Geophysical Journal International

Geophys. J. Int. (2020) **223,** 648–665
Advance Access publication 2020 July 13
GJI Geomagnetism, Rock Magnetism and Palaeomagnetism

doi: 10.1093/gji/ggaa336

Correlation based snapshot models of the archeomagnetic field

S. Mauerberger [1], M. Schanner [2], M. Korte [2] and M. Holschneider [1]

[1] *Applied Mathematics, University of Potsdam, 14469 Potsdam, Germany. E-mail: mauerber@uni-potsdam.de*
[2] *Deutsches GeoForschungsZentrum GFZ, Sektion 2.3, Potsdam, Germany*

Accepted 2020 July 4. Received 2020 July 4; in original form 2020 February 4

SUMMARY

For the time stationary global geomagnetic field, a new modelling concept is presented. A Bayesian non-parametric approach provides realistic location dependent uncertainty estimates. Modelling related variabilities are dealt with systematically by making little subjective *a priori* assumptions. Rather than parametrizing the model by Gauss coefficients, a functional analytic approach is applied. The geomagnetic potential is assumed a Gaussian process to describe a distribution over functions. *A priori* correlations are given by an explicit kernel function with non-informative dipole contribution. A refined modelling strategy is proposed that accommodates non-linearities of archeomagnetic observables: First, a rough field estimate is obtained considering only sites that provide full field vector records. Subsequently, this estimate supports the linearization that incorporates the remaining incomplete records. The comparison of results for the archeomagnetic field over the past 1000 yr is in general agreement with previous models while improved model uncertainty estimates are provided.

Key words: Geopotential theory; Archaeomagnetism; Magnetic field variations through time; Palaeomagnetism; Inverse theory; Statistical methods.

1 INTRODUCTION

Global geomagnetic field reconstructions of the past millennia are useful to investigate the geodynamo process or the complex interaction of the field with solar wind particles and cosmic rays, and they find application in archeomagnetic and palaeomagnetic dating. Reconstructions are typically built from volcanic and archeomagnetic samples collected at the Earth's surface providing records of the ancient Earth's magnetic field (EMF). Unfortunately, on a global scale records are clustered, unevenly distributed towards the Western Eurasian region and corrupted by various uncertainties. This considerably complicates the reconstruction of the ancient EMF.

Dating back to 1985, Gubbins & Bloxham were amongst the first to propose a Bayesian inference for modelling the EMF, already discussing non-linear observables and model uncertainties. Their parametrized implementation of a truncated spherical harmonic (SH) representation with norm optimization has become a widely used modelling scheme. The majority of historical and archeomagnetic field models published over the past years essentially rely upon the same inverse strategy. Early models such as Jackson *et al.* (2000), Constable *et al.* (2000) and Korte & Constable (2003) provide estimates without quantifying uncertainties. More recent attempts—for example Korte *et al.* (2009), Licht *et al.* (2013), Hellio & Gillet (2018) and Senftleben (2019)—describe variabilities by deriving ensembles of equivalent solutions. Roughly speaking, those models differ in two aspects: On the one hand, the error handling, data selection and outlier detection have been refined over the years (Licht *et al.* 2013). On the other hand, different strategies are used to incorporate *a priori* knowledge. Early models are typically starting off from an axial dipole and are regularized by a physically motivated norm. Since regularized field models are known to underestimate uncertainties at small length scales (Gillet 2019), more elaborate modelling concepts are under investigation. Recent attempts deduce a priory information including temporal dynamics from the statistics of satellite era models (e.g. Hellio & Gillet 2018) or from geodynamo simulations (e.g. Sanchez *et al.* 2016). Existing models, however, have in common that uncertainties related to modelling, in particular due to model parameters and the uneven data distribution, are not dealt with systematically.

This paper introduces an advanced concept to model snapshots of the EMF. This work should be considered as a fist step towards a new inverse strategy in which the notion of modelling related uncertainties is well defined. Therefore, we adapt the correlation based inversion developed by Holschneider *et al.* (2016) that is known from modelling observatory and satellite data. Several modifications are required to adjust the concept to archeo- and palaeomagnetic data.

We pursue a fully Bayesian approach that determines the EMF's posterior distribution which simultaneously encodes the most probable field model and its uncertainties. To obtain the posterior distribution we use a functional analytic approach where inference takes place directly in the space of functions. Observables and quantities of interest are expressed in terms of functionals that act on the geomagnetic potential. Rather than using a model that is parametrized by a finite SH basis, the geomagnetic potential is assumed a Gaus-

sian process (GP). The GP in use is non-parametric in the sense that it is a distribution over functions and is specified by a two-point covariance function.

From a parametric point of view, GPs have been used for a long time in modelling the EMF (e.g. Bouligand *et al.* 2005; Khokhlov *et al.* 2006), known under the term Giant Gaussian Process (GGP). That term was coined by Constable & Parker (1988) who proposed a GP based model focusing on the estimation of model parameters. Our approach may be seen as the functional analytic extension of the GGP model.

Our *a priori* distribution of the geomagnetic potential is characterized by its mean power spectral behaviour, which is represented by an explicit correlation function that takes all SH degrees into account. If a SH truncation was desired, transdimensional modelling (Livermore *et al.* 2018) may be applied to also infer the cut-off degree. Using an explicit kernel function—not truncated at a certain SH degree—circumvents that problem. This does not necessarily mean that our approach reaches a higher resolution at a global scale but, improves treatment of the uneven data coverage and allows the exploitation of the records to their fullest. In addition, a low SH degree truncation may lead to spurious oscillations and ringings if the data include pronounced local anomalies. We try to be the least subjective and specify the *a priori* field model using uninformative distributions, when possible. Our *a priori* model depends only on a single parameter that controls the *a priori* power spectral behaviour.

In the case of satellite and observatory data, EMF full vector components are observed directly and observables are linearly related to the geomagnetic potential. Thus the posterior distribution for the GP is explicit and may be computed using ordinary linear algebra. Archeomagnetic data, however, call for a refined modelling strategy that takes the non-linearity of declination, inclination and intensity into account. The majority of sites only have incomplete vector information so that a direct linearization of each record is not possible. Therefore, we propose a two step Bayesian update system: First, a rough field estimate is obtained considering only sites that provide complete field vector records. Subsequently, this estimate supports the linearization that incorporates the remaining incomplete records.

To demonstrate the potential of our modelling strategy we present a case study using archeomagnetic and volcanic data of the past 1000 yr. Joint maps of best prediction and point-wise uncertainty are presented, which allow an improved interpretation of the spatial field structure. Although our modelling is not based upon a SH basis, we predict Gauss coefficients and quantify their uncertainties. In addition, we calculate the posterior mean of the spatial power spectrum and estimate error bounds. Finally, we present the posterior probability density function (PDF) for the dipole strength and the location of the geomagnetic north pole. The latter results are obtained for records of the past millennium, coarsely sorted into 100-yr bins and arranged into a discrete time-series.

The structure of the document is as follows: Section 2 gives an overview of the modelling theory, which closely follows Holschneider *et al.* (2016). First, we introduce the general construction of our dipole and non-dipole priors and correlations kernels (Section 2.1). We describe the link between observations and model, and how to obtain the posterior distribution from linear observations (Section 2.2), and then discuss the need for linearization of archeomagnetic observables (Section 2.3). The treatment of data uncertainties is laid out in Section 2.4. The following two Sections 3 and 4, give details about the necessary adjustments to model archeo- and palaeomagnetic data. Section 3 focuses on the two step strategy used to handle the non-linearities, formulated as a Bayesian update system and includes synthetic tests. Section 4 covers the translation of *a priori* uncertainties in the model parameters to the posterior. Finally, Section 5 provides the case study. We close the document by drawing conclusions and showing future perspectives in Section 6.

2 MODELLING CONCEPT

In this section we review the non-parametric and correlation-based modelling strategy (Holschneider *et al.* 2016) that underlies our Bayesian approach for modelling time-stationary snapshots of the EMF. We lay out our field model together with our *a priori* assumptions and establish the nomenclature that we adopt throughout this paper. We point out difficulties arising when working with non-linear observables such as archeo- and palaeomagnetic records and discuss the general treatment of data uncertainties.

2.1 Magnetic field model

When modelling the geomagnetic core field on archeo- to palaeomagnetic timescales, further contributions to the geomagnetic field—the crust, ionosphere and magnetosphere—are neglected since archeo- and palaeomagnetic measurement errors are assumed to be significantly larger than the respective field contributions (Constable & Korte 2015).

Close to the surface—far from magnetic sources—the EMF, $\boldsymbol{B}$, can be approximated by the gradient of a scalar potential satisfying Laplace's equation

$$\boldsymbol{B} = -\nabla\Phi, \quad \nabla^2\Phi = 0\,, \tag{1}$$

where Φ is referred to as the geomagnetic potential (Backus *et al.* 1996, chap. 4). In contrast to the magnetic field, the potential is not directly observable. For an internal source and with respect to (w.r.t.) some reference sphere of radius R, the potential Φ at location $\boldsymbol{x}$, $|\boldsymbol{x}| > R$, can be expanded in SHs

$$\Phi(\boldsymbol{x}) = R\sum_{\ell}\left(\frac{R}{|\boldsymbol{x}|}\right)^{\ell+1}\sum_{-l\le m\le l} g_\ell^m\, Y_\ell^m(\hat{\boldsymbol{x}}), \tag{2}$$

where Y_ℓ^m refers to the real valued and Schmidt semi-normalized SH of degree l and order m with related Gauss coefficient g_ℓ^m. The dependence of g_ℓ^m on a reference radius R is not explicitly typed.

We use a spherical coordinate system with B_N pointing to geographic north, B_E to the east and B_Z vertically downward. The components of the magnetic field vector $\boldsymbol{B}$ at a location of radius r, colatitude θ and longitude ϕ, are

$$B_N = -\frac{1}{r}\frac{\partial\Phi}{\partial\theta}, \quad B_E = \frac{1}{r\sin(\theta)}\frac{\partial\Phi}{\partial\phi}, \quad B_Z = -\frac{\partial\Phi}{\partial r}. \tag{3}$$

The ellipticity of the Earth is neglected and geocentric coordinates are treated as if they were geodetic.

Motivated by magnetic field theory, we express our lack of knowledge by making *a priori* assumptions about all Gauss coefficients. Considering Holocene timescales, the dipole part of the field does not have the same statistical distribution as the non-dipole part (Constable & Parker 1988). Thus, our model of choice is dipole dominated with additional random field contributions.

The dipole part is specified by the Gauss coefficients of degree $\ell = 1$. *A priori* the coefficients are assumed normal

$$\boldsymbol{g}_1 \sim \mathcal{N}(\bar{\boldsymbol{g}}_1,\ \Sigma_1)\,, \tag{4}$$

with mean vector $\bar{\boldsymbol{g}}_1$ and covariance matrix Σ_1, nine model parameters to be determined (three for the mean and six for the covariance).

The subscript is hinting at the SH degree and is going to be generalized. The dipole potential is a GP with mean and covariance function

$$\mathbb{E}[\Phi_{\mathrm{DP}}] = \boldsymbol{Y}_1^{\top}\boldsymbol{g}_1, \quad \mathbb{V}[\Phi_{\mathrm{DP}}] = \boldsymbol{Y}_1^{\top}\Sigma_1\boldsymbol{Y}_1, \tag{5}$$

where $\boldsymbol{Y}_1$ refers to the SH basis of degree $\ell = 1$, that is

$$\boldsymbol{Y}_1(\boldsymbol{x})^T = R\left(Y_1^0(\boldsymbol{x}), Y_1^1(\boldsymbol{x}), Y_1^{-1}(\boldsymbol{x})\right)\left(\tfrac{R}{|\boldsymbol{x}|}\right)^2. \tag{6}$$

Section 4.1 deals with choosing *a priori* dipole parameters. The correlation pattern for independent and identically distributed (IID) Dipole coefficients is shown in the left-hand panel of Fig. 1.

Our model of the non-dipole part is similar to the one proposed by Constable & Parker (1988) but within a Bayesian setting. The non-dipole potential is assumed a GP of zero mean. Rather than truncating the SH basis at a certain degree, a covariance function of closed form is used. Therefore, we adopt the kernel construction method from Holschneider *et al.* (2016, section 4). With respect to a reference sphere, the potential is characterized by its mean power spectral behaviour (see Section 5.6). That reference sphere may be seen as a virtual source region and the reference radius has no particular physical significance (Constable & Parker 1988). Except for the dipole, at the reference radius Gauss coefficients are assumed IID normal random variables. The according Legendre type kernel is of the form

$$K_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y}) = \lambda^2 R^2 \sum_{\ell=2}^{\infty}\left(\frac{R^2}{|\boldsymbol{x}||\boldsymbol{y}|}\right)^{\ell+1}\sum_{-l\le m\le l} Y_\ell^m(\hat{\boldsymbol{x}})Y_\ell^m(\hat{\boldsymbol{y}}) \tag{7}$$

where we introduced λ, a scaling factor that controls the amount of the non-dipole contribution and its dimension. Dipole and non-dipole parts are assumed statistically independent. According to Holschneider *et al.* (2016, eq. 54), K_{L} can be expressed in closed form. We choose this kernel as it is computationally simple and depends only on two parameters, R and λ. The right-hand panel of Fig. 1 depicts an *a priori* correlation pattern for the non-dipole potential.

In principle an SH decomposition is possible but computationally limited. To cover highly localized modelling errors, the covariance has to map the characteristic length scales present in the data. Small scale correlations among records are a valuable source of information that should enter the model. In case of an expansion, the highest SH degree must be chosen such that the smallest spatial wavelength coincides with distances between sites. If some sites were clustered with distances of about 100 km, an expansion up to degree $\ell \le 175$ would be needed. Since high degree expansions are demanding, using an explicit kernel function is computationally beneficial and makes it feasible to adopt a global point of view while preserving the accuracy of local length scales.

