity database compiled by Korte et al. (2005), but restricted to the interval spanning 1590–1840 A.D. Jones (2005) checked this dataset against original references and corrected three minor transcription errors. The resulting set of 315 intensity observations with associated age and measurement error estimates is identical to that previously used by Gubbins et al. (2006).

This archaeointensity dataset contains measurements made by a large number of authors using a range of experimental procedures, who often presented their data in different ways. For an overview of the merits of the various experimental archaeointensity protocols, the interested reader should consult the reviews by Valet (2003) and Tauxe and Yamazaki (2007). A complementary account comparing a number of palaeointensity techniques applied to archaeological samples is given by Donadini et al. (2007). Korte et al. (2005) made a commendable effort to distill from such diversity a dataset that is internally consistent in its error estimations. Here, I briefly summarize how they arrived at their error estimates, since this will prove important for the interpretation of the resulting field models reported in Section 4.

Korte et al. (2005) developed criteria for the consistent allocation of measurement errors depending on the measurement method used, the relative dispersion of the mean of the measurements of different samples from the same site, whether tests for alteration were carried out, whether the TRM anisotropy effect was taken into account, the number of samples per site and the number of sub-samples analyzed per sample. It is worth emphasizing that they did not rely on the error estimates reported in the original papers. They further suggested that minimum measurement uncertainty estimates of 5% were appropriate on the basis of comparisons of the archaeointensity data to *gufm1* predictions for field intensity post-1840 A.D. The dataset used in the present study (spanning only 1590–1840 A.D.) contains estimated measurement errors ranging between 2.3 μT and 13.6 μT.

Archaeointensity data differ from the direct observations in that they also contain errors in their ages (an independent variable in field modelling). Korte et al. (2005) derived error estimates for the ages of the observations in their dataset, where possible using estimates from the original literature and re-calibrating Carbon-14 ages using the method of Stuiver and Reimer (1993). If no age uncertainties were given by the authors then uncertainties of 100, 50 or 10 years were assigned depending on the dating technique employed. The dataset used in this study (spanning 1590–1840 A.D.) contains estimated age errors ranging between 1 and 300 years. Gubbins et al. (2006) dealt with the challenge of accounting for age errors by employing a total least squares approach, finding the best-fitting straight line which minimized a weighted combination of age and measurement misfits. Unfortunately such an approach is incompatible with the B-spline temporal expansion method employed during time-dependent field modelling (see next section).

In Section 4.1, results from two alternative approaches for dealing with age errors are reported. The first (referred to here as AGE-CIT) involves an approach analogous to that adopted by Jackson et al. (2000) for dealing with location inaccuracies in maritime observations. Denoting total intensity by F, it includes using the present iteration of the field model to estimate $\partial F/\partial t$ at the observation location and time, then effective intensity errors are calculated from age errors using the linear approximation $\Delta F_t \simeq (\partial F/\partial t)\Delta t$. The second approach (referred to here as AGE-KC) follows the suggestion of Korte et al. (2005) and Korte and Constable (2005a) who encountered the same problem during the construction of their CALS7K.2 archaeomagnetic field model. They adopted a qualitative scale whereby for age uncertainties less than 10 years no extra intensity error was allocated ($\Delta F_t = 0\,\mu$T), for an age error of 10–50 years an effective intensity error of $\Delta F_t = 1.5\,\mu$T was allocated, for an error of 50–100 years $\Delta F_t = 2.5\,\mu$T was allocated, for an age error of 100–500 years $\Delta F_t = 5\,\mu$T was allocated, while if the age error was greater than 500 years the observation was rejected. In both cases, the estimated intensity error coming from age uncertainty (ΔF_t) and that from measurement uncertainty (ΔF_m) were added in quadrature to obtain a final intensity error estimate $\Delta F = \sqrt{\Delta F_m + \Delta F_t}$. In Section 4.1 the performance of these two approaches will be compared.

With the known sources of error in the archaeointensity data all mapped into an intensity error the archaeointensity dataset described above is now suitable to be included alongside historical datasets in inversions for time-dependent historical field models. An outline of the regularized inversion procedure used to derive such field models is given in the next section.

3. Field modelling methodology

The regularized inversion technique for time-dependent field modelling employed in this study was developed and previously used by Bloxham and Jackson (1992) and Jackson et al. (2000); these references should be consulted for a detailed exposition of the method as only a brief outline is given here.

Assuming the mantle to be (to a first approximation) an insulator, then the magnetic field outside Earth's core can be expressed as the gradient of a scalar potential V which satisfies Laplace's equation and can be expanded in a basis of spherical harmonics at a specified radius r such that

$$V(r,\theta,\phi,t) = a\sum_{l=1}^{L}\sum_{m=0}^{l}\left(\frac{a}{r}\right)^{l+1} g_l^m(t)Y_l^m(\theta,\phi), \tag{1}$$

where $Y_l^m(\theta,\phi)$ are the Schmidt quasi-normalized spherical harmonics of degree l and order m, L is the degree of truncation of the expansion and $a = 6371.2$ km is Earth's mean surface radius. Here the spherical harmonic expansions are truncated at degree $L = 14$, by which point the power in the model is already negligible due to the regularization applied (see below). The Gauss coefficients $g_l^m(t)$ are further expanded in a fourth order (cubic) B-spline basis such that

$$g_l^m(t) = \sum_n g_l^{mn} M_n(t), \tag{2}$$

where $M_n(t) > 0$ if $t \epsilon |t_n, t_{n+4}|$ and is zero otherwise. 163 knots are erected with a 2.5 years spacing to span the 400-year interval from 1590 A.D. to 1990 A.D. The resulting field model consists of 36512 free parameters. The observed magnetic field components X (northward field), Y (eastward field), Z (vertical field), D (declination), I (inclination), F (total intensity) and H (horizontal intensity) can straightforwardly be written in terms of the potential V and its gradients (see, for example, Bloxham et al., 1989, p. 418) so observables predicted by a given set of field model coefficients can easily be obtained. If the field observations are listed in a vector $\mathbf{d}$, and the model coefficients are listed in a vector $\mathbf{m}$ then the forward problem can be written in matrix form as

$$\mathbf{d} = \mathbf{f}(\mathbf{m}) + \mathbf{e}, \tag{3}$$

where **f** is the non-linear functional relating the data to the model and **e** is an error vector which gives the residual between the model predictions and the observations.

The inverse problem of geomagnetic field modelling involves finding the model **m** (consisting of the parameters g_l^{mn}) that best describes the evolution of the magnetic field at the core surface given the available geomagnetic observations. This problem is highly non-unique: many possible field models are capable of fitting the observations to within their estimated errors. Here, a regularization or minimum norm strategy is adopted (following Shure et al., 1982; Gubbins, 1983; Gubbins and Bloxham, 1985) that allows smooth, well-converged, field models to be obtained at the core surface. The technique requires the calculation of measures of model complexity; these are chosen to be the Ohmic dissipation norm of Gubbins (1975) measuring spatial field complexity

$$\mathbf{m}^{\mathrm{T}}\mathbf{S}^{-1}\mathbf{m} = \frac{4\pi}{t_e - t_s}\int_{t_s}^{t_e}\sum_{l=1}^{L} f(l)\sum_{m=0}^{l}[(g_l^m)^2 + (m_l^m)^2]\,\mathrm{d}t, \tag{4}$$

with

$$f(l) = \frac{(l+1)(2l+1)(2l+3)}{l}\left(\frac{a}{c}\right)^{2l+4}, \tag{5}$$

and the temporal curvature norm measuring temporal complexity

$$\mathbf{m}^{\mathrm{T}}\mathbf{T}^{-1}\mathbf{m} = \frac{1}{t_e - t_s}\int_{t_s}^{t_e}\oint_{\mathrm{CMB}}(\partial_t^2 B_r)^2\,\mathrm{d}\Omega\,\mathrm{d}t. \tag{6}$$

where t_s and t_e denote that start and end times for the model. Estimates of the total errors associated with each type of observation (including the effects of crustal fields, external fields, measurement errors, mislocation errors and dating errors) are used to build a data covariance matrix $\mathbf{C}_e$ which enters naturally in the generalized least-squares inversion formalism (see, for example, Gubbins, 2004, pp. 97–100). Further details regarding the form of the covariance matrix used for the historical data are given by Jackson et al. (2000). For the archaeointensity data, errors are assumed uncorrelated so the covariance matrix takes a diagonal form. In Section 4.2 the consequences of 'over-weighting' the archaeointensity data are discussed. By 'over-weighting' I mean that the diagonal elements of the matrix $\mathbf{C}_e^{-1}$ for this dataset are multiplied by a factor W_f^2 so that the error estimates are multiplied by a factor $1/W_f$ and the weighting given to the data is increased by a factor W_f.

Solving the regularized inverse problem involves finding a model that minimizes the objective function Θ defined as

$$\Theta(\mathbf{m}) = [\mathbf{d} - \mathbf{f}(\mathbf{m})]^{\mathrm{T}}\mathbf{C}_e^{-1}[\mathbf{d} - \mathbf{f}(\mathbf{m})] + \mathbf{m}^{\mathrm{T}}\mathbf{C}_m^{-1}\mathbf{m}, \tag{7}$$

with the model covariance matrix taking account of both the spatial and temporal norms being

$$\mathbf{C}_m^{-1} = (\lambda_S\mathbf{S}^{-1} + \lambda_T\mathbf{T}^{-1}), \tag{8}$$

where λ_S is the spatial damping parameter (chosen following Jackson et al., 2000, to be $1\times10^{-12}\,\mathrm{nT}^{-2}$) and λ_T is the temporal damping parameter (chosen following Jackson et al., 2000, to be $5\times10^{-4}\,\mathrm{nT}^{-2}\,\mathrm{year}^{-4}$). The optimization problem of finding the model that minimizes Θ is solved numerically using an iterative quasi-Newton scheme (LSQN); this is necessary because some of the observables (F, H, D and I) depend non-linearly on the model parameters g_l^{mn}.

The models discussed in Section 4.1 are constrained so that $g_1^0(t)$ evolves linearly before 1840 A.D. This constraint is implemented by creating synthetic g_1^0 data for each year working back from 1840 A.D. to 1590 A.D., derived using the specified linear rate of change (β). Models which depart from these synthetic data are then penalized via an extra penalty term which is added to the objective function Θ.