The reference radius R controls the predominant slope of the *a priori* power spectrum. The smaller R, the smoother the *a priori* field at the surface. The scaling factor λ causes a shift along the axis of ordinates. Similar to Constable & Parker (1988), the power spectrum is used to tune the reference radius. R is chosen such that the prior mean power spectrum roughly conforms with the International Geomagnetic Reference Field (IGRF, Thébault *et al.* 2015) time average from 1900 to 2020. Fig. 2 depicts the alignment, carried out by visual inspection. Throughout we are going to use the fixed value $R = 2800$ km. In contrast, as λ is highly uncertain, it receives special treatment in Section 4.2. A reference radius below the core–mantle boundary (CMB) implies a non-erratic covariance structure at the CMB. A virtual source region within the outer core may seem unconventional but is becoming more popular. To give an example, Sanchez *et al.* (2016) are using complex correlation patterns at the CMB obtained from dynamo simulations as prior information.

The EMF is modelled as the negative gradient of the potential (eq. 1). Differentiation is a linear operation and thus the field model is a GP as well. The *a priori* mean reads

$$\bar{\boldsymbol{B}}(\boldsymbol{x}) = -\nabla\boldsymbol{Y}_1^{\top}(\boldsymbol{x})\bar{\boldsymbol{g}}_1 = -\sum_{-1\le m\le 1}\nabla Y_1^m(\boldsymbol{x})\bar{g}_1^m \tag{8}$$

and the correlation kernel is composed of dipole and non-dipole covariance functions

$$\begin{aligned} K_{\boldsymbol{B}}(\boldsymbol{x}, \boldsymbol{y}) &= K_{\boldsymbol{B},\mathrm{DP}} + K_{\boldsymbol{B},\mathrm{ND}} \\ &= \nabla\boldsymbol{Y}_1^{\top}(\boldsymbol{x})\Sigma_1\boldsymbol{Y}_1(\boldsymbol{y})\nabla^{\dagger} + \nabla K_{\mathrm{L}}(\boldsymbol{x}, \boldsymbol{y})\nabla^{\dagger}, \end{aligned} \tag{9}$$

where the right-hand gradient acts on the left at the second argument.

2.2 Inference

In contrast to existing archeomagnetic field models we pursue a functional analytic approach and perform a regression directly in the space of functions. The corner stone of our modelling strategy is formed by a GP regression. GP regression is also known in the field of geostatistics as Kriging (see Rasmussen & Williams 2006, sections 2.2 and 2.8). For linear $\boldsymbol{B}$-field observations this strategy has already been adopted by Holschneider *et al.* (2016). We briefly recall the overall modelling concept before we introduce non-linear observations.

Suppose we observe the components B_N, B_E and B_Z of the EMF. Measurements are denoted by

$$o = \left\{\boldsymbol{o}_i = \left(B_N(\boldsymbol{z}_i), B_E(\boldsymbol{z}_i), B_Z(\boldsymbol{z}_i)\right)^{\top}\right\}_{i=1,\dots,n}, \tag{10}$$

recorded at locations $\boldsymbol{z}_i$. Observations are corrupted by additive noise $\boldsymbol{E}_i$ and the data model reads

$$O = \{\boldsymbol{B}(\boldsymbol{z}_i) + \boldsymbol{E}_i\}_{i=1,\dots,n}. \tag{11}$$

The error model is assumed normal of zero mean and covariance Σ_E. Recorded values o are assumed a realization of O. The data's *a priori* mean vector and covariance matrix are

$$\bar{O} = \left\{\bar{\boldsymbol{B}}(\boldsymbol{z}_i)\right\}_{i=1,\dots,n} \tag{12}$$

$$\Sigma_O = \left\{K_{\boldsymbol{B}}(\boldsymbol{z}_i, \boldsymbol{z}_j)\right\}_{i,j=1,\dots,n} + \Sigma_E \tag{13}$$

with typically diagonal error-covariance matrix Σ_E.

To obtain information about the EMF, we need to compute $\boldsymbol{B}$'s posterior distribution. If we assume the $\boldsymbol{B}$-field and measurement errors are independent, the cross covariance matrix follows to read

$$\Sigma_{\boldsymbol{B}(\boldsymbol{x})O} = \mathrm{Cov}[\boldsymbol{B}(\boldsymbol{x}), O] = \{K_{\boldsymbol{B}}(\boldsymbol{x}, \boldsymbol{z}_i)\}_{i=1,\dots,n} \tag{14}$$

for any design point $\boldsymbol{x}$ outside the reference sphere. Since O and $\boldsymbol{B}$ are jointly Gaussian, the posterior distribution is normal as well. It is fully determined by the conditional mean and conditional covariance

$$\mathbb{E}[\boldsymbol{B}(\boldsymbol{x})|o] = \bar{\boldsymbol{B}}(\boldsymbol{x}) + \Sigma_{\boldsymbol{B}O}\Sigma_O^{-1}\left(o - \bar{O}\right) \tag{15}$$

$$\mathrm{Cov}[\boldsymbol{B}(\boldsymbol{x}), \boldsymbol{B}(\boldsymbol{y})|o] = K_{\boldsymbol{B}}(\boldsymbol{x}, \boldsymbol{y}) - \Sigma_{\boldsymbol{B}O}\Sigma_O^{-1}\Sigma_{\boldsymbol{B}O}^{\top}. \tag{16}$$

Gauss coefficients are modelled analogously. Magnetic potential and Gauss coefficients are related through

$$g_\ell^m = \frac{2\ell+1}{4\pi R}\iint Y_\ell^m(\boldsymbol{x})\Phi(\boldsymbol{x})\,\mathrm{d}^2\boldsymbol{x}\,, \tag{17}$$

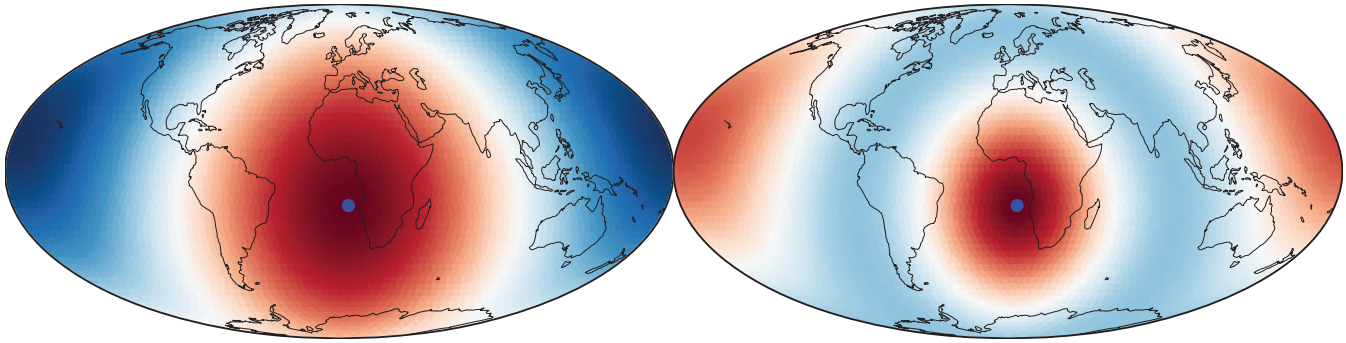

Figure 1. Visualization of dipole (left-hand panel) and non-dipole (right-hand panel) correlation structures at the Earth's surface. The reference location is indicated by the blue dot. The dipole pattern corresponds to IID dipole coefficients. The apparent correlation of antipodal points stems from the fact that this kernel describes the fluctuations around the removed dipole, which is dominated by the quadrupole contributions. Overall cross correlations result from superimposing dipole and non-dipole parts. Correlations are normalized and a diverging colourmap is used, ranging from −1 (blue) to 1 (red).

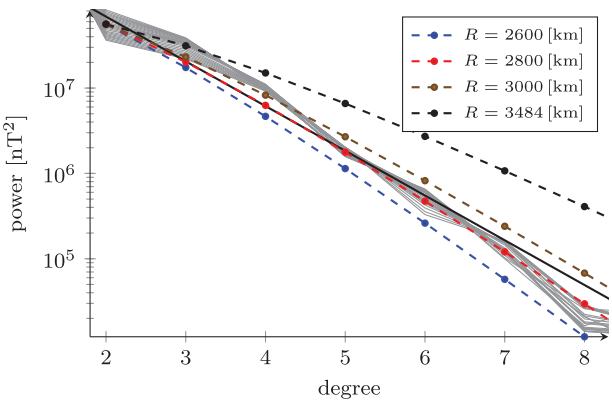

Figure 2. The mean power spectral behaviour chosen *a priori* is indicated by red dots. As a reference, the power spectra of the IGRF from 1900 to 2020 are drawn in grey. For comparison, the solid black line refers to the mean spectrum suggested by Constable & Parker (1988).

where integration is carried out over the sphere of radius R. The collection of Gauss coefficients up to SH degree ℓ is denoted by $\boldsymbol{g}_\ell$. Accordingly, Σ_ℓ refers to the prior covariance matrix of Gauss coefficients up to degree ℓ. Except for the dipole and according to eq. (7), the *a priori* covariance Σ_ℓ is diagonal. At the reference radius the *a priori* variance is λ^2. The cross covariance matrix between $\boldsymbol{g}_\ell$ and the observations reads

$$\Sigma_{\ell O} = \mathrm{Cov}\big[\boldsymbol{g}_\ell, O\big] = \{-\Sigma_\ell \nabla \boldsymbol{Y}_\ell(\boldsymbol{z}_i)\}_{i=1,\ldots,n} \tag{18}$$

and $\boldsymbol{Y}_\ell$ refers to the SH basis up to degree ℓ. Since Gauss coefficients and potential are linearly related, the posterior distribution is normal as well and, again, fully determined by the conditional mean and covariance

$$\mathbb{E}\big[\boldsymbol{g}_\ell | o\big] = \bar{g}_\ell + \Sigma_{\ell O}\Sigma_O^{-1}\big(o - \bar{O}\big) \tag{19}$$

$$\mathbb{V}\big[\boldsymbol{g}_\ell | o\big] = \Sigma_\ell - \Sigma_{\ell O}\Sigma_O^{-1}\Sigma_{\ell O}^\top. \tag{20}$$

It is worth mentioning that Bouligand *et al.* (2005) used geodynamo simulations to conclude that there are significant cross-correlations among the Gauss coefficients. Our statistical model places no restrictions on the posterior cross-correlations although the *a priori* assumptions are based on IID Gauss coefficients.

These are the formulae we are going to build our modelling strategy upon. However, they require an extension since archeomagnetic records and the magnetic potential are non-linearly related. Before elaborating our approach to this problem, let us first recall the observational functionals in question.

2.3 Observational functionals and linearization

Archeomagnetic data is not provided in the form of Cartesian field vector components. The quantities that are determined in laboratory experiments are the two angles, declination (D) and inclination (I) and intensity (F), acquired at locations $\boldsymbol{z}$. These quantities and the vector components are non-linearly related. The corresponding functionals read

$$H : \boldsymbol{B} \to \begin{pmatrix} D \\ I \\ F \end{pmatrix} = \begin{pmatrix} \arctan\left(\frac{B_E}{B_N}\right) \\ \arctan\left(\frac{B_Z}{F_H}\right) \\ \sqrt{B_N^2 + B_E^2 + B_Z^2} \end{pmatrix}, \tag{21}$$

where the horizontal intensity

$$F_H = \sqrt{B_N^2 + B_E^2} \tag{22}$$

is introduced as an auxiliary variable (Backus *et al.* 1996, eq. 1.2.1–1.2.4). The components of H are referred to as observation functionals and are denoted by $\mathrm{H}_i[\boldsymbol{B}]$ for $i = D, I, F$. The inverse map to magnetic field vector components reads

$$H^{-1} : \begin{pmatrix} D \\ I \\ F \end{pmatrix} \to \boldsymbol{B} = F \begin{pmatrix} \cos(I)\cos(D) \\ \cos(I)\sin(D) \\ \sin(I) \end{pmatrix}. \tag{23}$$

The inverse map is only unique if all three observables are at hand, that is three vector components relate uniquely to three observables and vice versa.

As already pointed out in the previous Section, the vital prerequisite for the modelling strategy is joint normality of observations and EMF. Certainly, the functionals D, I and F do not preserve $\boldsymbol{B}$'s normality, nor are the measurement errors Gaussian.

To adopt the modelling concept by Holschneider *et al.* (2016) we approximate D, I and F by a 1st order Taylor expansion

$$\mathrm{H}_i[\boldsymbol{B}] \approx \mathrm{H}_i[\tilde{\boldsymbol{B}}] + \nabla \mathrm{H}_i[\tilde{\boldsymbol{B}}]^\top\big(\boldsymbol{B} - \tilde{\boldsymbol{B}}\big), \tag{24}$$

where $\tilde{\boldsymbol{B}}$ refers to an arbitrary point of expansion (POE). This is the non-parametric counterpart compared with the approach presented by Gubbins & Bloxham (1985,eq. 10). The functionals approximating D, I and F arise to

$$D \approx \tilde{D} + \frac{1}{\tilde{F}_H^2} \begin{bmatrix} -\tilde{B}_E \\ \tilde{B}_N \\ 0 \end{bmatrix}^\top \boldsymbol{B}, \tag{25}$$

$$I \approx \tilde{I} + \frac{1}{\tilde{F}_H} \left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{\tilde{B}_Z}{\tilde{F}} \frac{\tilde{\boldsymbol{B}}}{\tilde{F}} \right)^\top \boldsymbol{B}, \tag{26}$$

$$F \approx \frac{\tilde{\boldsymbol{B}}^\top}{\tilde{F}} \boldsymbol{B} \tag{27}$$

where $\tilde{D}$, $\tilde{I}$ and $\tilde{F}$ are referring to the 0th order terms. From those equations it is obvious that performing an expansion about zero is not going to work. Existing models (e.g. Licht *et al.* 2013; Hellio & Gillet 2018) typically use an axial dipole as initial POE. Since a Taylor expansion performs better the less $\boldsymbol{B}$ deviates from the POE, we propose not to use an axial dipole. Section 3 is dedicated to which point to linearize about.

Since $\boldsymbol{B}$ is assumed a GP and due to linearity, the approximating observational functionals are normally distributed. However, to actually achieve joint normality amongst observations and $\boldsymbol{B}$-field, measurement errors need a Gaussian proxy as well.

2.4 Measurement errors

Although the observational functionals were linearized, the data model is still not normal since measurement errors are not necessarily Gaussian. It is common to characterize the uncertainty of archeomagnetic directions (D and I) using the von Mises-Fisher (vMF) distribution and the variability of intensity using the normal distribution (Love & Constable 2003). As long as intensity records report the standard error, linearizing F provides a normal proxy model which we are going to use for inference. However, we lack a Gaussian proxy for directional errors.

The commonly used approach is to calculate directional errors individually. The vMF distribution is parametrized by a concentration factor and a location parameter. The larger the value of the concentration factor, the more the distribution tends towards concentrating around the location parameter. Provided a large concentration factor and that the location parameter is not pointing towards high latitudes, the marginal probability densities for D and I are approximately Gaussian. Proxy errors for declination and inclination are assumed independent, of zero mean and standard deviation

$$\sigma_I = \frac{57.3^\circ}{140} \alpha_{95} \quad \text{and} \quad \sigma_D = \frac{1}{\cos o_I} \sigma_I \tag{28}$$

(Suttie & Nilsson 2019). Typically directional records report the 95 per cent confidence cone α_{95} of the vMF distribution. This is a pragmatic approach that does not necessarily reflect the Gaussian moment matching proxy, since correlations are dropped and—in general—the mean of a vMF distribution does not coincide with the location parameter. In case of isolated declinations this approach does not work and such records are rejected. Another drawback of a Gaussian proxy error model is the intolerance against outliers. In the context of optimization theory there exist more robust approaches (e.g. Farquharson & Oldenburg 1998; Walker & Jackson 2008; Hellio *et al.* 2014). For the vast majority of non-Gaussian likelihoods, however, no explicit solution to the Bayesian inverse problem is known.

Two philosophies have been used in previous work when modelling archeo- and palaeomagnetic data. Either very strict selection criteria are applied and these often contain tests that had not been applied in data published a few decades ago. As the absence of the test does not necessarily mean a result would not have passed the test, the other philosophy is to include as much data as possible without applying very strict criteria, aiming to increase the signal to noise ratio. However, it is likely that the reported measurement errors in several cases underestimate the true data uncertainties, which might contain systematic biases if corrections for, for example, cooling rate or anisotropy have not been performed.

For practical reasons, we adopt the second philosophy. To compensate possibly non-conforming error estimates we introduce a scaling parameter ϵ. Then, the individual proxy uncertainties are given by

$$E_i \to \epsilon E_i \sim \mathcal{N}\left(0,\, \epsilon^2 \sigma_i^2\right) \tag{29}$$

and ϵ is regarded as a model parameter. Although we do not know the specific value for ϵ, its order of magnitude is assumed to be one.

In addition, we introduce a residual term $\boldsymbol{P}$ that compensates modelling related errors and accounts for observational biases. Amongst others, effects that our model does not include are temporal correlations, dating errors and crustal field anomalies. Therefore, we assume $\boldsymbol{P} \sim \mathcal{N}(0,\, \boldsymbol{I})$, that is uncorrelated standard normal at every pair of distinct sites. Our final data model becomes

$$o_i = \mathrm{H}_i[\boldsymbol{B}_i + \rho \boldsymbol{P}_i] + \epsilon E_i \tag{30}$$

and the magnitude of the residual is controlled by ρ, another not yet known model parameter. The residual term can be thought of as an error term that describes the simplification of the underlying physics statistically. In other words, those real world contributions that are not covered by our model are treated as if they were random errors.

Because we focus on time stationary snapshots of the EMF, dating uncertainties are displaced into the residual term. Nonetheless, these errors are of importance if the temporal behaviour of the EMF is reconstructed. There already exist several approaches to accommodate dating errors, for example Jack-knife (Korte *et al.* 2009; Licht *et al.* 2013), MCMC sampling (Hellio *et al.* 2014) or transdimensional modelling (Livermore *et al.* 2018).

To proceed with the modelling concept outlined in Section 2.2, we combine the linearization and the Gaussian proxy for the directional errors. However, a POE for the linearization is still missing. The subsequent section covers this problem and also discusses the concrete incorporation of the error term and the residual. As the final ingredient of our modelling concept, the treatment of model parameters is subject of Section 4.

3 BAYESIAN UPDATE SYSTEM

The need for linearization described in Section 2.3 requires a suitable POE. The linearization as a Taylor expansion performs better the less the POE deviates from the *truth*. Archeo- and palaeomagnetic directional and intensity data are determined from different laboratory experiments, and the majority of records report either one or two field components (incomplete records). The complete vector information (D, I, F) is only available in rare cases. Noting that it is easier to determine a POE from full vector records, we introduce a Bayesian update system to treat complete and incomplete records successively.

The posterior distribution is computed by a two step strategy only considering a subset of observations at a time. Records are partitioned into two disjoint groups $o_{\mathcal{I}}$ and $o_{\mathcal{C}}$ where subscripts are referring to incomplete and complete measurements. Making use of the conditional probability rule—that is $p(X|Y)p(Y) = p(X, Y)$—and

according to Bayes' law the posterior $\boldsymbol{B}$-field factorizes

$$p(\boldsymbol{B}|o) = p(\boldsymbol{B}|o_{\mathcal{C}}, o_{\mathcal{I}}) = \frac{p(o_{\mathcal{I}}|\boldsymbol{B}, o_{\mathcal{C}})}{p(o_{\mathcal{I}}|o_{\mathcal{C}})} p(\boldsymbol{B}|o_{\mathcal{C}}), \tag{31}$$

that is, the posterior EMF based on the complete observations $o_{\mathcal{C}}$ serves as prior for the Bayesian posterior based on $o_{\mathcal{I}}$. Not to incorporate the data all at once appears to be a promising strategy due to strong magnetic field correlations.

The complexity of the developed algorithm is growing through this and the following section. Fig. 3 provides a schematic illustration so as not to loose the overview. The two step strategy is shown in the top panel of Fig. 3.

3.1 Complete records

In the initial step only complete records are taken into account. Triplets of declination, inclination and intensity are forming the set of complete records

$$o_{\mathcal{C}} = \left\{ \boldsymbol{o}_i = (D(\boldsymbol{z}_i), I(\boldsymbol{z}_i), F(\boldsymbol{z}_i))^{\top} \right\}_{i=1,\dots,n_{\mathcal{C}}}. \tag{32}$$

In order to apply the linearization, a POE is missing. The special case of knowing all three components allows the calculation of the inverse map (eq. 23)

$$\tilde{\boldsymbol{B}}_{\mathcal{C}} = \left\{ H^{-1}[\boldsymbol{o}_i] \right\}_{i=1,\dots,n_{\mathcal{C}}} \tag{33}$$

which will serve as the point to linearize about. That point is reasonable as long as the prior $\boldsymbol{B}$-field variance is large in comparison with the measurement errors. If measurement errors were not negligible w.r.t. the *a priori* distribution, then the POE is no longer known with confidence and we have to propagate uncertainties. To ensure this, Section 4 is devoted to choosing an uninformative *a priori* field. A Gaussian approximation in this way is also known as Laplace's method (Murphy 2012, section 8.4.1).

In order to apply the modelling scheme introduced in Section 2.2 we use the Gaussian proxy error model. The diagonal error covariance matrix is denoted by

$$\Sigma_{E,\mathcal{C}} = \mathrm{diag}\left(\sigma_1^2, \dots, \sigma_{n_{\mathcal{C}}}^2\right), \tag{34}$$

where σ_i refers to individual standard errors w.r.t. D, I and F. The linearized observation functionals translate between (D, I, F) and $\boldsymbol{B}$. To keep equations concise, the dipole basis and the Jacobi matrices are collected in big matrices

$$\acute{Y}_{1,\mathcal{C}} = \{\nabla \boldsymbol{Y}_1(\boldsymbol{z}_i)\}_{i=1,\dots,n_{\mathcal{C}}} \qquad (3n_{\mathcal{C}} \times 3) \tag{35}$$

$$\acute{\mathrm{H}}_{\mathcal{C}} = \left\{ \delta_{ij} \nabla H[\tilde{\boldsymbol{B}}_i] \right\}_{i,j=1,\dots,n_{\mathcal{C}}} \qquad (3n_{\mathcal{C}} \times 3n_{\mathcal{C}}), \tag{36}$$

where δ_{ij} refers to the Kronecker delta and $\acute{\mathrm{H}}_{\mathcal{C}}$ is 3×3 block-diagonal. For $o_{\mathcal{C}}$, the approximative prior mean vector is given by

$$\bar{O}_{\mathcal{C}} \approx o_{\mathcal{C}} + \acute{\mathrm{H}}_{\mathcal{C}}^{\top} \left(-\acute{Y}_{1,\mathcal{C}} \bar{\boldsymbol{g}}_1 - \tilde{\boldsymbol{B}}_{\mathcal{C}} \right), \tag{37}$$

where $\boldsymbol{B}_{\mathcal{C}}$ means evaluated at all the locations of observation. Due to assumed independence of error model and EMF, the linearized covariance matrix for complete records reads

$$\Sigma_{\mathcal{C}} = \mathbb{V}[O_{\mathcal{C}}] \approx \acute{\mathrm{H}}_{\mathcal{C}}^{\top} \left(\mathbb{V}[\boldsymbol{B}_{\mathcal{C}}] + \rho^2 1 \right) \acute{\mathrm{H}}_{\mathcal{C}} + \epsilon^2 \Sigma_{E,\mathcal{C}}, \tag{38}$$

where $\mathbb{V}[\boldsymbol{B}_{\mathcal{C}}]$ is constructed from the kernel (eq. 9) at any two locations of observation. Due to bi-linearity of the covariance $\acute{\mathrm{H}}_{\mathcal{C}}$ is factored out and all constant terms are stripped off. For arbitrary design points, the linearized cross covariance amongst EMF and measurements is given by

$$\Sigma_{\boldsymbol{B}\mathcal{C}} = \mathrm{Cov}[\boldsymbol{B}, O_{\mathcal{C}}] \approx \mathrm{Cov}[\boldsymbol{B}, \boldsymbol{B}_{\mathcal{C}}] \acute{\mathrm{H}}_{\mathcal{C}}. \tag{39}$$

According to eqs (15) and (16), a Gaussian proxy of $\boldsymbol{B}$'s posterior distribution is determined through the conditional mean $\mathbb{E}[\boldsymbol{B}|o_{\mathcal{C}}]$ and conditional covariance $\mathbb{V}[\boldsymbol{B}|o_{\mathcal{C}}]$. This first step of our modelling scheme is illustrated in the top left panel of Fig. 3.

Analogous to eq. (18), the linearized cross covariance amongst Gauss coefficients and observations is given by

$$\Sigma_{\ell\mathcal{C}} = \mathrm{Cov}\left[\boldsymbol{g}_{\ell}, O_{\mathcal{C}}\right] \approx \mathrm{Cov}\left[\boldsymbol{g}_{\ell}, \boldsymbol{B}_{\mathcal{C}}\right] \acute{\mathrm{H}}_{\mathcal{C}} = -\Sigma_{\ell} \acute{Y}_{\ell,\mathcal{C}} \acute{\mathrm{H}}_{\mathcal{C}}. \tag{40}$$

where $\acute{Y}_{\ell,\mathcal{C}}$ extends eq. (35) up to SH degree ℓ. The approximations of the conditional mean $\mathbb{E}\left[\boldsymbol{g}_{\ell}|o_{\mathcal{C}}\right]$ and conditional covariance $\mathbb{V}\left[\boldsymbol{g}_{\ell}|o_{\mathcal{C}}\right]$ are given through eqs (19) and (20).

Although the subset of complete records is comparatively small, as a first *guess*, we anticipate a reconstruction of the EMF's dominating features due to strong magnetic field correlations.

3.2 Incomplete records

Inference with the incomplete records only implicitly depends on the *a priori* mean and covariance function through the first step. The distribution of the EMF posterior to $o_{\mathcal{C}}$ may be understood as the prior to incorporate the remaining measurements $o_{\mathcal{I}}$. To carry out the linearization, the mean conditional on $o_{\mathcal{C}}$ will serve as POE

$$\tilde{\boldsymbol{B}}_{|\mathcal{C}} = \mathbb{E}[\boldsymbol{B}|o_{\mathcal{C}}]. \tag{41}$$

From a theoretical point of view this is a function that is going to be evaluated by the observational functionals. We expect those points to be well suited, as in a Gaussian model the mean is the most likely solution. However, this is an arbitrary choice and does not necessary imply any optimality.

Because the EMF's proxy posterior to $o_{\mathcal{C}}$ is normal we are going to use the same modelling concept as we have already done (Section 2.2). To facilitate the second step it is necessary not only to predict on design points but also on *auxiliary* quantities in the first step.

Incomplete records are treated individually even though a certain location may report more than one observable. The approximative mean is given by

$$\bar{O}_{\mathcal{I}|\mathcal{C}} \approx \left\{ \mathrm{H}_i\left[\tilde{\boldsymbol{B}}_{|\mathcal{C}}\right] \right\}_{i=1,\dots,n_{\mathcal{I}}} \tag{42}$$

and we use the subscript i to indicate both the location $\boldsymbol{z}_i$ and the type of record in H_i, either D, I or F. If we collect all gradients within one big matrix

$$\acute{H}_{\mathcal{I}|\mathcal{C}} = \left\{ \delta_{ij} \nabla \mathrm{H}_i[\tilde{\boldsymbol{B}}_{|\mathcal{C}}] \right\}_{ij=1,\dots,n_{\mathcal{I}}}, \tag{43}$$

then the covariance matrix for incomplete records reads

$$\Sigma_{\mathcal{I}|\mathcal{C}} \approx \acute{H}_{\mathcal{I}|\mathcal{C}}^{\top} \left(\mathbb{V}[\boldsymbol{B}_{\mathcal{I}}|o_{\mathcal{C}}] + \rho^2 1 \right) \acute{H}_{\mathcal{I}|\mathcal{C}} + \epsilon^2 \Sigma_{E,\mathcal{I}}. \tag{44}$$

The error covariance matrix is analogous to eq. (34). The auxiliary quantities we have to carry along are conditional mean vector and covariance matrix at points of observation.