Having described both the data and methodology employed in the historical field modelling, in the next section the results of new inversions are reported. First in Section 4.1 field models with linearly constrained $g_1^0(t)$ are described from a maximum likelihood perspective. Section 4.2 reports models obtained when the linear constraint is released, and archaeointensity data alone constrain $g_1^0(t)$ pre-1840 A.D. In Section 4.3 the fit of the models to datasets used in the inversions are reported, while Section 4.4 describes the fit of models to the independent dataset of Gallet et al. (2005). Section 4.5 gives an illustration of how Bayesian model comparison may be used to determine which of the reported field models is the most probable. Detailed interpretation and discussion of the results is delayed until Section 5.

4. Results

4.1. Linear constraint on $g_1^0(t)$ between 1590 A.D. and 1840 A.D.

As a first step, I constrain the axial dipole to evolve in a linear manner between 1590 A.D. and 1840 A.D., as specified by the equation

$$g_1^0(t) = g_1^0(1840) + \beta(t - 1840) \quad \text{with}$$
$$-150\,\mathrm{nT/year} \le \beta \le 150\,\mathrm{nT/year}. \tag{9}$$

In the inversions reported in this section, data rejection at the $3\sigma_i$ level was applied to the historical data (as in *gufm1*), but no rejection of archaeointensity data was carried. This facilitates a meaningful comparison of fit of the various models to the archaeointensity data. In reporting the inversion results it is useful to refer to a L2-norm measure of rms misfit between model predictions and the observations that is normalized by error estimates σ_i such that

$$M_2 = \frac{\chi}{\sqrt{N}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(d_i - [f(m)]_i)^2}{\sigma_i^2}} \tag{10}$$

where d_i is the ith observation, $[f(m)]_i$ is the prediction of the field model at the time and location of the ith observation and N is the number of observations considered. Fig. 2 shows how this normal-

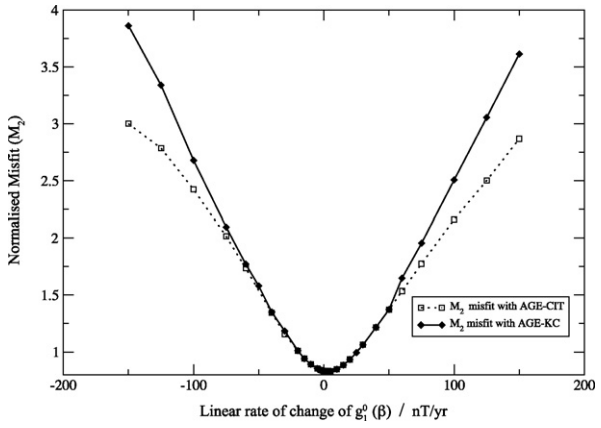

Fig. 2. Normalized root mean square misfit (M_2) between model predictions and the archaeointensity dataset described in Section 2 as a function of the imposed linear rate of change of $g_1^0(t)$ between 1590 A.D. and 1990 A.D. (i.e. β). The two curves show two different methods for converting age errors into effective intensity errors: AGE-CIT uses the value of $\partial F/\partial t$ at the observation location and time from the current model iterate while AGE-KC uses an a priori criteria (see Section 2 for further details).

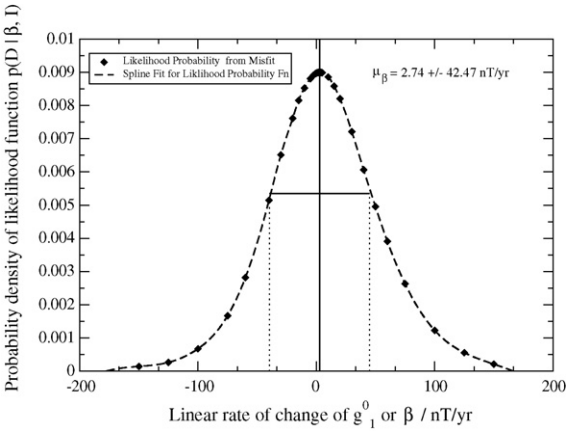

Fig. 3. Probability density function describing the likelihood of the 1590–1840 A.D. archaeointensity dataset of Korte et al. (2005) being explained by different values of β. Squares are results of inversions carried out with different values of β and the dashed line is a cubic B-spline fit to these. The maximum likelihood value of β with error bounds corresponding to 67% probability is $\mu_\beta = 2.74 \pm 42.27$ nT/year.

ized misfit M_2 for the archaeointensity dataset varies as a function of β, for the two different methods of converting age error estimates into intensity error estimates (AGE-CIT and AGE-KC) discussed in Section 2.

It can be seen that the two approaches give rather similar results regarding the value of β with smallest misfit and the range of values of β that fit the data satisfactorily ($M_2 < 1$). Differences between the two methods are found only when the chosen value of β is not compatible with the data. Consequently, in subsequent discussions only the results obtained using simpler method (AGE-KC) are considered. Presented in Fig. 3 is the probability density function describing the likelihood of obtaining the observed archaeointensity data, as a function of β. The continuous function plotted was obtained by fitting cubic B-splines to the values of $p(\mathbf{D}|\beta, I)$ (the conditional probability of the observed archaeointensity data $\mathbf{D}$ being obtained given a value for β and the setup of the inversion experiment I) which are shown as squares in the plot. $p(\mathbf{D}|\beta, I)$ is related to χ^2 and hence to M_2 through the expression

$$p(\mathbf{D}|\beta, I) = C\, \mathrm{e}^{-(\chi^2(\beta)/2)}, \tag{11}$$

where C is a positive constant, chosen so the probability density function integrated over all values of β equals 1.0.

The value of β with maximum likelihood (denoted by μ_β) occurs at 2.74 nT/year. An estimate of the error bar on this quantity can be obtained from the smallest range of β within which 67% of the probability density function lies (e.g. Sivia and Skilling, 2006, p. 22). This approach leads to the maximum likelihood estimate for β of $\mu_\beta = 2.74 \pm 42.27$nT/year. This result is compatible with the earlier study of Gubbins et al. (2006) who arrived at the estimate $\mu_\beta = 2.28 \pm 2.72$ nT/year. Although the estimated values of μ_β obtained here and that found by Gubbins et al. are in agreement, the difference in the error bounds is striking. It is therefore worth describing three factors that contribute to this discrepancy.

Firstly, Gubbins et al. (2006) calculated their error bars indirectly from g_1^0 values (and errors estimates) propagated from the original data using the ratio of the observed to *gufm1* predicted intensity. In contrast I calculated error bounds on μ_β directly from the likelihood probability density function for β, by finding the range of values of β which account for the archaeointensity data with a probability greater than 0.67. The propagation of errors to g_1^0 could lead to an under-estimate of the true error on β by Gubbins et al. (2006).

Secondly, the approach of Gubbins et al. (2006) explicitly took into account age errors via the total-least squares procedure of Williamson (1968). In contrast, I mapped age errors into intensity errors because the B-spline field modelling formalism is not capable of dealing with errors in the independent variable. This could conceivably lead to over-estimated errors associated with age inaccuracy, hence to over-estimated errors for μ_β. Two different mapping schemes (AGE-CIT and AGE-KC) were employed to try to combat this problem, each yielded identical error estimates on μ_β giving some indication that results found here are sensible.

Finally, Gubbins et al. (2006) artificially fixed the non-axial dipole spherical harmonic coefficients in their study. Such an approach is only strictly valid when the directional data used to construct the historical field model have perfect geographical coverage (Hulot et al., 1997). Here, I adopted a more self-consistent procedure; full inversions were carried out to find the field model that best fitted the archaeointensity data, given a particular value of β and a particular model regularization. The extra freedom in this approach undoubtedly permits a larger range of β to be consistent with the data and this will contribute to the discrepancy in the error estimates on μ_β. To summarize, I argue that the approach presented here is a more direct method of testing whether a particular linear constraint on $g_1^0(t)$ is compatible with both the historical and archaeointensity observations. The consequences of large error bounds on μ_β for model selection and for physical interpretations will be discussed in Sections 4.5 and 5.

It is also important to remark here that models with $\beta = 0$ nT/year are almost as probable as the maximum likelihood solution of $\mu_\beta = 2.74$ nT/year. This is relevant because $\beta = 0$ nT/year represents a simpler model (no change) than a linear variation model. Henceforth, the model constructed here with $\beta = 0.0$ nT/year 1590–1840 A.D. will be referred to as *gufm1-g10c*. Whether or not this model should be preferred to the maximum likelihood solution is discussed in Section 4.5.

4.2. Models without linear constraint on the evolution of $g_1^0(t)$

The obvious next step is to determine whether it is possible to go beyond a simple linear model for $g_1^0(t)$ and test whether archaeointensity data, together with model regularization, are sufficient to determine a physically plausible, freely evolving, $g_1^0(t)$. In this scenario, the g_1^{0n} spline coefficients are determined in the same way as all the other model coefficients through minimization of the objective function Θ defined in (7). It was found necessary to also reject archaeointensity data more than $3\sigma_i$ from current model iterate; the maximum likelihood model from Section 4.1 was used as the starting model. Convergence of the quasi-Newton iterative scheme typically occurred after three to four iterations. Illustrative examples of solutions for $g_1^0(t)$ found in a series of inversions with different weighting factors W_f applied to the archaeointensity data (see Section 3) are presented in Fig. 4.

Finding acceptable solutions while allowing $g_1^0(t)$ to evolve freely turns out to be somewhat difficult. Trouble arises because the constraint provided by the archaeointensity data is rather weak. This is a consequence of the larger error estimates for these data compared to those for historical data and because the spatial and temporal regularization has been designed to balance the desire to fit the historical data, not the weaker archaeointensity data. Trials showed that inversions using the archaeointensity error estimates suggested by Korte et al. (2005) resulted in unrealistic $g_1^0(t)$ histories that artificially decayed before 1840 A.D., when the archaeointensity data should have constrained the evolution of $g_1^0(t)$. This undesirable effect occurs because the objective function Θ can be minimized by adopting low values for g_1^{0n} (resulting in a small spatial Ohmic heating norm), while the penalty for not fitting the archaeointensity data is not

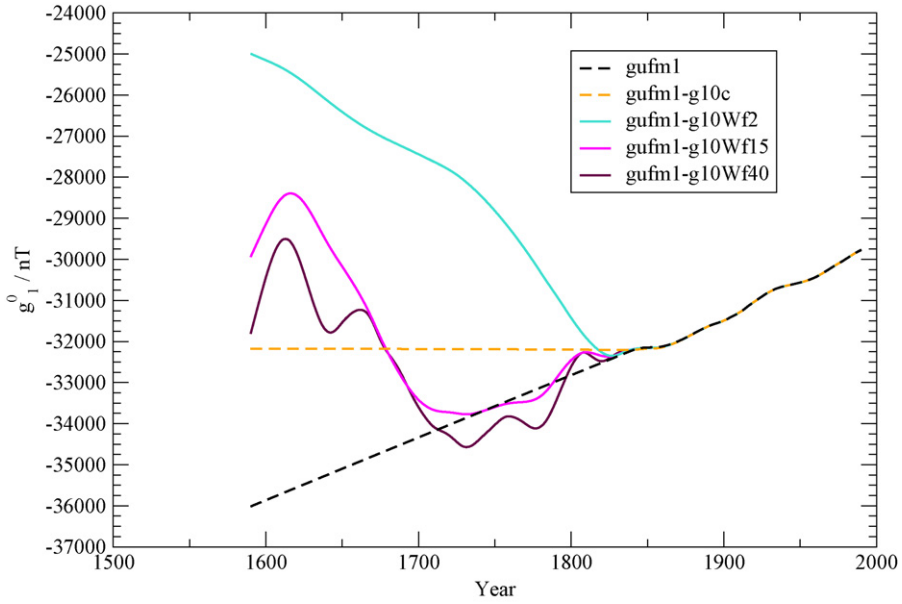

Fig. 4. Examples of $g_1^0(t)$ from three field models constructed without any linear constraint applied and with three different choices of weighting of the archaeointensity data (*gufm1-g10Wf2*, *gufm1-g10Wf15* and *gufm1-g10Wf40*). Also shown for reference are *gufm1* which had a $\beta = 15$ nT/year linear constraint imposed and *gufm1-g10c* which had $\beta = 0$ nT/year imposed.

prohibitive because archaeointensity error estimates are rather large.