To model the posterior $\boldsymbol{B}$-field we again have to calculate linearized cross correlations amongst $\boldsymbol{B}|o_{\mathcal{C}}$ and $O_{\mathcal{I}}$. If we store the matrix $\mathrm{Cov}[\boldsymbol{B}, \boldsymbol{B}_{\mathcal{I}}|o_{\mathcal{C}}]$ within the first step, linearized cross correlations amongst design points and incomplete records are given by

$$\Sigma_{\boldsymbol{B}\mathcal{I}|\mathcal{C}} = \mathrm{Cov}[\boldsymbol{B}, O_{\mathcal{I}}|o_{\mathcal{C}}] \approx \mathrm{Cov}[\boldsymbol{B}, \boldsymbol{B}_{\mathcal{I}}|o_{\mathcal{C}}] \acute{H}_{\mathcal{I}}. \tag{45}$$

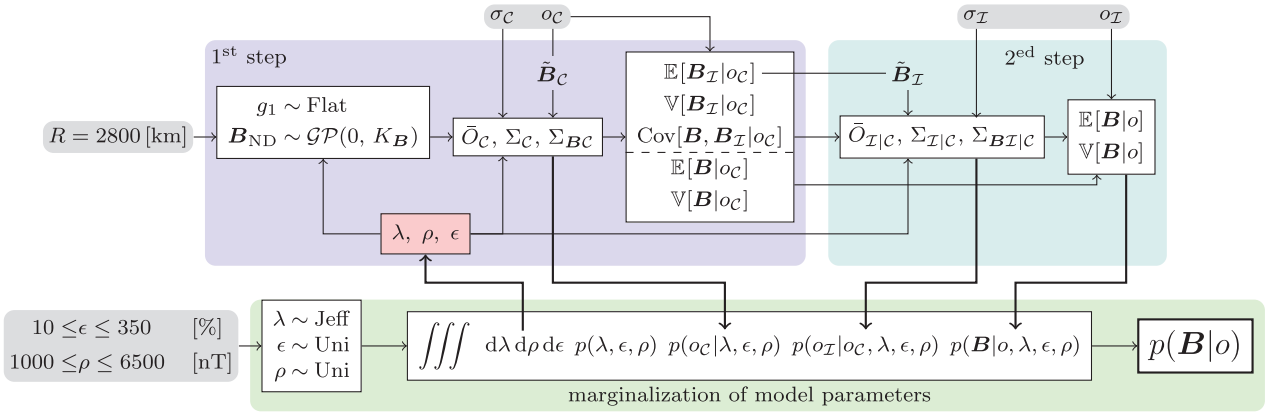

Figure 3. Illustration of how the two-step strategy and a marginalization intertwine to compute the posterior compound distribution. Invariants—such as observations—are shaded in grey. The upper part refers to the update system where model parameters are highlighted in red. Top left-hand panel shows the initial step (Section 3.1) whereas top right-hand illustrates the update (Section 3.2). The marginalization is highlighted in green (Section 4.2). Arrows indicate how information is passed.

The Gaussian proxy for the EMF's posterior distribution is again determined through eqs (15) and 16. Conditional mean and covariance read

$$\mathbb{E}[\boldsymbol{B}|o] = \mathbb{E}[\boldsymbol{B}|o_{\mathcal{C}}] + \Sigma_{\boldsymbol{B}\mathcal{I}|\mathcal{C}}\Sigma^{-1}_{\mathcal{I}|\mathcal{C}}(o_{\mathcal{I}} - \bar{O}_{\mathcal{I}|\mathcal{C}}) \tag{46}$$

$$\mathbb{V}[\boldsymbol{B}|o] = \mathbb{V}[\boldsymbol{B}|o_{\mathcal{C}}] - \Sigma_{\boldsymbol{B}\mathcal{I}|\mathcal{C}}\Sigma^{-1}_{\mathcal{I}|\mathcal{C}}\Sigma^{\top}_{\boldsymbol{B}\mathcal{I}|\mathcal{C}}. \tag{47}$$

The top panel of Fig. 3 illustrates how the posterior EMF is built within two steps.

Gauss coefficients are estimated analogously. If we store the matrix $\mathrm{Cov}\big[\boldsymbol{g}_\ell, \boldsymbol{B}_{\mathcal{I}}|\mathcal{C}\big]$ while performing the first step, then, the linearized cross correlations are given by

$$\Sigma_{\ell\mathcal{I}|\mathcal{C}} = \mathrm{Cov}\big[\boldsymbol{g}_\ell, O_{\mathcal{I}}|o_{\mathcal{C}}\big] \approx \mathrm{Cov}\big[\boldsymbol{g}_\ell, \boldsymbol{B}_{\mathcal{I}}|o_{\mathcal{C}}\big]\,\acute{H}_{\mathcal{I}|\mathcal{C}}. \tag{48}$$

Approximations of conditional mean and covariance translate according to eqs (19) and (20).

Unfortunately, we cannot yet directly apply that algorithm since we do not know specific values for the model parameters $\bar{\boldsymbol{g}}_1$, Σ_1, λ, ρ and ϵ. Before we deal with abandoning these parameters (see Section 4) we carry out tests with synthetic data to validate the proposed framework.

3.3 Tests

The current IGRF coefficients are used as a realistic reference field. Synthetic data is generated from the reference field and is corrupted by artificial noise. A gamma distribution is used to corrupt intensity records whereas directional data is randomly drawn from a vMF distribution. To check for robustness, the measurement errors are chosen on purpose not to coincide with the proposed Gaussian proxy error model. For the tests we use several data sets that differ in the error level, the complete/incomplete ratio and the spatial distribution. Since the reference model is known, we use the mean absolute error (MAE)

$$\mathrm{MAE} = \frac{1}{n}\sum_{j}^{n} \big\| \boldsymbol{B}(\boldsymbol{z}_j) - \mathbb{E}\big[\boldsymbol{B}(\boldsymbol{z}_j)|o\big] \big\|_1 \tag{49}$$

as a test characteristic. Reconstruction and reference field are compared at the Earth's surface, sampled at a rate that accounts for the length scales present in the reference field. Our modelling strategy is able to recover the reference field for all data sets considered.

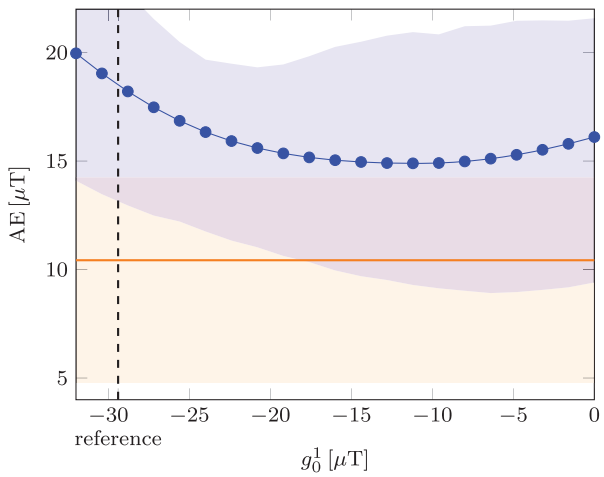

Figure 4. Linearization about an axial dipole compared to the strategy proposed. The horizontal red line indicates the MAE of our approach whereas the blue dots refer to the MAE by varying the axial dipole strength. The shaded areas extend from the 25 to 75 per cent quantile to illustrates the spread of absolute errors. The dashed vertical line refers to the axial dipole of the reference field.

Furthermore the influence of the POE on the linearization is examined. We compare the proposed strategy to linearization about an axial dipole of $g_1^0 = -23\,\mu\mathrm{T}$. For all data sets under consideration, the MAE of our modelling strategy falls below linearization about the axial dipole. Especially if the data set mimics reality, the proposed strategy performs better.

To make sure that $g_1^0 = -23\,\mu\mathrm{T}$ is not a coincidence, we also vary the axial dipole strength. For the synthetic data that mimic reality, Fig. 4 compares the MAE of the proposed strategy with linearizing about a range of axial dipoles. Even the axial dipole featuring the smallest MAE has a value that is still above the proposed POE built from the complete data. Throughout the synthetic data sets considered, the proposed strategy outperforms the linearization about the *best* suited axial dipole. A detailed description and the full test results are available together with the source code by Schanner & Mauerberger (2019).

4 MODEL PARAMETERS

Before applying the outlined modelling strategy we need to address several model parameters. As the posterior EMF must not depend on a certain choice of model parameters, the *a priori* mean $\bar{\boldsymbol{g}}_1$ and covariance Σ_1 of the dipole and the scaling factors λ, ρ and ϵ are considered so-called *nuisance* parameters, that ias parameters that are not of primary interest. To abandon those quantities from the EMF's posterior distribution we are going to use two different techniques: Section 4.1 analytically eliminates $\bar{\boldsymbol{g}}_1$ and Σ_1 by exploring the limit of a flat dipole prior. Section 4.2 addresses a hierarchical Bayesian approach to marginalize the scaling factors λ, ρ and ϵ.

4.1 Uninformative dipole

To make as few assumptions as possible we aim to explore the limit of an uninformative prior dipole. This approach is beneficial in two different ways: (1) It prevents us from accidentally choosing an overly confident prior. (2) We abandon all of the nine dipole-parameters. We expect the data to be strong enough to estimate $\boldsymbol{g}_1$ with confidence.

Loosely speaking, a flat prior may be understood as the limit of a Gaussian with variance sent to infinity. Since an unbounded covariance is not well defined the standard approach is to express formulae w.r.t. the precision and explore the limiting case of vanishing dipole precision, that is $\Sigma_1^{-1} \to 0$. In the following, we closely follow Rasmussen & Williams (2006, section 2.7), additionally taking our update system and the linearization into account.

Even though the prior dipole precision is sent to zero, the posterior distribution remains normal. The second step in our update system depends only implicitly on the *a priori* assumptions made, thus, only the first step needs modifications. This means that the overall modelling concept persists, although we have to modify the conditional mean and covariance (eqs 15 and 16).

First of all, we have to partition the covariance matrix into dipole and non-dipole contributions

$$\Sigma_{\mathcal{C}} = \Sigma_{\mathcal{C},\mathrm{DP}} + \Sigma_{\mathcal{C},\mathrm{ND}} = \acute{H}_{\mathcal{C}} \acute{Y}_{1,\mathcal{C}} \Sigma_1 \acute{Y}_{1,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} + \Sigma_{\mathcal{C},\mathrm{ND}}, \tag{50}$$

where $\Sigma_{\mathcal{C},\mathrm{ND}}$ is constructed from $K_{\boldsymbol{B},\mathrm{ND}}$ also containing measurement errors and residuals. To keep equations concise, it is beneficial to predict the dipole coefficients first. Making use of the matrix inversion lemma (Rasmussen & Williams 2006, A.9), conditional mean and inverse of the covariance result in

$$\mathbb{V}[\boldsymbol{g}_1|o_{\mathcal{C}}]^{-1} = \Sigma_1^{-1} + \acute{Y}_{1,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \acute{H}_{\mathcal{C}} \acute{Y}_{1,\mathcal{C}} \tag{51}$$

$$\mathbb{E}[\boldsymbol{g}_1|o_{\mathcal{C}}] = \mathbb{V}[\boldsymbol{g}_1|o_{\mathcal{C}}]^{-1} \left(\Sigma_1^{-1} \bar{\boldsymbol{g}}_1 - \acute{Y}_{1,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \acute{H}_{\mathcal{C}} \tilde{B}_{\mathcal{C}} \right) \tag{52}$$

and the data $o_{\mathcal{C}}$ enters through the POE (see eq. 37). Considering the limit of the uninformative dipole yields

$$\Sigma_{1|\mathcal{C}}^{-1} := \lim_{\Sigma_1^{-1} \to 0} \mathbb{V}[\boldsymbol{g}_1|o_{\mathcal{C}}]^{-1} = \acute{Y}_{1,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \acute{H}_{\mathcal{C}} \acute{Y}_{1,\mathcal{C}} \tag{53}$$

$$\bar{g}_{1|\mathcal{C}} := \lim_{\Sigma_1^{-1} \to 0} \mathbb{E}[\boldsymbol{g}_1|o_{\mathcal{C}}] = -\Sigma_{1|\mathcal{C}}^{-1} \acute{Y}_{1,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \acute{H}_{\mathcal{C}} \tilde{B}_{\mathcal{C}}. \tag{54}$$

Interestingly, $\boldsymbol{g}_1|o_{\mathcal{C}}$ does not depend on $\bar{\boldsymbol{g}}_1$, rendering the dipole's prior mean irrelevant.

Predicting Gauss coefficients of higher SH degree is straightforward since *a priori* dipole and non-dipole contributions are assumed independent. For $\ell > 1$ cross correlations are of no concern as they do not depend on $\boldsymbol{g}_1$. Analogous to eq. (40), conditional mean and covariance result in

$$\mathbb{E}\left[\boldsymbol{g}_{2:\ell}|o_{\mathcal{C}}\right] = -\Sigma_{2:\ell} \acute{Y}_{2:l,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} \Omega_{\mathcal{C}} \acute{H}_{\mathcal{C}} \tilde{B}_{\mathcal{C}} \tag{55}$$

$$\mathbb{V}\left[\boldsymbol{g}_{2:\ell}|o_{\mathcal{C}}\right] = \Sigma_{2:\ell} - \Sigma_{2:\ell} \acute{Y}_{2:l,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} \Omega_{\mathcal{C}} \acute{H}_{\mathcal{C}} \acute{Y}_{2:l,\mathcal{C}} \Sigma_{2:\ell}, \tag{56}$$

where $\Omega_{\mathcal{C}}$ refers to the limiting precision matrix

$$\Omega_{\mathcal{C}} = \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} - \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \acute{H}_{\mathcal{C}} \acute{Y}_{1,\mathcal{C}} \Sigma_{1|\mathcal{C}} \acute{Y}_{1,\mathcal{C}}^{\top} \acute{H}_{\mathcal{C}}^{\top} \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \tag{57}$$

and we again made use of the matrix inversion lemma. To perform the second step, the whole posterior covariance matrix is necessary. However, the posterior cross covariance amongst dipole and non-dipole coefficients is missing. The difficulty is to find the limiting case. According to eq. (20) we have

$$\mathrm{Cov}\left[\boldsymbol{g}_1, \boldsymbol{g}_{2:\ell}|o_{\mathcal{C}}\right] = -\mathrm{Cov}[\boldsymbol{g}_1, O_{\mathcal{C}}]\, \Sigma_{\mathcal{C}}^{-1} \mathrm{Cov}\left[O_{\mathcal{C}}, \boldsymbol{g}_{2:\ell}\right] \tag{58}$$

since *a priori* dipole and non-dipole are assumed independent. Plugging in the linearization we end up with

$$\mathrm{Cov}\left[\boldsymbol{g}_1, \boldsymbol{g}_{2:\ell}|o_{\mathcal{C}}\right] = -\Sigma_{1|\mathcal{C}} \acute{Y}_{1,\mathcal{C}} \acute{H}_{\mathcal{C}} \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \acute{H}_{\mathcal{C}}^{\top} \acute{Y}_{2:l,\mathcal{C}}^{\top} \Sigma_{2:\ell}, \tag{59}$$

which no longer depends on Σ_1. To see this, use the matrix inversion lemma to express the precision matrix, factor in the left hand side and expand by $\Sigma_{1|\mathcal{C}}^{-1} \Sigma_{1|\mathcal{C}}$. Using eq. (51) and re-arranging terms yields eq. (59).

To predict on the EMF, we divide the cross covariance into dipole and non-dipole contributions. Since *a priori* the dipole term is assumed independent of the non-dipole contribution we have

$$\Sigma_{\boldsymbol{B}\mathcal{C}} = \Sigma_{\boldsymbol{B}\mathcal{C},\mathrm{DP}} + \Sigma_{\boldsymbol{B}\mathcal{C},\mathrm{ND}} = \acute{Y}_1^{\top} \Sigma_1 \acute{Y}_{1,\mathcal{C}} \acute{H}_{\mathcal{C}} + \Sigma_{\boldsymbol{B}\mathcal{C},\mathrm{ND}} \tag{60}$$

and $\Sigma_{\boldsymbol{B}\mathcal{C},\mathrm{ND}}$ is constructed from $K_{\boldsymbol{B},\mathrm{ND}}$. Using the same strategy as we did with the Gauss coefficients, the posterior mean and covariance arise to

$$\mathbb{E}[\boldsymbol{B}|o_{\mathcal{C}}] = -\acute{Y}_1^{\top} \bar{g}_{1|\mathcal{C}} + \Sigma_{\boldsymbol{B}\mathcal{C},\mathrm{ND}} \Sigma_{\mathcal{C},\mathrm{ND}}^{-1} \left(\tilde{B}_{\mathcal{C}} - \acute{Y}_{1,\mathcal{C}}^{\top} \bar{g}_{1|\mathcal{C}} \right) \tag{61}$$

$$\mathbb{V}[\boldsymbol{B}|o_{\mathcal{C}}] = \acute{Y}_1^{\top} \Sigma_{1|\mathcal{C}} \acute{Y}_1 + K_{\boldsymbol{B},\mathrm{ND}} - \Sigma_{\boldsymbol{B}\mathcal{C},\mathrm{ND}} \Omega_{\mathcal{C}} \Sigma_{\boldsymbol{B}\mathcal{C},\mathrm{ND}}^{\top}. \tag{62}$$

From those equations we can obtain all quantities needed to proceed with the second step incorporating incomplete records.

The importance in the result is that conditional mean and covariance no longer depend on the choice of the *a priori* dipole, since we assumed zero precision.

4.2 Compound distribution

Although we are not particularly interested in reconstructing the probability distribution of $\vartheta = (\lambda, \rho, \epsilon)$, their variabilities must be taken into account. The final result of the proceeding is the EMF's compound distribution

$$p(\boldsymbol{B}|o) = \int p(\boldsymbol{B}|o, \vartheta)\, p(\vartheta|o)\, \mathrm{d}\vartheta \tag{63}$$

which results from marginalizing the scaling factors. Fig. 3 illustrates the interaction of the two-step strategy and the compound distribution, with the marginalization being depicted in the bottom part. This approach makes it possible to escape from Gaussianity. The integral will not be tractable analytically and must be evaluated by numerical methods.

Let us inspect the two PDFs we want to integrate over. For a certain choice of parameters, $p(\boldsymbol{B}|o, \vartheta)$ is calculated according to our two step strategy. To keep track of the scaling factor's posterior $p(\vartheta|o)$ we add a hierarchical stage. Applying Bayes' law, the posterior density is given by

$$p(\vartheta|o) \propto p(o|\vartheta) p(\vartheta), \tag{64}$$

where we neglected the normalizing constant. To suitably normalize we have to carry out another quadrature since $p(o) =$

$\int p(o|\vartheta)p(\vartheta)\,\mathrm{d}\vartheta$ is unknown. Because of the flat dipole prior, calculating $p(o|\mathfrak{I})$ needs special attention. Not to distract from our endeavour, this is the focus of Section 4.3.

A priori, all three parameters are assumed statistically independent. Both, residual and error level are considered uniformly distributed. As we are roughly aware of magnitudes, the chosen range is well-spaced (weakly informative). On the contrary, λ is a scale parameter bearing across orders of magnitude. The according uninformative prior—representing the state of no prior information—Jeffrey's prior

$$p(\lambda) \propto \frac{1}{\lambda}\,[\mathrm{nT}] \tag{65}$$

that is, values 10 times larger are just as likely as values 10 times smaller (Murphy 2012, section 5.4.2). However, there is a subtlety arising for the compound prior PDFs. The hierarchical approach has an impact because the *a priori* $\boldsymbol{B}$-field depends on λ. The scale invariance is passed on rendering the compound prior distributions improper, that is the density can not be normalized. As an example, the compound PDF for the *a priori* non-dipole Gauss coefficients reads

$$p(g_\ell^m) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\lambda^2} \exp\left\{-\frac{1}{2}\frac{(g_\ell^m)^2}{\lambda^2}\right\} \mathrm{d}\lambda \propto \frac{1}{|g_\ell^m|}\,. \tag{66}$$

That density remains centred at zero whereas the variance does not exist. Nevertheless, the compound posterior PDFs are normalizable and well defined. Marginalizing λ results in an extremely weak *a priori* assumption.

Although the compound distribution is not normal, calculating higher-order moments is one of its potentials. Posterior mean and covariance are given by

$$\mathbb{E}[\boldsymbol{B}|o] = \int p(\vartheta|o)\mathbb{E}[\boldsymbol{B}|o,\vartheta]\;\mathrm{d}\vartheta \tag{67}$$

$$\mathbb{V}[\boldsymbol{B}|o] = \int p(\vartheta|o)\big(\mathbb{E}[\boldsymbol{B}|o,\vartheta]^2 + \mathbb{V}[\boldsymbol{B}|o,\vartheta]\big)\,\mathrm{d}\vartheta - \mathbb{E}[\boldsymbol{B}|o]^2 \tag{68}$$

and the mean and covariance we are integrating over are explicitly given. For modelling Gauss coefficients the above formulae translate analogously.

An actual implementation requires to perform three numeric integrations, in total. Since numeric quadrature in three dimensions is feasible, both the fully Bayesian posterior density and also the Gaussian moment matching proxy are right at hand.

4.3 Marginal likelihood

We postponed the calculation of the *marginal likelihood* until here. The terms *marginal* and *likelihood* refer to the marginalization over the EMF as a function in $\mathfrak{I}$. To actually discretize and integrate the compound distribution, $p(o|\mathfrak{I})$ is still missing. We again separate into complete and incomplete records. The marginal likelihood factorizes

$$p(o|\vartheta) = p(o_\mathcal{I}|o_\mathcal{C},\vartheta)p(o_\mathcal{C}|\vartheta), \tag{69}$$

where we used the conditional probability rule. The major benefit is that we already know the quantities needed to evaluate the PDFs on the right-hand side of eq. (69) through our two step strategy. Fig. 3 illustrates where eq. (69) enters the modelling scheme and how quantities are passed between marginalization and two-step strategy.

For the sake of simplicity we analyse incomplete records, first. For a certain choice of parameters, $\boldsymbol{B}|o_\mathcal{C},\vartheta$ and $O_\mathcal{I}|\vartheta$ are jointly normal due to linearization and error approximation. By taking advantage of the joint normality, we can directly observe the prior predictive distribution of the incomplete records. The PDF follows to read

$$p(o_\mathcal{I}|o_\mathcal{C},\vartheta) = \frac{\exp\left\{-\frac{1}{2}(o_\mathcal{I}-\bar{O}_{\mathcal{I}|\mathcal{C}})^\top\Sigma_{\mathcal{I}|\mathcal{C}}^{-1}(o_\mathcal{I}-\bar{O}_{\mathcal{I}|\mathcal{C}})\right\}}{\sqrt{(2\pi)^{n_\mathcal{I}}|\Sigma_{\mathcal{I}|\mathcal{C}}|}} \tag{70}$$

and the mean and covariance are given by eqs (42) and (44), implicitly depending on λ, ρ and ϵ through the first step in our update system. Regarded as a function in $\mathfrak{I}$, the marginal likelihood is certainly not normal.

Because of the flat dipole prior, the marginal likelihood for complete records needs special attention. We proceed analogously to the incomplete data, however, having to bear the limiting case

$$p(o_\mathcal{C}|\vartheta) = \lim_{\Sigma_1^{-1}\to 0} \frac{\exp\left\{-\frac{1}{2}(o_\mathcal{C}-\bar{O}_\mathcal{C})^\top\Sigma_\mathcal{C}^{-1}(o_\mathcal{C}-\bar{O}_\mathcal{C})\right\}}{\sqrt{(2\pi)^{n_\mathcal{C}}|\Sigma_\mathcal{C}|}}. \tag{71}$$

We closely follow Rasmussen & Williams (2006, section 2.7) and again split into dipole and non-dipole contributions. As already mentioned in Section 4.1, in the limit $\tilde{\boldsymbol{g}}_1$ is irrelevant and we set $\bar{O}_\mathcal{C} = o_\mathcal{C} - \acute{H}_\mathcal{C}\tilde{B}_\mathcal{C}$ (see eq. 37). The limiting precision matrix is of no concern and given by eq. (57). The big concern, however, is the determinant as the dipole variance tends to infinity. According to Rasmussen & Williams (2006, eq. 2.45), the marginal likelihood for the complete records results in

$$p(o_\mathcal{C}|\vartheta) = \frac{\exp\left\{-\frac{1}{2}\tilde{B}_\mathcal{C}^\top\acute{H}_\mathcal{C}^\top\Omega_\mathcal{C}\acute{H}_\mathcal{C}\tilde{B}_\mathcal{C}\right\}}{\sqrt{(2\pi)^{n_\mathcal{C}-3}|\Sigma_{\mathcal{C},\mathrm{ND}}|\,|\Sigma_{1|\mathcal{C}}^{-1}|}} \tag{72}$$

and we already computed all relevant quantities in the first step of our update system. Although not explicitly indicated, the non-dipole covariance, the limiting precision and the dipole covariance depend on $\mathfrak{I}$.

5 APPLICATION

As a practical proof of concept we apply the suggested method to the archeomagnetic and volcanic data offered by the GEOMAGIA50.v3 database (Brown *et al.* 2015). We used all directional and intensity records between 1650 AD and 1750 AD that were included in GEOMAGIA50 version 3.3 in November 2019. To simplify matters, the Earth is assumed a sphere of radius $R_E = 6\,371.2$ km and coordinates are treated as if they were spherical. We estimate the committed error to be less than $\frac{1}{2}$ μT which is small compared to modelling uncertainties.

We use the individual, originally reported error estimates. If uncertainties are not available (*ca.* 8.4 per cent of the data), we assign $\alpha_{95} = 4.5^\circ$ as directional errors and $\sigma_F = 8.25$ μT as intensity errors (Licht *et al.* 2013, section 2.2). Single unpaired declination records are not used.

As our model does not yet account for time dependence we group the data into disjoint bins of 100 yr. The decisive factor for the window length is the number of complete records per bin. Our two step strategy and the linearization are the basic rationale behind. The number of complete records has to determine a reasonable POE for incorporating the incomplete records. Let us assume that at the surface the EMF is dipole dominated (about 90 per cent) for the timespan under consideration. As a rule of thumb, a Taylor expansion performs reasonably well when deviating less than 10 per cent

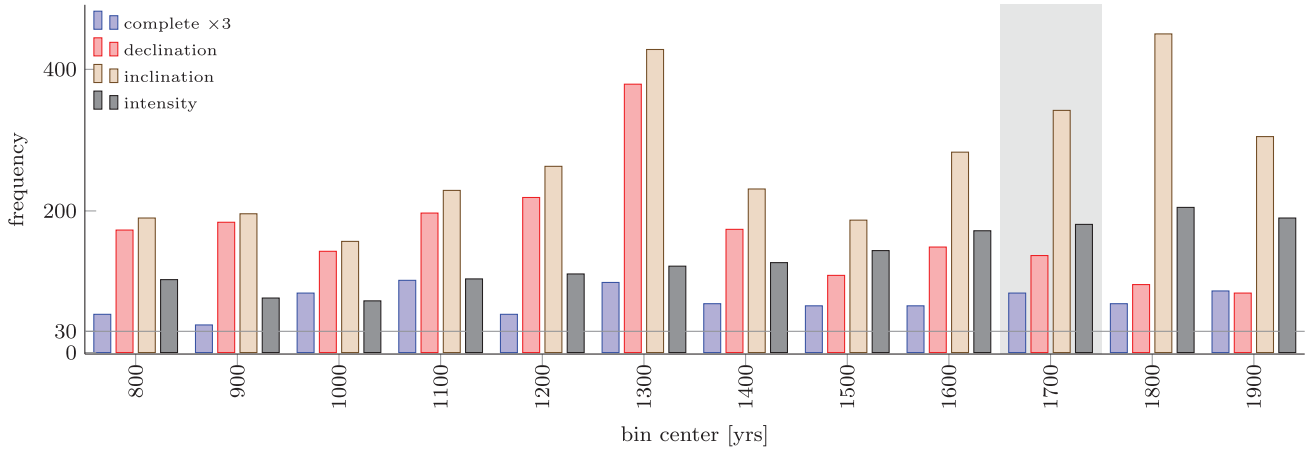

Figure 5. Number of data per 100 yr bin (D, I and F counts are without complete records). For comparability each complete record is counted as three data points. The bin that we use as example in the discussion is shaded in grey. The horizontal line refers to the criterion of how the window length is chosen.