In an effort to force the field models to better fit the archaeointensity data, but without explicitly changing the global regularization, I explored artifical over-weighting of the archaeointensity dataset by a factor W_f. This was simply achieved by dividing archaeointensity error estimates by W_f (see Section 3). The introduction of W_f is certainly ad hoc and a horribly artificial procedure. Nevertheless, it was considered important to determine whether field models fitting both the historical and archaeointensity data and possessing physically plausible $g_1^0(t)$ histories could be found. Furthermore, the error estimates chosen by Korte et al. (2005) apparently err on the side of being cautious, so it is reasonable to experiment with over-weighting the archaeointensity data, in fact Korte and Constable, 2005a, also found it necessary to do this.

Considering the post-1840 total intensity measurements used to construct *gufm1*, a mean intensity error estimate of 234.59 nT was allocated; this is a factor of 39.3 times smaller than the mean error estimate (including age errors) used in the present study for archaeointensity data (9225.94 nT). Therefore $W_f = 40$ represents the maximum factor by which archaeointensity data could sensibly be over-weighted assuming the mean accuracy of the historical intensity measurements is better than the mean accuracy of the archaeointensity data.

The $g_1^0(t)$ histories resulting from inversions with illustrative examples of $W_f = 2$, 15 and 40 are shown in Fig. 4 together with $g_1^0(t)$ from *gufm1* (where $\beta = 15$ nT/year) and from *gufm1-g10c* (where $\beta = 0.0$ nT/year). For reference purposes, the models with $W_f = 2$, 15 and 40 will be henceforth be referred to as *gufm1-g10Wf2*, *gufm1-g10Wf15* and *gufm1-g10Wf40*, respectively.

Considering $g_1^0(t)$ for $W_f = 2$ it is apparent that the unrealistic decay of $g_1^0(t)$ before 1840 A.D. is still a problem. However, when $W_f = 15$ and 40 more plausible histories, involving an oscillation of g_1^0 (t) around the solution of $\beta = 0.0$ nT/year, are found. For these solutions g_1^0 is typically higher than its 1840 A.D. level before 1675 A.D. and lower after 1675 A.D. The larger W_f is made, the higher the amplitude and shorter the timescale of oscillations in $g_1^0(t)$ become, as the model is forced more and more strongly to fit the archaeointensity data. This causes models to become increasingly complex as measured by the Ohmic heating norm (4).

In Fig. 5 the temporal evolution of the Ohmic heating norm (4) is presented for the field models previously shown in Fig. 4. Model *gufm1-g10Wf40* with $W_f = 40$ (strongly forced to fit the archaeointensity data) is found to possess a very complex field structure, with the Ohmic heat norm taking values up to twice that in the linearly constrained models and much greater than the values found post-1840 A.D. Model *gufm1-g10Wf2* has Ohmic heating norm much lower than linearly constrained models, reflecting the unphysical decay in $g_1^0(t)$ found in Fig. 4 and discussed earlier. Finally, model *gufm1-g10Wf15* has Ohmic heating norm rather similar to that of the linearly constrained models, with neither rapid growth nor decay before 1840 A.D. It is therefore considered to be the most physically plausible of the free evolution models discussed here.

Fig. 6 further illustrates the point regarding changes in the spatial complexity of the field models with W_f. It shows snapshots of vertical field (Z) at the core surface in 1650 for each of the models presented in Figs. 4 and 5. Note that the contour levels are the same in all the plots. Interestingly, the field morphology is remarkably

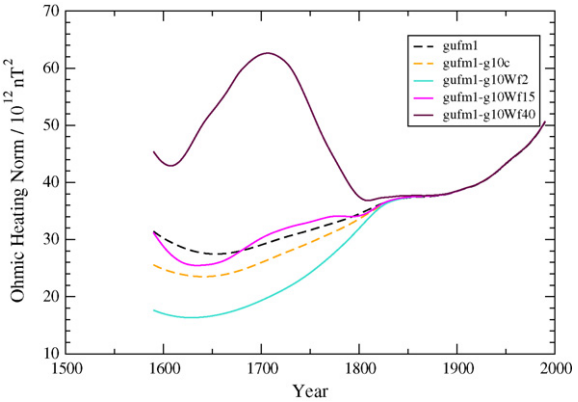

Fig. 5. Temporal evolution of the Ohmic heating norm (a measure of spatial complexity defined in eqs. (4) and (5)) for the models previously presented in Fig. 4.

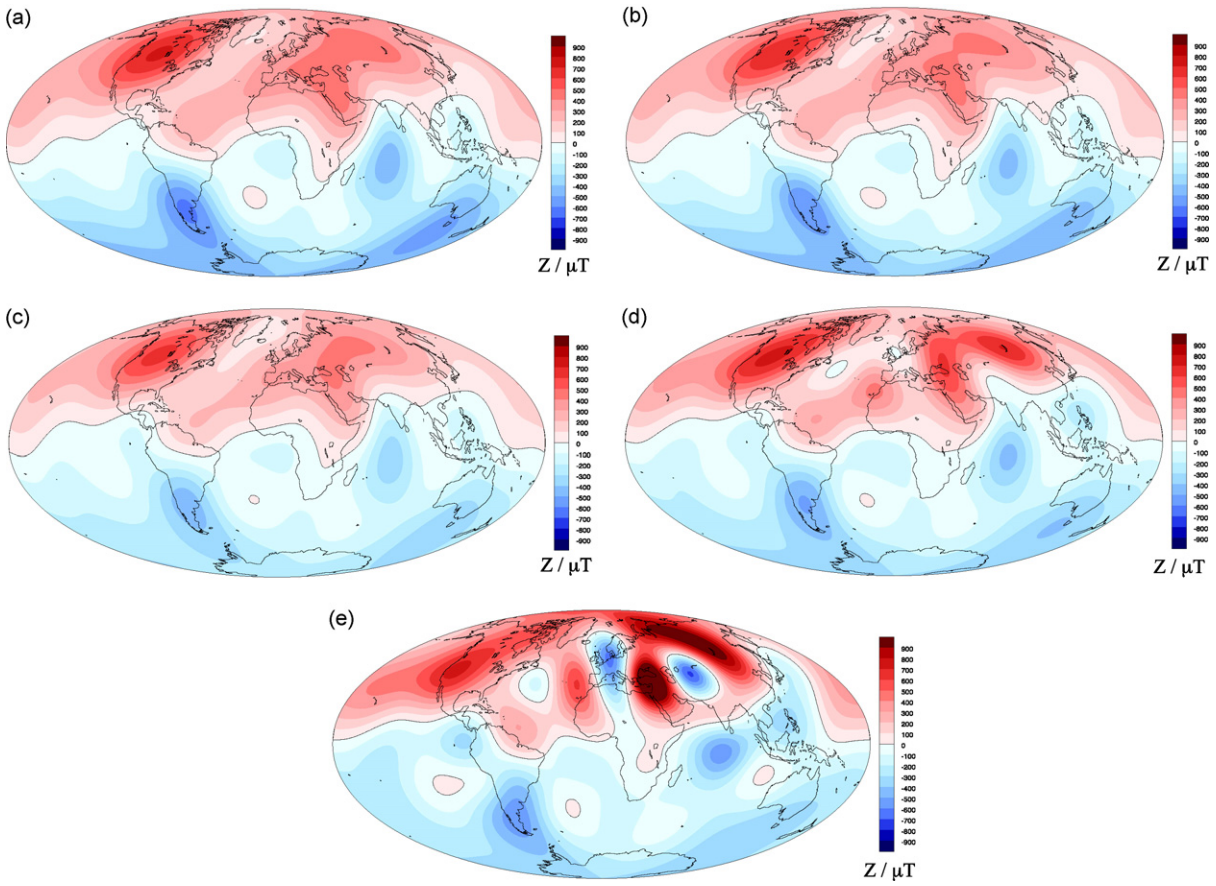

Fig. 6. Contour maps of the vertical (Z) component of the core surface geomagnetic field in 1650 A.D. in units of μT. The map projection is Hammer–Aitoff equal area. These plots were produced using the magmap and colour plotting software developed by R.L. Parker. (a) Shows *gufm1* in 1650, (b) shows *gufm1-g10c* in 1650, (c) shows *gufm1-g10Wf2* in 1650, (d) shows *gufm1-g10Wf15* in 1650 and (e) shows *gufm1-g10Wf40* in 1650.

similar in all cases, with differences between *gufm1*, *gufm1-g10c* and *gufm1-g10Wf2* very minor. *gufm1-g10Wf15* has relatively high amplitude non-dipole field features in 1650 A.D., particularly in the northern hemisphere, but the field morphology still looks similar to that familiar from *gufm1*. In *gufm1-g10Wf40* the amplitude and gradients of the non-dipole field features are much more prominent, with some additional reverse flux features becoming visible. Nonetheless, the morphology and location of field features remains largely similar to that found in the other models. The higher amplitude non-dipolar field structure found in *gufm1-g10Wf40* are apparently necessary to fit the archaeointensity data precisely. The similarity of the field morphologies to the post-1840 A.D. field of *gufm1* demonstrates that moderate (up to say $W_f = 20$) over-weighting of the archaeointensity data still results in plausible core surface field structures.

In order to find historical field models that have $g_1^0(t)$ before 1840 A.D. determined by archaeointensity data, it therefore appears necessary to artificially over-weight the archaeointensity data relative to the error estimates suggested by Korte et al. (2005). If one does not over-weight the data in this manner, $g_1^0(t)$ artificially decays before to 1840 A.D. as a result of the regularization applied during field modelling. On the other hand, if one over-weights the archaeointensity data too much, unphysically large variations in the field complexity develop. Of the models investigated here that fit the archaeointensity and historical data satisfactorily, *gufm1-g10Wf15* was found to have a temporally smooth $g_1^0(t)$ evolution, as well as physically plausible Ohmic heating and core surface field morphology. It is therefore preferred to *gufm1-g10Wf2* and *gufm1-g10Wf40*. In the following sections, I will discuss whether this model should also be preferred over simpler field models with linearly constrained $g_1^0(t)$.

4.3. Comparison of models to datasets included in inversions

A number of field models with different pre-1840 A.D. $g_1^0(t)$ have been reported in the previous sections. In Table 1 the rms (M_2)

Table 1

Model	Misfit to Arch. Int	Misfit to pre-1840 Historical	Global misfit	No. Accepted Data pre-1840	Total No. Accepted Data	S_N	T_N
gufm1	0.839	1.551	1.156	113,564	365,694	3.51E+13	6.18E+04
gufm1-g10c	0.835	1.536	1.153	114,008	367,323	3.28E+13	5.93E+04
gufm1-g10Wf2	1.058	1.493	1.136	114,624	367,937	2.98E+13	5.88E+04
gufm1-g10Wf15	0.693	1.545	1.156	113,565	366,877	3.44E+13	6.52E+04
gufm1-g10Wf40	0.625	1.575	1.164	112,421	365,736	4.63E+13	1.17E+05

M_2 misfit between the model predictions and the archaeointensity dataset, the historical datasets spanning 1590–1840 A.D. and all datasets as well as the number of data accepted pre-1840 A.D. and the total number of accepted data. Note that *gufm1* was constructed without archaeointensity data. The final two columns document the global spatial and temporal norms: S_N denotes the Ohmic heating norm (4) in units of $(\mathrm{nT})^{-2}$ and T_N denotes the temporal curvature norm (6) in units of $(\mathrm{nT})^{-2}(\mathrm{year})^{-4}$.

misfit of these models to the data involved in the inversions is presented. The misfit of the models to the archaeointensity dataset, to historical data from 1590 A.D. to 1840 A.D. and to all the data used in the inversions are reported. The number of data accepted between 1590 A.D. and 1840 A.D., and the total number of accepted data are also reported. Again, note that for the over-weighted models, data greater than three times the error estimate from the previous model iterate are rejected; models with linearly constrained $g_1^0(t)$ evolution have data rejection for the historical datasets only. The global spatial and temporal norms for the models are also presented in Table 1.

gufm1-g10c fits the archaeointensity data and pre-1840 A.D. historical data slightly better than *gufm1* and less data is rejected during its construction. The free evolution model *gufm1-g10Wf2* fits the archaeointensity data much worse than *gufm1-g10c* and *gufm1*. *gufm1-g10Wf40* fits the archaeointensity data best of all the models considered, but does not fit the pre-1840 A.D. historical data as well as *gufm1-g10c* probably because incorrect variations in field amplitudes have been introduced in the attempt to fit archaeointensity data which probably contain significant errors. *gufm1-g10Wf15* fits the pre-1840 A.D. historical data better than *gufm1* (but not as well as *gufm1-g10c*) and fits the archaeointensity data much better than either *gufm1* or *gufm1-g10c* which are limited by their linearly constrained $g_1^0(t)$. Once more, the conclusion is that *gufm1-g10Wf15* is the most reasonable of the free evolution models tested in detail, though it fits the pre-1840 A.D. historical data no better that the linearly constrained models. Examining the fit of the various field models to the data used in the inversions thus suggests that simple axial dipole evolution models, with linear constraints on $g_1^0(t)$, are capable of accounting for both archaeomagnetic and historical data within their estimated errors.

4.4. Comparison with the independent dataset of Gallet et al. (2005)

In this section, the fit of the field models to an independent dataset not included in the inversions is reported. The relative success of models in fitting an independent dataset is a classic criterion for model comparison; in the present case this is a difficult task because there are very few archaeointensity data spanning the interval 1590–1840 A.D. and almost all the reliable data have already been included in the dataset used in the inversion. Fortunately, a new, high quality, dataset described by Gallet et al. (2005) and not included in the compilation of Korte et al. (2005) is available. This was acquired using a particularly strict archaeointensity protocol (Genevey and Gallet, 2002) and the authors helpfully provided error estimates. Unfortunately, it consists of only a very small number of data and the observations are all located in France. Despite these limitations, this dataset can give a very useful indication of how good a job the various field models are doing at capturing intensity variations pre-1840 A.D. in Europe, where they might be expected to perform best. Table 2 reports the M_2 misfit of the models to the Gallet et al. (2005) dataset; the predicted evolution of intensity in Paris and the measured intensities relocated to Paris (with error bars) are plotted in Fig. 7. Note that the misfits reported in Table 2 are calculated at the original data site locations; relocation does not contribute to these misfits.

Table 2
M_2 misfit (see (10) for definition) of historical field models to the independent Gallet et al. (2005) dataset, not included in the inversions

Model	M_2 misfit to Gallet et al. (2005)
gufm1	1.926
gufm1-g10c	1.914
gufm1-g10Wf2	3.351
gufm1-g10Wf15	2.618
gufm1-g10Wf40	3.865

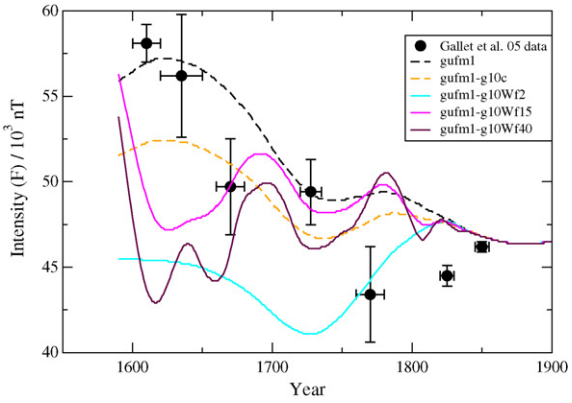

Fig. 7. Independent dataset of Gallet et al. (2005) (black dots) with the author's error bars added. Dashed lines are the field models with a linear $\dot{g}_1^0(t)$ imposed, solid lines are models without linear constraint on $g_1^0(t)$, but with different weightings of the archaeointensity data. Note that for the purposes of comparison, the data of Gallet et al. (2005) has been relocated to Paris and that the field model predictions for Paris are presented here. The misfits reported in Table 2 refer to data and model predictions at the original data locations, before relocation.

It is apparent that none of the historical field models presented in the previous section satisfactorily explain all the Gallet et al. (2005) archaeointensity data to within their suggested error estimates. The models *gufm1*, *gufm1-g10c* and *gufm1-g10Wf15* do fairly well, while *gufm1-g10Wf2* and *gufm1-g10Wf40* perform poorly. Evidently, forcing field models to fit the Korte et al. (2005) archaeointensity dataset very tightly does not help them to explain this independent intensity dataset. This leads to the conclusion the presently available archaeointensity data are insufficient to produce reliable global models of rapid intensity variations during the interval 1590–1840 A.D.

This (admittedly limited) test also indicates that the best free evolution model *gufm1-g10Wf15* does worse at predicting independent data than simpler models with $g_1^0(t)$ linearly constrained. *gufm1-g10c* does marginally better than *gufm1*, though there is not a large difference. On the basis of this test involving independent data, *gufm1-g10c* should therefore be the favoured model. I note however that this test is not as rigorous as one might wish due to the small and geographically limited nature of the Gallet et al. (2005) dataset. It would be worthwhile to carry out more detailed tests in the future when further independent data become available.

The inability of present historical field models to fit data showing rapid intensity variations in the 17th, 18th and 19th century is concerning, especially since these variations have been heavily interpreted with regard to possible archaeomagnetic jerks and potential mechanisms of climate change (Gallet et al., 2005; Courtillot et al., 2007a). None of the models reported here capture the 15 μT amplitude change in intensity suggested by the Gallet et al. (2005) data during the interval 1600–1800 A.D. Two explanations for this discrepancy seem likely. First, the error estimates of Gallet et al. (2005) for their intensity measurements (based on the scatter of multiple samples from a single location and time) are likely over-optimistic since sources of systematic error are ignored. Secondly, the historical field models are perhaps over-smoothed in time in an effort to deal with the sparse data, thus are unable to fit very rapid local changes in field intensity. It is nonetheless interesting to note that the model *gufm1-g10c* predicts a trend of decreasing intensity in Paris between 1600 and 1800, with an oscil-

lation about this trend leading to higher intensities around 1630 and lower intensities around 1750. This is indeed the general pattern suggested by the Gallet et al. (2005) data.

4.5. Bayesian model comparison

Thus far, the conventional frequentist approach to model comparison has been adopted; this involves determining which model best explains the data in a maximum likelihood sense and preferring it. However, the field models presented involve different parameterizations and different weightings of the data; these differences as well as prior beliefs concerning what constitutes a physically reasonable model should also be taken into account when models are quantitatively compared. Here, I propose that the technique of Bayesian model comparison (Jeffreys, 1961; Jaynes, 2003; MacKay, 2003; Sivia and Skilling, 2006; Trotta, 2007) is a useful framework for carrying out such comparisons. The advantages of this approach include that it naturally takes into account the complexity of models (in terms of the size of the model space) with assumptions concerning what is physically reasonable being encoded through prior probability density functions. To my knowledge this technique has not been used before in the context of comparing geomagnetic field models; I therefore illustrate the method by applying it to a very simple problem: comparing models of only $g_1^0(t)$. The language and notation of Sivia and Skilling (2006) is adopted in this section; their chapter 4 should be consulted by those readers desiring a more detailed exposition.