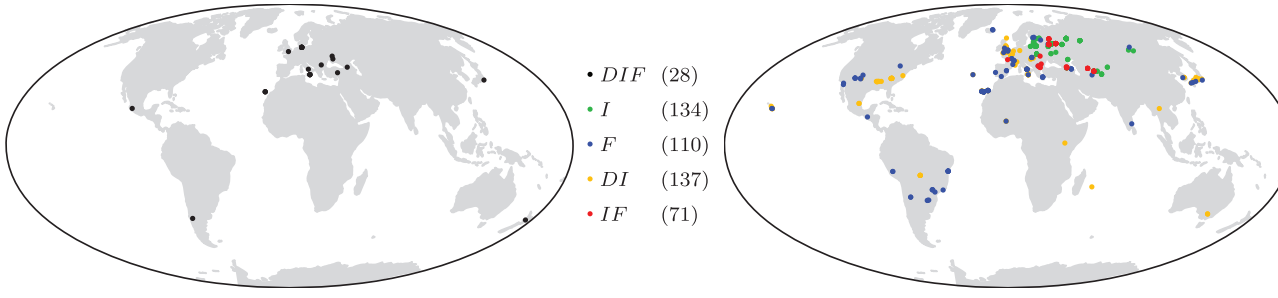

Figure 6. Illustration of the very irregular data coverage for the 1700 epoch. The left-hand panel shows complete records whereas the right-hand panel illustrates combinations of D, I and F measurements.

from the *truth*. Having a parametric view back in mind, the field's dominating features may be described by only nine parameters ($\boldsymbol{g}_1$ and Σ_1). With a minimum of ten complete records per bin we anticipate a coarse field estimate suited as POE. This consideration leads to a window length of 100 yr. Fig. 5 shows the temporal data distribution.

To demonstrate the potential of our modelling strategy in recovering the stationary field we use the interval [1650, 1750] as an example. In total, this bin summarizes 744 observations acquired at 480 sites. Fig. 6 depicts the highly irregular data coverage which is dominated by the northern hemisphere.

We are going to compare our findings to three previous, continuous magnetic field models. The models ARCH10k.1 (Constable *et al.* 2016), arhimag1k (Senftleben 2019) and COV-ARCH (Hellio & Gillet 2018) are considered eligible competitors as they stem from a similar data basis, although arhimag1k additionally includes direct historical observations. All three models report Gauss coefficients up to SH degree $\ell = 10$. ARCH10k.1 and arhimag1k are continuous in time, but do not report modelling errors. In contrast, COV-ARCH provides an ensemble of 50 realizations but uses the same coarse time steps as we do. To compare with COV-ARCH's ensemble we calculate sample mean and sample variance. Histograms are computed according to *Scott's rule of thumb* (Scott 1979).

5.1 Numeric integration

To evaluate the compound distribution $p(\boldsymbol{B}|o)$ we discretise the integral in eq. (63). The posterior PDF is approximated by a simple Riemann sum

$$p(\boldsymbol{B}|o) \approx \sum_i p(\boldsymbol{B}|o, \vartheta_i)\, p(\vartheta_i|o)\Delta\vartheta_i . \tag{73}$$

Although not highly accurate, we favour using a Riemann sum because of its simplicity. Since by construction the mixture components $p(\boldsymbol{B}|o, \vartheta_i)$ are normally distributed, the r.h.s. of eq. (73) gives a (finite) Gaussian mixture distribution (Murphy 2012, section 11.2.1), for which moments are readily available by translating eqs (67) and (68) accordingly.

To actually discretise the parameter space, we use a regular and equally spaced grid. Thus we have the posterior

$$p(\boldsymbol{B}|o) \approx \sum_{i,j,k=0}^{N} p(\boldsymbol{B}|o, \lambda_i, \epsilon_j, \rho_k)\, p(\lambda_i, \epsilon_j, \rho_k|o) \times \Delta\lambda\Delta\epsilon\Delta\rho, \tag{74}$$

where N^3 is the number of gridpoints for the Riemann sum, $\Delta\lambda$ is the step width, λ_0 and λ_N specify the bounds of the interval we integrate over and $\lambda_i = \lambda_0 + i\Delta\lambda$ (Similar for ϵ and ρ). As mentioned above, we carry out a second quadrature to normalize the parameter posterior $p(\lambda_i, \epsilon_j, \rho_k|o)$. This also serves the purpose of having a well defined Gaussian mixture density. We normalize it w.r.t. the Riemann sum, that is such that

$$\sum_{i,j,k=0}^{N} p(\lambda_i, \epsilon_j, \rho_k|o)\Delta\lambda\Delta\epsilon\Delta\rho = 1 . \tag{75}$$

Under these considerations eq. (74) can be seen as the discrete version of eq. (63), where all terms are replaced by their respective discrete equivalent.

We consider the approximation (eq. 73) reasonable, since we found the probability mass of the parameter posterior $p(\vartheta|o)$ to be unimodal and localized in a finite region. Thus to calculate the Gaussian mixture proxy (eq. 74) for the compound distribution, we perform two steps:

5.1.1 *Exploration of the parameter space*

We need to identify the region in which the parameter posterior is localized. Hence we span a coarse grid over all values we believe to be (physically) reasonable and calculate the posterior $p(\vartheta|o)$ on this grid. We choose $N = 25$ gridpoints along each of the three dimensions and choose the bounds as

$$\begin{aligned} \lambda_0 &= 100\,\mathrm{nT} & \lambda_N &= 150\,000\,\mathrm{nT} \\ \epsilon_0 &= 10\,\text{per cent} & \epsilon_N &= 350\,\text{per cent} \\ \rho_0 &= 1000\,\mathrm{nT} & \rho_N &= 6500\,\mathrm{nT}. \end{aligned} \tag{76}$$

Although in principle a wider extent of the grid may be 'physically reasonable', by trial and error we found these bounds to be sufficient (i.e. outside of the bounds the probability mass was approximately zero for all considered cases). Finally, from the posterior we calculate the marginal distribution for each parameter λ, ϵ, ρ via another Riemann sum. We then calculate each (empirical) mean and standard deviation from these coarse marginals. The region where the probability mass is located is then covered by the cuboid centred at the empirical mean with edge-lengths given by two empirical standard deviations.

5.1.2 *Calculation*

The actual numerical integration only takes place within the cuboid that is derived in the exploration step. Inside that cuboid a refined grid is spanned with $N = 15$ nodes per dimension. By calculating not only the posterior on this new grid, but running for each gridpoint an inversion for the EMF, the Gauss coefficients and other quantities of interest, we can calculate a proxy for the full compound posterior for each of these quantities, using eq. (74). This proxy, which is a Gaussian mixture distribution, is the final result of our modelling strategy.

For all epochs the parameter posterior is of good-nature only featuring a single mode. For epoch 1700, Fig. 7 depicts the posterior parameter PDF. Presented are all combinations of 2-D marginal parameter posterior at the coarse and refined grid. The white rectangle refers to the edges of the refined grid that is used for numeric integration. The maximum of the error level ϵ is slightly above 100 per cent at a rather small spread. Taking the approximations into account, a shift toward higher values seems reasonable. The residual term ρ has its maximum at about 4 μT and is also relatively sharp. It is interesting to note that the residual is of the same order of magnitude as the equatorial dipole. This need not only be due to unexplained sources, it can also be an effect of the linearization or not taking time dependencies into account. The scaling parameter λ features the widest distribution. In comparison with the IGRF, the magnitude and range appear reasonable. However, the posterior PDF of the model parameters is not sharp enough to justify a point estimate. Therefore we integrate out model parameters and thus the uncertainties from the parameter posteriors are translated into the posterior of the quantities of interest. This way the posterior variance, which is easily available due to the Gaussian mixture structure, does not only reflect uncertainties arising from the data, but also reproduces the model uncertainties.

5.2 Vector field predictions

The discretized versions of $\boldsymbol{B}$'s posterior mean and variance translate according to eqs (67) and (68) and are given by

$$\mathbb{E}[\boldsymbol{B}|o] = \sum p(\vartheta_i|o)\mathbb{E}[\boldsymbol{B}|o,\vartheta_i]\,\Delta\vartheta \tag{77}$$

$$\begin{aligned} \mathbb{V}[\boldsymbol{B}|o] = \sum p(\vartheta_i|o)\left(\mathbb{E}[\boldsymbol{B}|o,\vartheta_i]^2 + \mathbb{V}[\boldsymbol{B}|o,\vartheta_i]\right)\Delta\vartheta \\ -\mathbb{E}[\boldsymbol{B}|o]^2\,. \end{aligned} \tag{78}$$

The pointwise posterior standard deviation serves for realistic location dependent uncertainty estimates. Again, $\boldsymbol{B}$-field predictions are non-parametric and do not depend on Gauss coefficients.

The top row of Fig. 8 depicts both posterior mean and standard deviation of the down component at the Earth's surface. The field is evaluated at 2000 design points, equally distributed over the sphere (Deserno 2004). We are able to quantify what previous studies are suggesting: The EMF is reconstructed with confidence within areas of dense data coverage, for example in Europe. Structures of large parts of the southern hemisphere, however, remain vague. We see quite similar patterns across all epochs under investigation.

5.3 Declination, inclination and intensity

Because the EMF shares a non-linear relation with declination, inclination and intensity we do not see the possibility of analytically deriving properties such as the posterior mean for D, I and F. For a moderate number of design points we can obtain the distribution by sampling strategies. However, trying to draw from a high dimensional normal mixture distribution appears absurd. For the 2000 design points we are interested in, storing all the covariance matrices of the entire parameter grid is hardly possible, let alone drawing samples from the resulting mixture.

The idea behind uncovering the posterior mean and variance of D, I and F is once more a linearization. To do so the same strategy as pointed out in Section 2.3 is used. The linearization of D, I and F are given by eqs (25), (26) and (27). We build upon the EMF estimate from the previous section and the posterior mean $\mathbb{E}[\boldsymbol{B}|\boldsymbol{o}]$ serves as POE. Because of the linearization approximations for mean and (co)variance are explicit.

To provide an example we consider the intensity, only. Utilizing eq. (27), the approximate posterior mean and variance are given by

$$\mathbb{E}[F|o] \approx \|\,\mathbb{E}[\boldsymbol{B}|\boldsymbol{o}]\,\| \tag{79}$$

$$\mathbb{V}[F|o] \approx \frac{\mathbb{E}[\boldsymbol{B}|\boldsymbol{o}]^\top \mathbb{V}[\boldsymbol{B}|\boldsymbol{o}]\mathbb{E}[\boldsymbol{B}|\boldsymbol{o}]}{\|\,\mathbb{E}[\boldsymbol{B}|\boldsymbol{o}]\|^2}. \tag{80}$$

We pursue a non-parametric approach but the concept is similar to that of Hellio *et al.* (2014, appendix A). For the 1700 epoch, the bottom row of Fig. 8 depicts mean and standard deviation of the intensity at the Earth's surface. Although not directly observed by a single record, the evolving area of weak field known as the South Atlantic Anomaly (e.g. Mandea *et al.* 2007; Hartmann & Pacca 2009) is certainly significant within one standard deviation. Looking at the other epochs, the westward drift is also visible. This feature is constrained here by only a few points around the actual anomaly, which shows that our modelling approach is capable of uncovering features that are known from models with stronger data

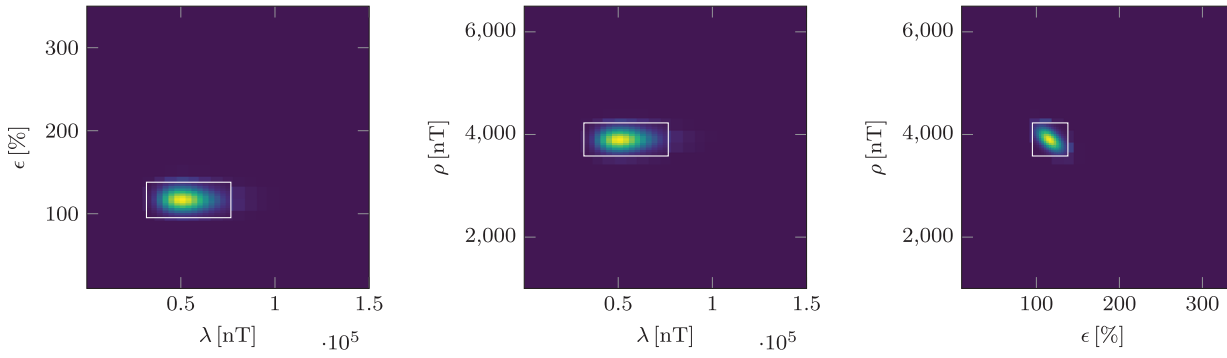

Figure 7. Discretization of the parameter space for the epoch 1700. The accurately sampled marginal densities for the integration phase (white framed box) are drawn atop of the coarse exploratory grid.