From a Bayesian perspective, the relative merit of two competing models **A** and **B** given a particular dataset **D** and an experimental setup I, can be compared by computing the (odds) ratio of their posterior probabilities

$$\frac{p(\mathbf{A}|\mathbf{D},I)}{p(\mathbf{B}|\mathbf{D},I)} = \frac{p(\mathbf{D}|\mathbf{A},I)\,p(\mathbf{A}|I)}{p(\mathbf{D}|\mathbf{B},I)\,p(\mathbf{B}|I)} \tag{12}$$

where $p(\mathbf{A}|\mathbf{D},I)$ and $p(\mathbf{B}|\mathbf{D},I)$ are the conditional posterior probabilities of each model given the data **D**, $p(\mathbf{D}|\mathbf{A},I)$ and $p(\mathbf{D}|\mathbf{B},I)$ are the probabilities of the data given the model, and $p(\mathbf{A}|I)$ and $p(\mathbf{B}|I)$ are the prior probabilities of the models before the arrival of the data, derived from previous and independent knowledge. Eq. (12) arises from the simple application of Bayes' theorem to the posterior probabilities. It has proved extremely useful in a wide variety of model comparison problems in the physical sciences (see, for example, Gregory, 2005).

In the present simple example, **D** is the archaeointensity dataset of Korte et al. (2005) spanning 1590–1840 A.D. as described in Section 2, while the models to be compared (denoted **A** and **B** above) are examples of classes of field models with particular assumed parameterizations for $g_1^0(t)$ between 1590 A.D. and 1840 A.D. In principle, one could quantitatively compare any of the previously constructed models using the apparatus of Bayesian model comparison. In practice however, it would be a major undertaking to deal correctly with the effects of temporal regularization and variable weighting factors W_f in the free evolution models for $g_1^0(t)$: such nuisance parameters should really be integrated over in the Bayesian approach as discussed by MacKay (1992). In the interests of a simple demonstration of the method, I accept the findings of Sections 4.3 and 4.4 that the linearly varying models are preferable to the free evolution models, and focus instead on the question raised at the end of Section 4.1. There it was asked whether a model with no change in $g_1^0(t)$ between 1590 A.D. and 1840 A.D. (such as *gufm1-g10c*) should be preferred to the maximum likelihood model with a linearly varying g_1^0 (t). To answer this question, three models are considered.

(i) $\mathbf{m_1}$: $g_1^0(t)$ continues at its average post-1840 A.D. rate of decay (15 nT/year as assumed in *gufm1*):

$$g_1^0(t) = g_1^0(1840) + 15.0(t - 1840) \quad 1590 \le t \le 1840 \tag{13}$$

(ii) $\mathbf{m_2}$: $g_1^0(t)$ remains unchanged from 1840 A.D. to 1590 A.D. (referred to as *gufm1-g10c* in previous sections):

$$g_1^0(t) = g_1^0(1840) \quad 1590 \le t \le 1840 \tag{14}$$

(iii) $\mathbf{m_3}$: $g_1^0(t)$ can have any linear rate of change pre-1840 A.D. as explored in Section 4.1:

$$g_1^0(t) = g_1^0(1840) + \beta(t - 1840) \quad 1590 \le t \le 1840 \tag{15}$$

Models $\mathbf{m_1}$ and $\mathbf{m_2}$ have no free parameters, while $\mathbf{m_3}$ is more complex since it has one free parameter β. It is further assumed that there is no reason to favour of any of these models in advance, so that $p(\mathbf{m_1}|I) = p(\mathbf{m_2}|\mathrm{I}) = p(\mathbf{m_3}|I)$. The posterior odds ratio needed to compare of any two of the models (for example, **A** and **B** which could be any two of $\mathbf{m_1}$, $\mathbf{m_2}$ or $\mathbf{m_3}$) in this case reduces to

$$\frac{p(\mathbf{A}|\mathbf{D},I)}{p(\mathbf{B}|\mathbf{D},I)} = \frac{p(\mathbf{D}|\mathbf{A},I)}{p(\mathbf{D}|\mathbf{B},I)}. \tag{16}$$

To calculate quantities such as $p(\mathbf{D}|\mathbf{A},I)$, often known as the evidence or marginal likelihood for a model **A**, see Sambridge et al. (2006), one must integrate over free parameters in the model **A**; doing this enables one to compare the complete ability of all models of the form **A** to account for the data **D**.

For models $\mathbf{m_1}$ and $\mathbf{m_2}$ the calculation of the evidence is very straightforward because there are no free parameters to be integrated over; the likelihood probability density function for particular choices of the linear constraint (β) must simply be evaluated from Fig. 3. This is permissible because $\mathbf{m_1}$ and $\mathbf{m_2}$ are formally equivalent to having Dirac delta functions at particular β (at $\beta = 15.0$ nT/year for $\mathbf{m_1}$ and at $\beta = 0.0$ nT/year for $\mathbf{m_2}$) as prior probability density functions for β.

For model $\mathbf{m_3}$ things are a little more complicated. To obtain the evidence, the posterior probability density function must be integrated over the free parameter β

$$p(\mathbf{D}|\mathbf{m_3},I) = \int p(\mathbf{D}|\mathbf{m_3},\beta,I)p(\beta|\mathbf{m_3},I)\,\mathrm{d}\beta \tag{17}$$

where $p(\mathbf{D}|\mathbf{m_3},\beta,I)$ is the likelihood probability density function of obtaining the observed data, given the model $\mathbf{m_3}$ and a particular choice of β; this is what was previously presented in Fig. 3. A prior probability density function for β, written ($p(\beta|\mathbf{m_3},I)$), is also required in order to complete the calculation. In choosing the prior probability density function, one must remember that only knowledge available before new data (the archaeointensity data in this case) arrives may be take into account. For example, it is not legitimate to choose a prior probability density function with mean value matching the maximum likelihood value of 2.74 nT/year obtained after analysis of the archaeointensity dataset as in Section 4.1, or to test many such priors and choose (after the fact) that which one performs best.

Respecting such considerations, I choose the prior probability density function for β to be a Gaussian distribution with a mean of 15 nT/year. 15 nT/year is the average value of β from 1840 A.D. to 1990 A.D., which is legitimate prior information regarding models for the pre-1840 axial dipole evolution because it comes from an independent post-1840 dataset. Choosing the width of the Gaussian prior is more troublesome. I therefore decided to investigate three possible prior widths of 20 nT/year, 100 nT/year and 500 nT/year. The 20 nT/year width expresses a prior belief that the β will likely be very similar to its post-1840 A.D. value, the 100 nT/year width encompasses a wide range of physically (and palaeomagnetically)

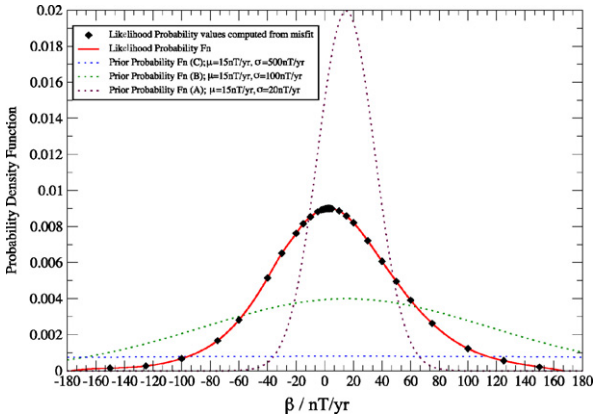

Fig. 8. Likelihood probability density function (solid line), and three Gaussian priors with widths 500 nT/year, 100 nT/year and 20 nT/year; these are used in the computation of the posterior probability density functions for model $\mathbf{m_3}$.

plausible values for β, while the width 500 nT/year quantifies a state of relative ignorance regarding the likely value β. With these three choices I aim to span a wide range of possible prior beliefs concerning the value of β; these prior beliefs play an important role in Bayesian inference, especially when the available data are not as informative as one would wish. The three prior probability density functions are shown together with the likelihood probability density function for β in Fig. 8. Note that assumptions regarding the prior are here stated explicitly: this is the hallmark of a Bayesian analysis.

Calculating the posterior odds ratios in the manner described above yields the results presented in Table 3 for the ratio of the posterior probability of $\mathbf{m_3}$ compared to that for $\mathbf{m_1}$ and $\mathbf{m_2}$. The important point is that these posterior odds ratios are always less than 1 meaning that model $\mathbf{m_3}$ is always less probable than $\mathbf{m_1}$ and $\mathbf{m_2}$. This result holds for all reasonable choices of prior probability density function. It is observed that the higher the degree of prior ignorance in β (i.e. the larger of the width of the prior probability density function), the less probable model $\mathbf{m_3}$ becomes relative to $\mathbf{m_1}$ and $\mathbf{m_2}$. This result demonstrates how Bayesian model comparison naturally penalizes more complex models: the larger the model space, the easier it is for a model to explain any particular set of observations, and consequently the less faith one has in it generalizing to other scenarios. Mathematically this is incorporated in calculation by the integral carried out over the free parameter β.

Comparisons between the normalized posterior and prior probability density functions for the three choices of prior explored for $\mathbf{m_3}$ are presented in Fig. 9. This illustrates how the information supplied by the data has updated and improved the prior probability density functions for β. Notice that in all cases the maximum of the posterior probability density functions is shifted towards lower values of β than the assumed prior mean of 15 nT/year. This plot firmly demonstrates that the data favour a value of β lower than 15 nT/year, as is also seen in the likelihood probability density function plotted in Figs. 3 and 8.

Table 3

Model	$\mathbf{m_3}$-20	$\mathbf{m_3}$-100	$\mathbf{m_3}$-500
$\mathbf{m_1}$	0.92	0.42	0.093
$\mathbf{m_2}$	0.87	0.40	0.088

Posterior odds ratios for model $\mathbf{m_3}$ compared to models $\mathbf{m_1}$ and $\mathbf{m_2}$. The ratios are always less than 1 meaning that models $\mathbf{m_1}$ and $\mathbf{m_2}$ are always more probable than model $\mathbf{m_3}$, regardless of the choice of prior.