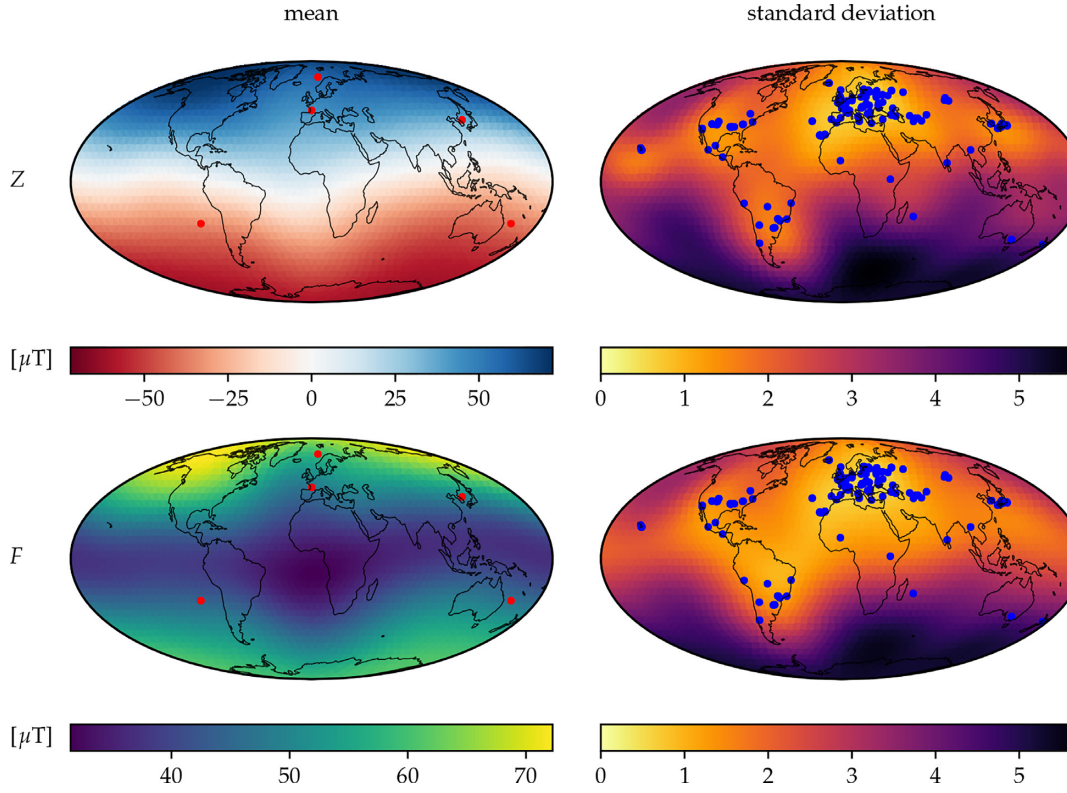

Figure 8. Posterior mean (left-hand panel) and standard deviation (right-hand panel) of the EMF at the Earth's surface for the 1700 epoch. The top row shows the down component B_Z. The locations at which we perform normality tests are indicated in red (see Table 1). The bottom row shows the field intensity F. Blue dots refer to the underlying data sites.

basis. This ability stems from long ranging spatial correlations of the kernel and is controlled by the reference radius R (see Fig. 1).

When comparing epoch 800 with fig. 11(a) in Hellio & Gillet (2018) two things stand out. The magnitudes are similar but a low intensity patch is found in the Pacific rather than the Atlantic. In the standard deviation we also see a rather sharp transition between north and south. Since this is a proof of concept a detailed comparison or even an interpretation would be premature.

Since our posterior distributions are certainly not normal it is questionable if the standard deviation is adequate to describe uncertainties. If the posterior is highly skewed or even multimodal, the standard deviation would not be well suited to quantify modelling errors. One would preferably use percentiles which, however, cannot be derived analytically. As long as the number of design points is moderate, one brute-force method is to calculate percentiles by sampling. This works well for percentiles that are not far from the bulk of the probability mass.

A simplistic approach to obtain samples from a mixture distribution is the following algorithm. First, by chance the kth random variable is selected from the mixing distribution $p(\vartheta_k|o)$ (categorical). Then the value of the selected random variable $p(\boldsymbol{B}|o, \vartheta_k)$ is realized (multivariate normal). Repeat until the desired amount of samples is achieved.

Although we have to radically reduce the number of design points we can gain insight into how strong the proxy standard deviation and percentiles deviate. We compare 16- and 84-percentiles to mean $\pm$ standard deviation via

$$\Delta_{16} = 1 - \frac{\mu - \alpha_{16}}{\sigma} \quad \text{and} \quad \Delta_{84} = 1 - \frac{\alpha_{84} - \mu}{\sigma}, \tag{81}$$

Table 1. Estimation of the quality of the proxy standard deviation as a measure of error, compared to the 16-/84-percentiles. If Δ is larger (smaller) than 0, the standard deviation overestimates (underestimates) the error. Values are given in percent of one standard deviation.

Lat	−26°	45°	71°	39°	−26°
Lon	−88°	0°	9°	131°	159°
D Δ_{16}	2.5	0.7	0.3	−0.2	1.5
Δ_{84}	−0.3	2.9	3.9	1.8	0.4
I Δ_{16}	3.6	−0.3	−0.4	−1.5	3.8
Δ_{84}	−3.0	1.0	3.0	6.7	−2.3
F Δ_{16}	4.6	−2.4	0.7	1.5	7.1
Δ_{84}	−1.7	0.5	−0.1	−0.3	0.6
Z Δ_{16}	2.9	−2.6	−1.1	−0.4	1.2
Δ_{84}	1.6	0.1	0.6	0.6	1.9

where α refers to percentiles, μ and σ indicate mean and standard deviation. The interpretation of Δ is as follows: If Δ is larger (smaller) than zero, the standard deviation overestimates (underestimates) the uncertainty given by the percentile, by $|\Delta|$ standard deviations. Note, that to keep this interpretation consistent for both Δ_{16} and Δ_{84} the sign of the second term in (81) changes.

At 5 randomly selected locations we apply this check to the down component B_Z and to the three commonly used archeomagnetic observables D, I and F. As can be seen in Table 1, at all locations the deviation is well below 10 per cent. Thus we believe the proxy standard deviation is qualified to describe uncertainties. Furthermore, the proxy standard deviation obtained by linearization is computationally feasible and easy to visualize and interpret.

5.4 Predictions at the CMB

Until here predictions were carried out at the Earth's surface. It is straightforward to predict the EMF at arbitrarily chosen design points outside of the reference sphere of radius R. Fig. 9 presents mean and standard deviation of the down component B_Z at the CMB. We again use 2000 equidistributed design points but at radius $R_{\mathrm{CMB}} = 3480\,\mathrm{km}$. One observes that uncertainties are greater than at the Earth's surface. Roughly speaking, the standard deviation is about 40 per cent compared to the mean, while at the Earth's surface relative errors only amount up to 10 per cent (see right-hand column of Fig. 8). The modelling error strongly depends on how far design points are from the reference sphere.

This behaviour is best understood thinking in terms of Gauss coefficients. In eq. (2) the term that is raised to the power of $\ell + 1$ causes this effect. Since design points lie outside the reference sphere, the ratio $\frac{R}{|x|}$ is smaller than one. Thus, Gauss coefficients are more *penalized* the larger the SH degree, and the ratio $\frac{R}{|x|}$ determines the rate of decline. In turn, the closer to the reference radius the more impact higher SH degrees have. Related to smaller structures, the higher the SH degree the less certain we are. When scaling down—for example to the CMB—higher SH degrees increase and so do the respective uncertainties. This idea translates to our non-parametric model.

Compared to what Hellio & Gillet (2018, fig. 8c) found, our reconstruction of epoch 1400 looks very similar. Because of the large variabilities, a detailed comparison does not make much sense.

5.5 Gauss coefficients

We want to stress again that our model is inherently non-parametric. The fundamental quantity of the inference is the geomagnetic potential. Nevertheless, our approach allows to infer Gauss coefficients since geomagnetic potential and Gauss coefficients are linearly related (see eq. 17). The procedure is similar to inferring the EMF. The discretized versions of the compound mean (eq. 67) and (co)variance (eq. 68) for the Gauss coefficients g_ℓ^m read

$$\mathbb{E}\left[g_\ell^m|o\right] = \sum p(\vartheta_i|o)\mathbb{E}\left[g_\ell^m|o,\vartheta_i\right]\Delta\vartheta \tag{82}$$

$$\mathbb{V}\left[g_\ell^m|o\right] = \sum p(\vartheta_i|o)\left(\mathbb{E}\left[g_\ell^m|o,\vartheta_i\right]^2 + \mathbb{V}\left[g_\ell^m|o,\vartheta_i\right]\right)\Delta\vartheta - \mathbb{E}\left[g_\ell^m|o\right]^2. \tag{83}$$

In principle we can predict up to arbitrary SH degree. However, storing the component's mean and covariances for the whole parameter grid becomes memory intense. If we restrict our selves to a moderate SH degree—for example $\lesssim 10$—we are able to calculate the full mixture PDF, sample from the posterior and calculate percentiles.

Although not explicitly indicated, Gauss coefficients depend on the reference radius R. Given the g_ℓ^m at the reference radius, we can scale them to a radius $\tilde{R} > R$ by

$$\tilde{g}_\ell^m = g_\ell^m\left(\frac{R}{\tilde{R}}\right)^{\ell+2}. \tag{84}$$

Mean and covariances translate accordingly.

For comparison, we scale Gauss coefficients w.r.t. the Earth's surface. The mixture PDFs for the dipole coefficients are shown in Fig. 10. As can be seen, the mixture is quite close to the moment matching normal proxy. Across all epochs the dipole coefficients are close to being normally distributed, except for the bin [850, 950], for which the mixture distribution deviates slightly from the normal proxy. However, the higher the SH degree the less normal the posterior is and the more the scale invariant prior dominates (see eq. 66). We further see that for the epoch under consideration, the histogram built from the COV-ARCH ensemble is in good agreement with the mixture and the mean aligns with the ARCH10k.1 estimate. It is not too much of a surprise that the arhimag1k prediction deviates since it additionally incorporates historical records, which are not included in our data set.

Although well suited to compare with existing models, care has to be taken when interpreting Gauss coefficients. A finite set of Gauss coefficients does not represent the full information contained in our non-parametric modelling approach. Areas of dense data coverage may feature a resolution than can not be captured by an expansion, for example until degree 10. However, this is less relevant in the context of geomagentic core field modelling, where SH degrees larger than around 14 are dominated by lithospheric field signals, which cannot be resolved by a sparse data distribution. Fig. 11 compares the predicted Gauss coefficients to the selected reference models until SH degree 5, which is considered the approximate global resolution of the spherical harmonics based models (Korte *et al.* 2009; Licht *et al.* 2013; Constable *et al.* 2016; Sanchez *et al.* 2016; Hellio & Gillet 2018). With few exceptions, our findings are on a par with existing models within one standard deviation.

5.6 Spatial Power Spectrum

It is hard to digest all the information contained in a collection of Gauss coefficients. Therefore, it has become common to consider the geomagnetic power spectrum (Backus *et al.* 1996, section 4.4.2),

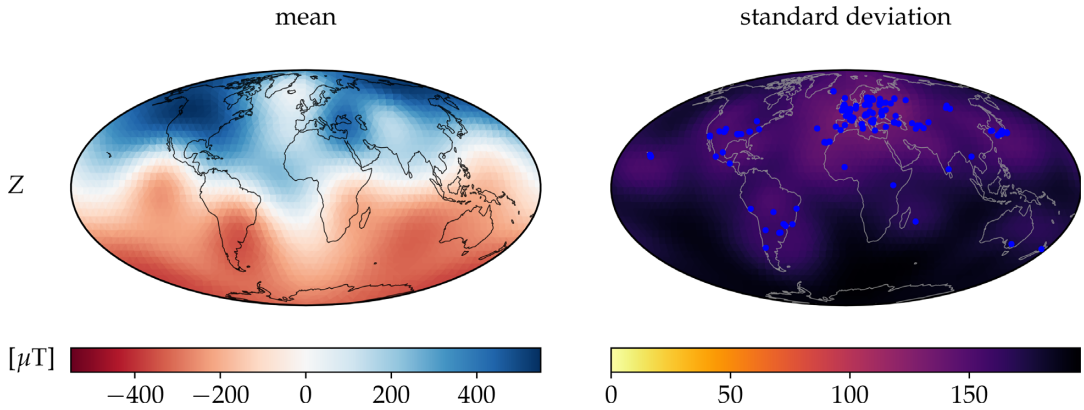

Figure 9. Posterior mean (left-hand panel) and standard deviation (right-hand panel) of the EMF down component B_Z at the CMB for the 1700 epoch. Blue dots refer to the sites where records are acquired.

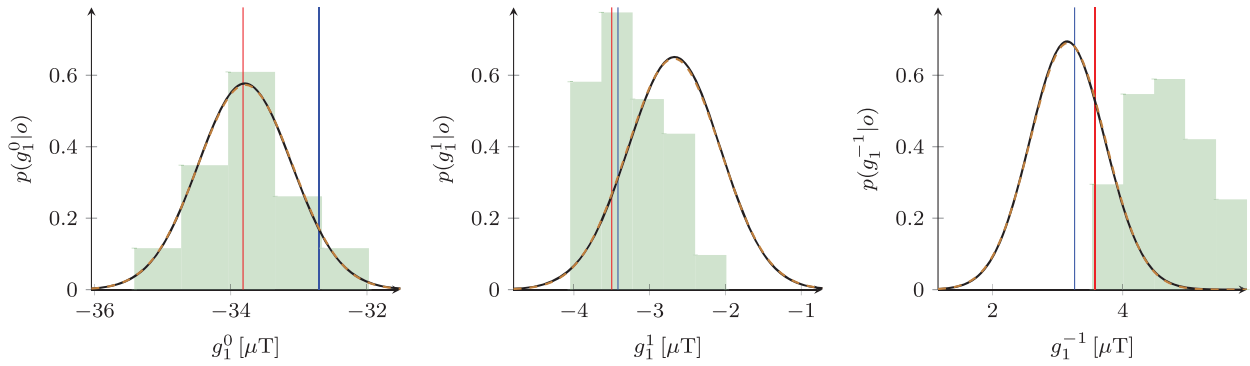

Figure 10. Distribution of the dipole coefficients at the Earth's surface for epoch 1700. The resulting mixture PDF is indicated by the solid black lines whereas the, nearly identical, dashed lines refers to the moment matching Gaussian proxies. The green histograms illustrates the COV-ARCH ensemble coefficients. The blue vertical lines refer to arhimag1k and the red ones to ARCH10k.1 .

that reflects the contributions of different spatial wavelengths in the SHs. For degree ℓ the corresponding wavelength is $\lambda_\ell \approx \frac{4\pi R}{2\ell+1}$ (Langel & Hinze 1998, section 4.3.5). The EMF is apportioned such that

$$\boldsymbol{B} = \sum_\ell \boldsymbol{B}_\ell = -\nabla \sum_\ell \Phi_\ell, \tag{85}$$

where

$$\Phi_\ell(\boldsymbol{x}) = R\left(\frac{R}{|\boldsymbol{x}|}\right)^{\ell+1} \sum_{-\ell\le m\le \ell} g_\ell^m\, Y_\ell^m(\hat{\boldsymbol{x}})\,. \tag{86}$$

The components Φ_ℓ are certainly orthogonal, since SHs form an orthogonal system. A characteristic that describes $\boldsymbol{B}_\ell$'s variations is the so-called *average square value*. Since $\boldsymbol{B}$ is divergence free, the net total flux through a closed surface is zero. Thus, the *average* over the sphere vanishes

$$\langle \boldsymbol{B}_\ell \rangle_R = 0\,, \tag{87}$$

where the angle brackets are an abbreviation for the surface integral (see eq. 17). The definition of the average can be extended to the (centred) average square value

$$\left\langle (\boldsymbol{B}_\ell - \langle \boldsymbol{B}_\ell \rangle_R)^2 \right\rangle_R = \left\langle \boldsymbol{B}_\ell^2 \right\rangle_R . \tag{88}$$

Due to Parseval's Theorem (Backus *et al.* 1996, eq. 4.4.21), the average square can be expressed in terms of Gauss coefficients. Regarded as a function in ℓ, the quantity

$$r_\ell := \left\langle B_\ell^2 \right\rangle_R = (\ell+1) \sum_{-\ell\le m\le \ell} (g_\ell^m)^2 \tag{89}$$

is called the *geomagnetic power spectrum* (Lowes 1974). Again, Gauss coefficients g_ℓ^m depend on the reference radius and so does r_ℓ.