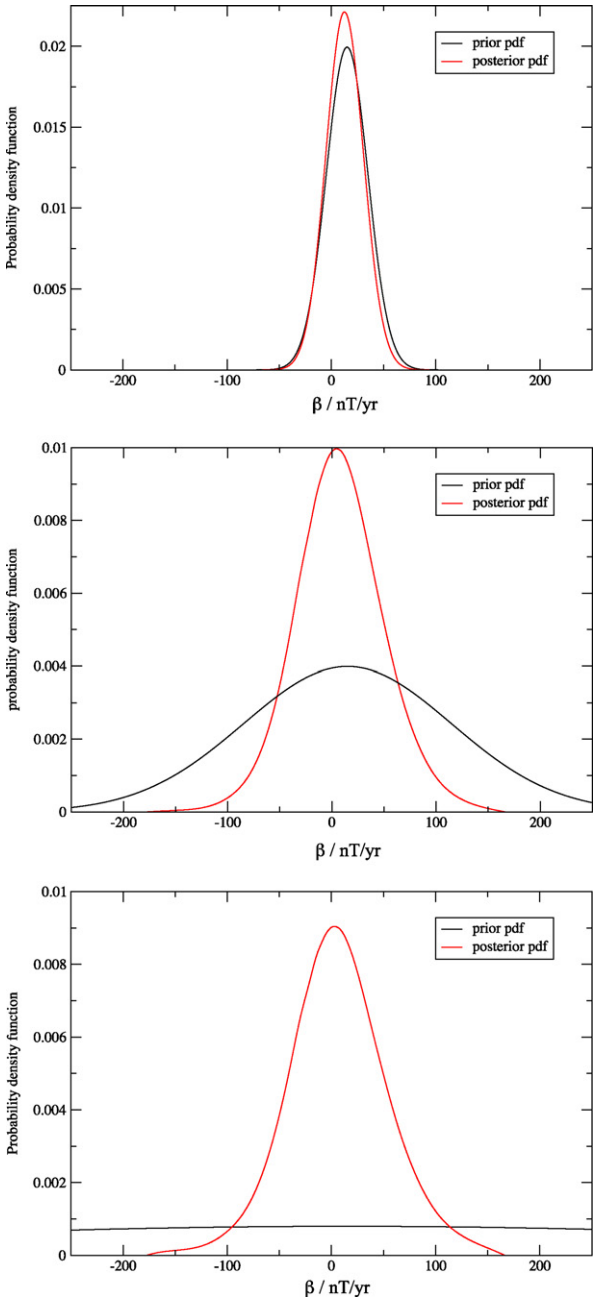

Fig. 9. Prior and posterior probability density functions for model $\mathbf{m_3}$ showing how observers knowledge is updated after the arrival of the data. In all cases the maximum of the posterior probability density function is shifted to lower values than the 15 nT/year mean assumed in the prior probability density functions.

Finally, it is worth pointing out that the ratio of posterior probabilities of models $\mathbf{m_1}$ and $\mathbf{m_2}$ is 0.95, meaning that model $\mathbf{m_2}$ with constant $g_1^0(t)$ is more probable than model $\mathbf{m_1}$ which involved the assumption that $\beta = 15$ nT/year as in *gufm1*. This result is independent of any choice of prior since neither $\mathbf{m_1}$ nor $\mathbf{m_2}$ has any free parameters to be integrated over. This result could in fact already have been deduced from the likelihood probability density function of Fig. 3. Obviously, if one had proposed a model with the constant rate of change $\beta = 2.74$ nT/year (the maximum likelihood value of β) this model would be more probable than $\mathbf{m_2}$. However, it is hard to envisage how this value of β could have been proposed before the arrival of the data, while $\beta = 15$ nT/year and 0 nT/year (the assumptions of models $\mathbf{m_1}$ and $\mathbf{m_2}$) represent conceivable prior

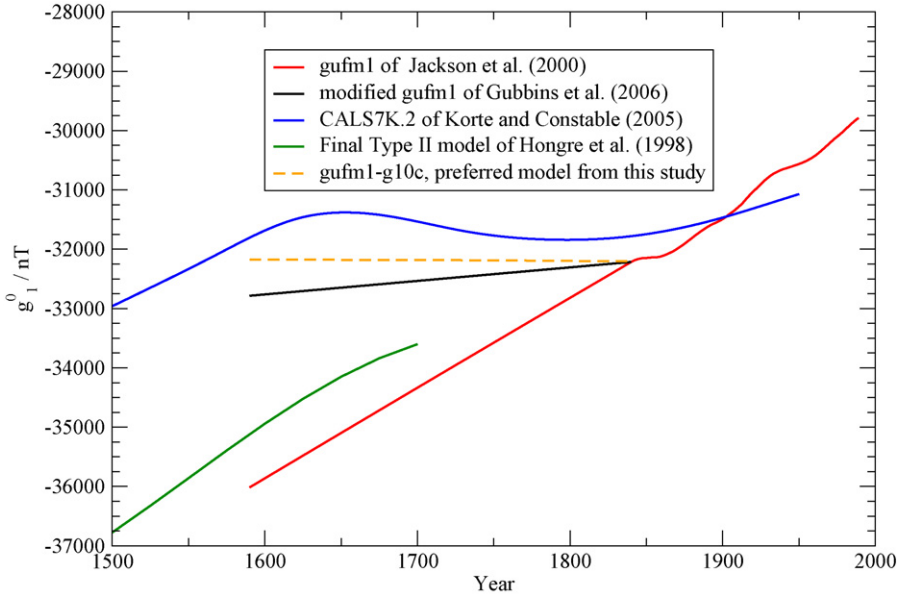

Fig. 10. Summary plot of previous inferences of $g_1^0(t)$ evolution together with the preferred model *gufm1-g10c* from the present study.

assumptions given previous knowledge and simplicity principles. For these reasons, I conclude that model $\mathbf{m_2}$ as incorporated in the field model *gufm1-g10c* and described in Figs. 4–7 and Tables 1 and 2 is the most probable model for $g_1^0(t)$ given the Korte et al. (2005) archaeointensity dataset.

5. Discussion

5.1. Comparison to previous models of $g_1^0(t)$ 1590–1840 A.D.

The most probable model identified in the previous sections (*gufm1-g10c*) involves no change of $g_1^0(t)$ between 1590 A.D. and 1840 A.D. In Fig. 10 this model is plotted together with previously published models discussed in the introduction, in order to aid comparisons.

gufm1-g10c is consistent with the result of $\beta = 2.28 \pm 2.72$ nT/year obtained by Gubbins et al. (2006), who analyzed the same archaeointensity dataset using a different method. The minor difference between their result and the findings of this study arise because I favour a simpler model with $\beta = 0$ nT/year on the basis of the Bayesian model comparison analysis of Section 4.5. It is important to emphasize that both the result of this study and that of Gubbins et al. (2006) are better compatible with the shallowing of $\dot{g}_1^0$ found in *gufm1* between 1860 A.D. and 1840 A.D. than is the $\beta = 15$ nT/year assumption employed by Jackson et al. (2000).

The lack of trend in $g_1^0(t)$ from 1590 A.D. to 1840 A.D. suggested by *gufm1-g10c* is also in agreement with the axial dipole evolution found during this time in the CALS7K.2 archaeomagnetic field model of Korte and Constable (2005a). In contrast, axial dipole field decay over the interval 1590–1750 A.D. was inferred by Hongre et al. (1998). However, this model ran only to 1750 A.D., relied on a rather limited dataset in its last few hundred years, and likely has endpoint problems during the 17th and 18th centuries.

5.2. Prospects for future improvements in models of $g_1^0(t)$

The constant pre-1840 $g_1^0(t)$ present in *gufm1-g10c* is obviously a gross simplification of the true evolution. This result should be interpreted as indicating that the average rate of change of $g_1^0(t)$ 1590–1840 A.D. is zero, and not mistaken as evidence for a truly constant $g_1^0(t)$ from 1590 A.D. to 1840 A.D. Recourse to inferences concerning the an average rate of change over a 250-year interval is unfortunately all that appears possible given the presently available data. Of course, the conclusions drawn in this study are dependent on the archaeointensity database of Korte et al. (2005) and the limitations of the data it includes. Is there any prospect of eventually being able to do better in retrieving more detail concerning $g_1^0(t)$, and how might this be achieved? The present study certainly suggests that there is room for improvement. I would like to emphasize two issues on which progress is required in order for improved models of the historical evolution of the axial dipole field to be possible.

Firstly and most obviously, further major efforts are required towards the collection of more accurate and better dated archaeointensity measurements. Good geographical coverage is especially important during the 16th and 17th centuries because the geographical coverage of historical directional data is less comprehensive during this time interval (Korte and Constable, 2006b; Hulot et al., 1997). This task is undoubtedly a major undertaking, but presents a chance to foster close collaborations between the geomagnetic and palaeomagnetic communities in the coming decades towards achieving a common goal. The prospect of new methods for extracting accurate absolute palaeointensity measurements (e.g. Genevey and Gallet, 2002; Le Goff and Gallet, 2004), exploitation of alternative recorders of magnetization (e.g. bacterial magnetosomes could be useful single domain recorders in lake sediments Snowball, 1994; Snowball and Sandgren, 2004; Snowball et al., 2007) and improvements in modelling of currently under-exploited relative palaeointensity records (Korte and Constable, 2006b) all give hope for the future. As a specific point, archaeointensity measurements from the 15th and 16th century would be especially valuable for testing the validity of *gufm1-g10c*.

Secondly, further efforts are needed towards better characterization of the errors inherent in the archaeointensity data, and how these can best be taken into account during global field modelling. Over-weighting the archaeointensity data relative to the estimates of Korte et al. (2005), it was found in this study that field models with physically plausible Ohmic heating and temporal curvature norms could be obtained. This indicates that the error estimates of

Korte et al. (2005) for archaeointensities may be slightly too conservative and uncritically adopting these estimates may prevent the full information content of the data from being exploited.

5.3. Implications for core flow inversions pre-1840 A.D.

Core flow inversions are routinely conducted using historical field models as input (see for example, Jackson, 1997; Amit and Olsen, 2006), but they are usually limited to the post-1840 A.D. epoch. A major reason for this limitation is the ambiguity introduced by the unjustified assumption regarding $g_1^0(t)$ pre-1840 A.D. made in the *ufm* and *gufm1* field models (Jackson, 2000). Although a step forward has been taken in this study by constructing a model *gufm1-g10c* that involves a simple $g_1^0(t)$ history constrained by archaeointensity data, it is far from capturing the complete details of the true $g_1^0(t)$ evolution produced by core processes. It therefore remains inadvisable to undertake core flow inversions before 1840 A.D., even using improved historical field models such as *gufm1-g10c*.

5.4. Centennial timescale variations in $g_1^0(t)$

Detailed examination of $\dot{g}_1^0(t)$ post-1840 A.D. (Gubbins, 1987; Bloxham and Jackson, 1992), has shown decadal variations occurring around a longer term trend. The important finding of Gubbins et al. (2006) was that this longer term trend changed from rapid decay to being almost constant between the intervals 1590–1840 A.D. and 1840–1990 A.D. (centre of intervals separated by 200 years). This evidence suggested major and rapid centennial variations in $g_1^0(t)$ and it is supported by the results of the present study as embodied by the field model *gufm1-g10c*.

When did the change in the centennial trend of $g_1^0(t)$ occur and how rapidly? Considering a quantity averaged over two successive epochs, if a difference is observed, it is tempting to conclude the change happened rapidly at the boundary between the epochs. This would of course be an incorrect inference. Gubbins et al. (2006) deduced that the change in the centennial trend in $g_1^0(t)$ occurred rapidly (over several decades) close to 1840 A.D. not on this basis, but because they noted the clear shallowing of $g_1^0(t)$ to an almost constant level between 1860 A.D. and 1840 A.D. This shallowing coincided with concomitant changes in the growth rate of reversed flux features in the Southern hemisphere at the core surface, which fitted with a known physical mechanism involving episodic flux expulsion. The fact that these changes occurred close to the boundary between the analysis intervals at 1840 A.D. was in their view simply an unfortunate coincidence.

From the perspective of the present study, the new field model *gufm1-g10c* can be interpreted in a very similar way: it contains the same shallowing $g_1^0(t)$ close to 1840 A.D., and its reversed flux features evolve in a similar manner to that reported in Gubbins et al. (2006) for their model with $\beta = 2.28$ nT/year. It is remarkable that the morphology of reversed and normal flux features (but not their amplitudes) are similar for all the models of $g_1^0(t)$ investigated in Section 4 as illustrated by Fig. 6.

Can anything further be concluded about variations in $g_1^0(t)$ from free evolution models such as *gufm1-g10Wf15*? Unfortunately not: the oscillations found pre-1840 A.D. in these models clearly depend on how strongly the archaeointensity data are weighted and it would therefore be foolhardy to interpret them, no matter how tempting this may be. At present the best one can do when studying centennial variations in $g_1^0(t)$ is to consider average rates of change over several hundred years (this study) or alternatively to look at archaeomagnetic field models such as CALS7K.2 (Korte and Constable, 2005a) which involve implicit smoothing kernels with widths of several hundred years (Korte and Constable, 2008).

5.5. Physical origin of centennial timescale variations in $g_1^0(t)$

Detailed investigation of the physical mechanism underlying rapid axial dipole change is beyond the scope of the present study, but some remarks seem appropriate. It is surprising that the mechanism producing temporal variations of a first order geophysical observable (the axial dipole field) is still very poorly understood, especially given the important consequences for solar-terrestrial interactions and space technologies. One useful contemporary paradigm within which to think about the problem is that of Zhang and Gubbins (2000): they see the geodynamo process as being intrinsically unstable due to the sensitivity of dynamo action to small changes in core flow, while at the same time small changes in magnetic field structure can cause major alterations in the form of the convection-driven, rapidly rotating flow. They envisaged these feedbacks as leading to fluctuations occurring across a wide range of timescales, most prominently as excursions and reversals but also accounting for shorter term, smaller amplitude variations such as the centennial variations of $g_1^0(t)$ discussed in the previous section. Their suggestion certainly seems useful, but lacks a detailed mechanistic account of exactly how such variations might come about. Fortunately, some more precise proposals of how variations in $g_1^0(t)$ could occur have recently been made on the basis of examinations of numerical geodynamo simulations.

Episodic expulsion of toroidal flux by convective upwelling was suggested by Gubbins et al. (2006) as an explanation for their observation of erratic variations in the rate of change of the axial dipole field during the past 400 years. Aubert et al. (2008) have shown how this might occur using novel visualization techniques to analyze numerical geodynamo simulations: they were able to follow magnetic upwellings (field lines with large magnetic energy) from their birth place at the inner-core boundary to their expulsion at the core-mantle boundary. They found this to be a highly variable process, with the evolution in any given case dependent on the precise flow and field configurations present. Another possibility has been highlighted by Busse and Simitev (2006). They describe how dynamo waves (Parker, 1955) and convective relaxation oscillations (Busse, 2002) can interact to produce time-variations in the poloidal field; it is however unclear whether this mechanism can produce variations on the centennial timescales of interest. Finally, Rotvig (2007) has proposed another mechanism on the basis of numerical experiments of rotating convection at lower (though still far from Earth-like) Ekman number of order 10^{-6} and at highly supercritical Rayleigh numbers. He finds convectively driven flow patterns involving strong and latitudinally drifting zonal winds that can act as kinematic dynamos and exhibit time variations in the poloidal field energy as the zonal jets (ω-effect) drift through the region where columnar convection (α-effect) occurs. The timing of field oscillations in this scenario depends on the occurrence of bursts of zonal flow that could apparently operate on timescales as short as centuries.

At the moment all the above mechanisms remain rather speculative, primarily because it is not yet clear whether the geodynamo is operating via an α^2 mechanism (like most current numerical models) or via an $\alpha\omega$ mechanism which seems plausible at low Ekman numbers where strong zonal flows could be favoured. Perhaps some input from observations could be helpful in testing the plausibility of these theoretical alternatives? If archaeointensity observations and field modelling methods improve sufficiently in the future, and if the prediction of the theoretical mechanisms can be suitably parameterized, it should be possible to use the apparatus of Bayesian model comparison to carry out such tests. Unfortunately the prospect of such tests seems remote at present; as illustrated in this study, all that can be determined with current data is the linear rate of change of $\dot{g}_1^0$ over a 250 year interval—it is impos-

sible to distinguish between the above mechanisms on the basis of such limited information. A pragmatic alternative to quantitative hypothesis testing is a qualitatively comparison of patterns of field evolution at the core surface with the predictions of different mechanisms found in numerical models. Such comparisons appear to favour the flux expulsion mechanism proposed by Gubbins et al. (2006).

5.6. Archaeomagnetic jerks since the 16th century and implications for the hypothesis of geomagnetic control of climate change

What can the models derived in Section 4 say concerning hypothesis of Gallet et al. (2005), that increases in geomagnetic field intensity observed at mid-latitudes on centennial timescales (known archaeomagnetic jerks, though not to be confused with geomagnetic jerks) are caused by a tilt of the geomagnetic dipole to lower latitudes? Gallet et al. (2005) used archaeomagnetic measurements from France to suggest that such events occurred around 1600 A.D. and 1800 A.D. These events lie within the timespan of the models constructed here, so I can test whether or not large movements of the geomagnetic dipole to lower latitudes occur in historical field models at these times.

In Fig. 11 the evolution of the latitudinal position of the northern geomagnetic dipole for the historical field models *gufm1*, *gufm1-g10c*, *gufm1-g10Wf2*, *gufm1-g10Wf15* and *gufm1-g10Wf40* as well as the archaeomagnetic field model CALS7K.2 are presented. All historical field models display the same trend: the dipole latitude gradually moves to slightly lower latitudes (from around 85° to around 78°) between 1600 A.D. and 1800 A.D.; since then it has stayed approximately constant, with only minor variations in the last 200 years. This gradual trend is in agreement with the monotonic decrease in dipole latitude between 1600 A.D. and 1800 A.D. found in CALS7K.2 (see Fig. 10 of Korte and Constable, 2008).

It is noteworthy that neither *gufm1* nor *gufm1-g10c* (with their $g_1^0(t)$ constrained to evolve linearly at rather different rates) are capable of producing the amplitude of dipole tilt required by the Gallet et al. (2005) hypothesis. Since $g_1^1(t)$ and $h_1^1(t)$ are well constrained by historical directional data, the only possibility for producing large dipole tilts unconstrained by historical data is through rapid, large amplitude, variations in $g_1^0(t)$. However, as shown in Fig. 11, even field models such as *gufm1-g10Wf15* and *gufm1-g10Wf40* that contain rapid oscillations in $g_1^0(t)$ (see Fig. 4), and also fit satisfactorily the observed historical and archaeomagnetic data, are again not capable of producing the required large dipole tilts. Instead they are only capable of producing small amplitude dipole latitude oscillations around the trend which is found in the linear evolution models.

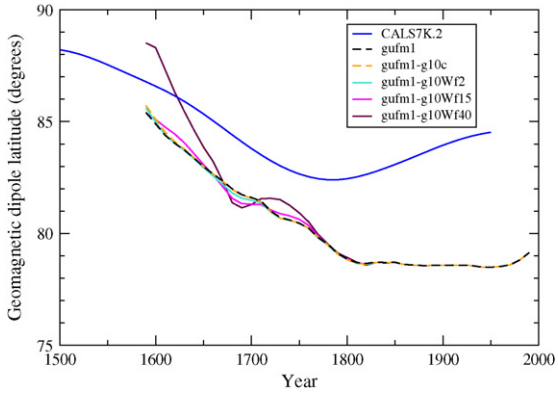

Fig. 11. Latitude of northern geomagnetic dipole for *gufm1*, *gufm1-g10c*, *gufm1-g10Wf2*, *gufm1-g10Wf15*, *gufm1-g10Wf40* and CALS7K.2.

It should be acknowledged that, as shown in Section 4.4, none of the field models investigated here satisfactorily fit the Gallet et al. (2005) intensity data to within the error estimates suggested by the authors. To confirm whether sharp changes in field intensity occurred over a wide geographical region in 1600 A.D. and 1800 A.D., further high accuracy, well dated, archaeointensity observations from the 16th to 18th centuries in Europe are needed. It is possible that current historical field models are too smooth in time (due to the poor time resolution of the data) and therefore incapable of capturing very rapid intensity variations. However, another possibility is that the error estimates of Gallet et al. (2005) are too optimistic. The observations in the archaeointensity database of Korte et al. (2005) appear most consistent with the later possibility; this is reflected in the field models presented here.

In summary, I find no evidence from time-dependent geomagnetic field models, constructed from global databases of historical and archaeomagnetic data spanning the past 400 years, for the rapid large amplitude motions of the geomagnetic dipole towards low latitudes in 1600 A.D. and 1800 A.D. required by the tilting dipole hypothesis of Gallet et al. (2005). Such large dipole motions would created a clear global signature in historical and archaeointensity observations that is not evident. The dipole tilt mechanism for explaining archaeomagnetic jerks is the keystone in the story of geomagnetic control of climate change promoted by Courtillot et al. (2007a,b). Since there is no convincing evidence for large dipole tilts having occurred in the last 400 years, their proposed explanation for cooling events around 1600A.D and 1800 A.D. (spanning the little ice age) currently lacks observational support from geomagnetism. The findings of this study therefore add a further angle to the criticisms already made by Bard and Delaygue (2007), who argued on other grounds that the case for geomagnetic control of climate has been over-stated.

6. Conclusions

The investigations carried our here, using both historical and archaeointensity observations, indicate that between 1590 A.D. and 1840 A.D. no change in $g_1^0(t)$ is the most probable model for the evolution of the geomagnetic axial dipole. This conclusion is consistent with a decrease in $\dot{g}_1^0(t)$ observed between 1860 A.D. and 1840 A.D. in *gufm1*, with the value of $\dot{g}_1^0(t) = 2.28 \pm 2.72$ nT/year obtained in the earlier study of Gubbins et al. (2006) by an alternative method, and with the CALS7K.2 archaeomagnetic field model of Korte and Constable (2005a). The large errors, both in dating and measurement, inherent in the presently available archaeointensity data dictate that only very simple models of $g_1^0(t)$ are defensible between 1590 A.D. and 1840 A.D.; consequently it is not yet possible to reliably determine details of rapid time variations of $g_1^0(t)$ before 1840 A.D. In future work it is planned to include the simple working hypothesis of an unchanged axial dipole between 1590 A.D. and 1840 A.D. as a constraint in the next generation of historical geomagnetic field models that will extend into the 21st century.

Acknowledgments

This work was first presented during a meeting organized by Keke Zhang held at the University of Exeter to celebrate David Gubbins' 60th birthday in May 2007. I wish to thank David for stimulating my interest in this problem while I was a Ph.D. student at the University of Leeds. Andy Jackson provided helpful advice regarding Bayesian methods. Two anonymous reviewers are thanked for a careful reading that helped to improve the manuscript. This

research was initiated while the author held a postdoctoral position at the University of Leeds funded by the NERC GEOSPACE program (grant NER/O/S/2003/00674).

References

Amit, H., Olsen, P., 2006. Time-average and time-dependent parts of core flow. Phys. Earth Plan. Inter. 155, 120–139.

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

Bard, E., Delaygue, M., 2007. Comment on "Are there connections between Earth's magnetic field and climate? by Courtillot, V., Gallet, Y., LeMouël, J.-L., Fluteau, F., and Genevey, A. Earth Plan. Sci. Lett. 253, 328–339, 2007". Earth Plan. Sci. Lett. 265, 302–307.

Barraclough, D.R., 1974. Spherical harmonic analyses of the geomagnetic field in eight epochs between 1600 and 1910. Geophys. J. R. Astron. Soc. 36, 497–513.

Bloxham, J., Gubbins, D., Jackson, A., 1989. Geomagnetic secular variation. Philos. Trans. R. Soc. Lond. A 329, 415–502.

Bloxham, J., Jackson, A., 1989. Simultaneous stochastic inversion for geomagnetic main field and secular variation 2. 1820–1980. J. Geophys. Res. 97, 19537–19563.

Bloxham, J., Jackson, A., 1992. A time-dependent mapping of the magnetic field at the core-mantle boundary. J. Geophys. Res. 94, 15753–15769.

Busse, F.H., Simitev, R.D., 2006. Parameter dependences of convection-driven dynamos in rotating spherical fluid shells. Geophys. Astrophys. Fluid. Dyn. 100, 136–341.

Busse, F.H., 2002. Convective flows in rapidly rotating spheres and their dynamo action. Phys. Fluids 14, 1301–1314.

Constable, C., 2007. Centennial-to millennial-scale geomagnetic field variations. In: Kono, M. (Ed.), Treatise on Geophysics, Elsevier, vol. 5, pp. 337–372.

Courtillot, V., Gallet, Y., LeMouël, J.-L., Fluteau, F., Genevey, A., 2007a. Are there connections between Earth's magnetic field and climate? Earth Plan. Sci. Lett. 253, 328–339.

Courtillot, V., Gallet, Y., LeMouël, J.-L., Fluteau, F., Genevey, A., 2007b. Response to comment on "Are there connections between Earth's magnetic field and climate? Earth Plan. Sci. Lett. 253, 328–339, 2007" by Bard, E. and Delaygue, M. Earth Plan. Sci. Lett. 265, 308–311.

Donadini, F., Kovacheva, M., Kostadinov, M., Casas, L., Pesonen, L.J., 2007. New archaeointensity results from Scandinavia and Bulgaria. Rock magnetic studies inference and geophysical application. Phys. Earth Plan. Inter. 165, 229–247.

Gallet, Y., Genevey, A., Fluteau, F., 2005. Does Earth's magnetic field secular variation control centennial climate change? Earth Plan. Sci. Lett. 236, 159–173.

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, 2285.

Gregory, P.C., 2005. Bayesian Logical Data Analysis for the Physical Sciences. Cambridge University Press, 468 pp.

Gubbins, D., 1975. Can the Earth's magnetic field be sustained by core oscillations? Geophys. Res. Lett. 2, 409–512.

Gubbins, D., 1983. Geomagnetic field analysis—I. Stochastic inversion. Geophys. J. R. Astron. Soc. 73, 641–652.

Gubbins, D., 1987. Mechanism for geomagnetic polarity reversals. Nature 326, 167–169.

Gubbins, D., 2004. Time Series Analysis and Inverse Theory for Geophysicists. Cambridge University Press, 255 pp.

Gubbins, D., Bloxham, J., 1985. Geomagnetic field analysis—III. Magnetic fields on the core-mantle boundary. Geophys. J. R. Astron. Soc. 80, 695–713.

Gubbins, D., Jones, A.L., Finlay, C.C., 2006. Fall in Earth's magnetic field is erratic. Science 312, 900–902.

Hongre, L., Hulot, G., Khokhlov, A., 1998. An analysis of the geomagnetic field over the past 2000 years. Phys. Earth Plan. Inter. 106, 311–335.

Hulot, G., Khokhlov, A., LeMouël, J.-L., 1997. Uniqueness of mainly dipolar magnetic fields recovered from directional data. Geophys. J. Int. 129, 347–354.

Jackson, A., 1997. Time-dependency of tangentially geostrophic core surface motions. Phys. Earth Plan. Inter. 103, 293–311.

Jackson, A., 2000. Comment on "Time evolution of the fluid flow at the top of the core. Geomagnetic jerks" by Le Huy, M., Mandea, M., Le Mouel, J.-L., and Pais, A. Earth Plan. Space 52, 649.

Jackson, A., Jonkers, A.R.T., Walker, M.R., 2000. Four centuries of geomagnetic secular variation from historical records. Philos. Trans. R. Soc. Lond. A 358, 957–990.

Jaynes, E.T., 2003. Probability Theory—The Logic of Science. Cambridge University Press, 727 pp.

Jeffreys, H., 1961. Theory of Probability, 2nd edition. Clarendon Press, Oxford, 459 pp.

Jones, A.L., 2005. The fall in the Earth's magnetic dipole moment over the last four hundred years, M. Geophys. Dissertation, University of Leeds, 68 pp.

Jonkers, A.R.T., Jackson, A., Murray, A., 2003. Four centuries of geomagnetic data from historical records. Rev. Geophys. 41, doi:10.1029/2002RG000115.

Korte, M., Constable, C.G., 2005a. Continuous geomagnetic field models for the past 7 millennia. 2. CALS7K. Geochem. Geophys. Geosyst. 6, doi:10.1029/2004GC000801.

Korte, M., Constable, C.G., 2005b. The geomagnetic dipole moment over the last 7000 years—new results from a global model. Earth Plan. Sci. Lett. 236, 348–358.

Korte, M., Constable, C.G., 2006a. Centennial to millennial geomagnetic secular variation. Geophys. J. Int. 167, 43–52.

Korte, M., Constable, C.G., 2006b. On the use of calibrated relative paleointensity records to improve millennial-scale geomagnetic field models. Geochem. Geophys. Geosyst. 7, doi:10.1029/2006GC001368.

Korte, M., Constable, C.G., 2008. Spatial and temporal resolution of millennial scale geomagnetic field models. Adv. Space Res. 41, 57–69.

Korte, M., Genevey, A., Constable, C.G., Frank, U., Schnepp, E., 2005. Continuous geomagnetic field models for the past 7 millennia: 1. A new global data compilation. Geochem. Geophys. Geosyst. 6, doi:10.1029/2004GC000800.

Le Goff, M., Gallet, Y., 2004. A new three-axis vibrating sample magnetometer for continuous high-temperature magnetization measurements: applications to paleo- and archeo-intensity determinations. Earth Plan. Sci. Lett. 229, 31–43.

MacKay, D.J., 1992. Bayesian interpolation. Neural Comp. 4 (3), 415–447.

MacKay, D.J., 2003. Information Theory, Inference, and Learning Algorithms. Cambridge University Press, 628 pp.

Parker, E.N., 1955. Hydromagnetic dynamo models. Astrophys. J. 122, 293–314.

Rotvig, J., 2007. Multiple zonal jets and drifting: thermal convection in a rapidly rotating spherical shell compared to a quasigeostrophic model. Phys. Rev. E 76, 046306(9 pages).

Sambridge, M., Gallagher, K., Jackson, A., Rickwood, P., 2006. Trans-dimensional inverse problems, model comparison and the evidence. Geophys. J. Int. 167, 528–542.

Shure, L.A., Parker, R.L., Backus, G., 1982. Harmonic splines for geomagnetic modelling. Phys. Earth Plan. Inter. 28, 215–229.

Sivia, D.S., Skilling, J., 2006. Data Analysis. A Bayesian Tutorial. Oxford University Press, 246 pp.

Snowball, I., Sandgren, P., 2004. Geomagnetic field intensity changes in Sweden between 9000 and 450 cal BP: extending the record of "archaeomagnetic jerks" by means of lake sediments and the pseudo-Thellier technique. Earth Plan. Sci. Lett. 227, 361–376.

Snowball, I., Zillen, L., Ojala, A., Saarinen, T., Sandgren, P., 2007. FENNOSTACK and FENNOPRIS: varve dated Holocene palaeomagnetic secular variation and relative palaeointensity stacks for Fennoscandia. Earth Plan. Sci. Lett. 255, 106–115.

Snowball, I., 1994. Bacterial magnetite and magnetic properties of sediments in a Swedish lake. Earth Plan. Sci. Lett. 126, 129–142.

Stuiver, M., Reimer, P.J., 1993. Extended Carbon-14 data base and revised CALIB 3.0 Carbon-14 age calibration. Radiocarbon 35, 215–230.

Tauxe, L., Yamazaki, T., 2007. Paleointensities. In: Kono, M. (Ed.), Treatise on Geophysics, Elsevier, vol. 5. pp. 510–555.

Trotta, R., 2007. Applications of Bayesian model selection to cosmological parameters. Mon. Not. R. Astron. Soc. 378, 72–82.

Valet, J.-P., 2003. Time variations in geomagnetic intensity. Rev. Geophys. 41, doi:10.1029/2001RG000104.

Williamson, J.H., 1968. Least-squares fitting of a straight line. Can. J. Phys. 46, 1845–1847.

Zhang, K., 2000. Is the geodynamo process intrinsically unstable? Geophys. J. Int. 140, F1–F4.

Zhang, K., Gubbins, D., 2000. Scale disparities and magnetohydrodynamics in the Earth's core. Philos. Trans. R. Soc. Lond. A 358, 899–920.