Within the setting of statistical inversions, calculating the power spectrum requires special attention since uncertainties have an appreciable effect. Squaring and summing normal distributed Gauss coefficients gives a random variable (RV) that is distributed as the sum of weighted non-central χ^2 RVs. Unfortunately, for the PDF of a linear combination of non-central chi-square RVs no closed, analytic expression is known (Bausch 2013). Nonetheless, using the algebraic formula for the variance $\mathbb{E}[XY] = \mathrm{Cov}[X, Y] + \mathbb{E}[X]\,\mathbb{E}[Y]$ we obtain an expression for the expectation of the power spectrum

$$\mathbb{E}[r_\ell|o] = (\ell+1) \sum_m \left(\mathbb{E}\big[g_\ell^m|o\big]^2 + \mathbb{V}\big[g_\ell^m|o\big] \right) \tag{90}$$

and it is obvious that variances play an important role. In other words, the larger the uncertainties the bigger the impact on the spectrum's mean.

Although the second moment may be accessible, the standard deviation is not suited to quantify errors. Roughly speaking, if standard deviations of the g_ℓ^ms dominate over the mean, the PDF of r_ℓ is highly skewed with wide tails. In turn, a moment matching Gaussian proxy would violate the positivity constraint of r_ℓ. For an increasing SH degree this is certainly the case. Percentiles are better suited to estimate the error level, but impossible to access analytically. We again calculate the distribution and percentiles empirically by brute-force sampling Gauss coefficients from the posterior. In Section 5.3, the sampling strategy is described.

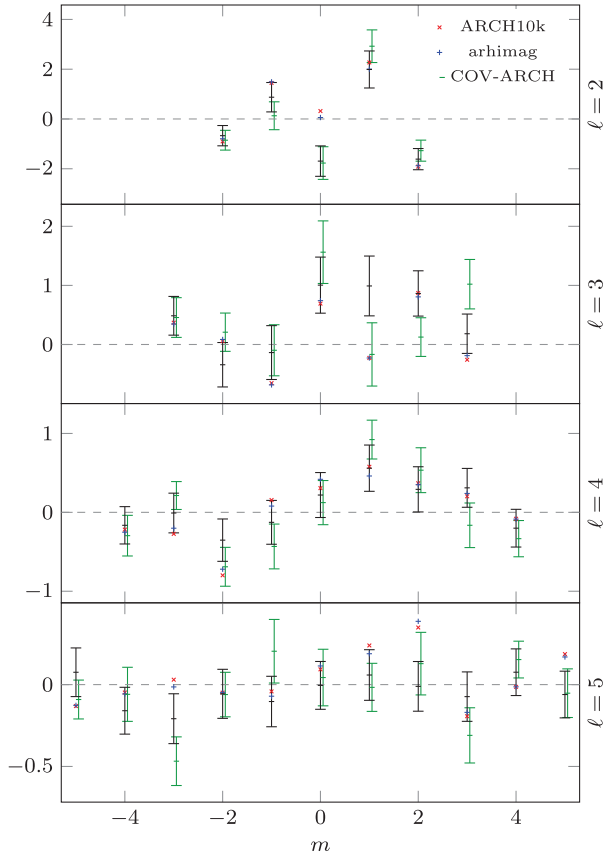

Figure 11. Comparison of Gauss coefficients for SH degrees 2–5 at Earth's surface for epoch 1700. Results from this study are shown in black. Error bars indicate one standard deviation.

For epoch 1700 the resulting power spectrum is shown in Fig. 12. Within the error margins our findings and the COV-ARCH model are in good agreement. Differences arise comparing with ARCH10k.1 and arhimag1k. Both models report less power at degree $\ell = 3$ and feature a rapid loss for degrees $\ell \geq 8$. While the degree 3 deviation might be due to differences in the underlying data basis, the latter likely is caused by the influence of the global regularization in the spherical harmonic models. Nonetheless, for $\ell \leq 8$ ARCH10k.1 and arhimag1k are potential realizations from what we find.

5.7 Dipole Moment

To a first approximation, the EMF is dipolar. This is, its shape is similar to that of a hypothetical bar magnet placed at the centre of the Earth. With respect to a Cartesian coordinate frame, the corresponding vector dipole moment is given by

$$\boldsymbol{m} = \frac{4\pi R_E^3}{\mu_0}(g_1^1\hat{\boldsymbol{x}} + g_1^{-1}\hat{\boldsymbol{y}} + g_1^0\hat{\boldsymbol{z}}) \,, \tag{91}$$

where $\mu_0 \approx 4\pi \times 10^{-7}\,[\mathrm{Tm\,A^{-1}}]$ refers to the permeability of free space (Backus *et al.* 1996, eq. 4.4.17). Vector components are indicated by subscripts x, y and z, for example $m_x \propto g_1^1$. Considering the discretization, the vector dipole moment $\boldsymbol{m}$ is a Gaussian mixture borrowing its statistical properties from the dipole coefficients. The

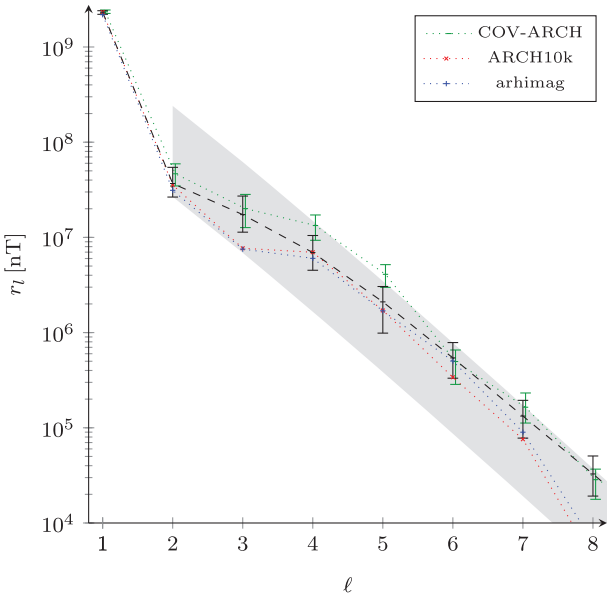

Figure 12. Power spectrum at the Earth' surface for epoch 1700. The mean of the posterior power spectrum is indicated by the dashed line. The top and bottom of the error bars refer to 16 and 84 percentiles, respectively. Except for a constant offset, the grey shaded area indicates the slope of the *a priori* power spectrum.

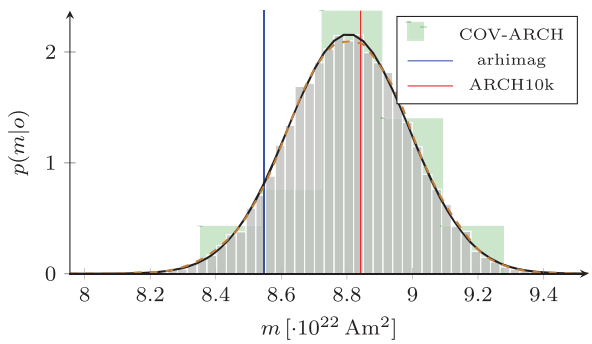

Figure 13. Distribution of the dipole moment for epoch 1700 based on 10 000 drawn samples. The kernel density estimate is indicated by the red dashed line, that is nearly identical to the Gaussian proxy (black line).

dipole's magnitude can be computed directly via

$$\tau = \|\boldsymbol{m}\| = \frac{4\pi R_E^3}{\mu_0}\sqrt{(g_1^1)^2 + (g_1^{-1})^2 + (g_1^0)^2} \tag{92}$$

and is proportional to the square root of the power r_1. We are interested in statistical properties of τ. Because of the square root, we are not able to derive an analytic expression for $\mathbb{E}[\tau|o]$. Nonetheless, if we knew $\mathbb{E}[\tau|o]$, the variance is right at hand

$$\mathbb{V}[\tau|o] = \mathbb{E}\left[\tau^2|o\right] - \mathbb{E}[\tau|o]^2 \,, \tag{93}$$

since $\mathbb{E}\left[\tau^2\right] \propto \mathbb{E}[r_1]$, which is given by eq. (90). To obtain a proxy of $\mathbb{E}[\tau|o]$ and to calculate the empiric distribution we again use sampling. An ordinary Gaussian kernel density estimate is used to smoothen the histogram (Murphy 2012, section 14.7.2). The bandwidth is selected by *Scott's rule of thumb* which—due to its simplicity—strongly influences the estimate.

For the epoch of choice, Fig. 13 compares the density estimate, the according histogram and the Gaussian proxy. The density estimate looks rather normal and is approximated well by the moment matching Gaussian. This is also the case for all other epochs

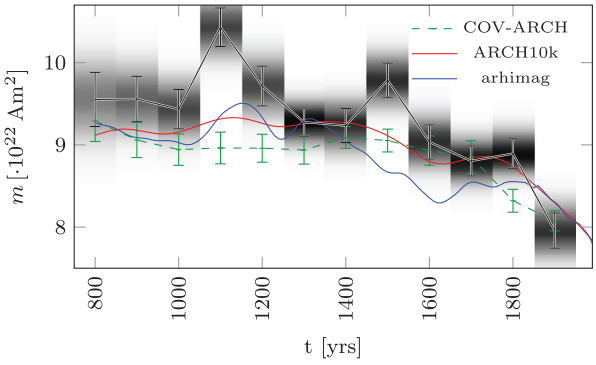

Figure 14. Temporal variation of the dipole moment magnitude based on 10 000 drawn samples. The posterior means are connected by black lines and the error bars indicate one standard deviation. The gray shaded background refers to the kernel density estimate. Results from models COV-ARCH, ARCH10k and arhimag are shown for comparison.

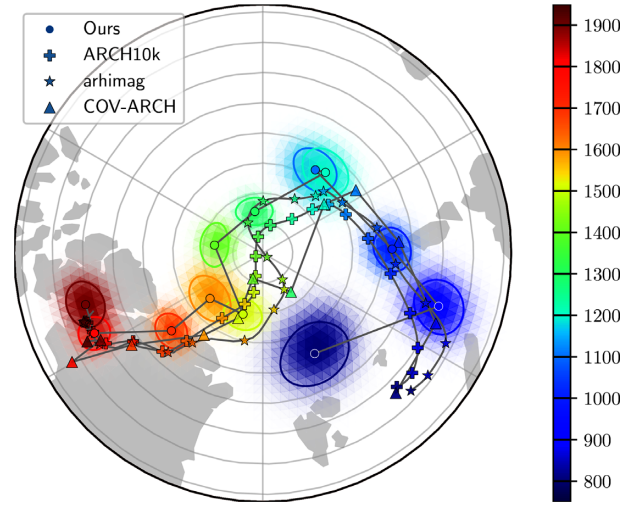

Figure 15. Wander of the geomagnetic north pole. For each epoch in the study, the probability distribution of the location is shown with the mean marked by dots and one sigma marked by an ellipse. The colour varies according to time. For comparison, wander paths for the reference models are shown, without uncertainties in order not to clutter the figure.

in our study. At least for the epoch 1700 our findings agree with ARCH10k.1 and COV-ARCH. Presumably, due to its stronger data basis arhimag1k deviates. However, the agreement varies through epochs.

As the dipole moment is known to change with time, Fig. 14 displays the time-series of all epochs under consideration. Compared to COV-ARCH, ARCH10k.1 and arhimag1k, we find a similar temporal evolution of the dipole moment. Presumably the most recent epoch deviates from the other models, since we do not include historical information in our model. Especially for the earlier epochs we see slightly higher intensities than reported by existing models. The strong deviation of epochs 1100 and 1500 may be caused by outliers, which we did not test for so far, but it is also possible that the variations are real and earlier models underestimate strong variations due to the treatment of dating uncertainties and temporal regularization. It is beyond the scope of our proof-of-concept model to resolve this question. For both epochs the deviation is caused by g_1^0. The parameter distributions $p(\vartheta|o)$ for these epochs do not show any noticeable problems. However, epoch 1500 features a rather weak data basis.

We are further interested in the distribution of the dipole's north pole a.k.a. *geomagnetic north*. The geomagnetic north pole is given as the antipode of the projection of the vector dipole moment onto the sphere. In other words, the intersection of the axis of the hypothetical bar magnet with the Earth's surface. The location w.r.t. spherical coordinates is given by

$$\theta_m = \arccos\left(-\frac{m_z}{\tau}\right), \quad \phi_m = \arctan\left(\frac{-m_y}{-m_x}\right), \tag{94}$$

where θ refers to colatitude and ϕ to longitude. We proceed with a similar approach as presented in Khokhlov *et al.* (2006) and translate the dipole's PDF that is interpreted w.r.t. a Cartesian reference frame into spherical coordinates. The PDF of the vector dipole moment transformed to spherical coordinates is given by

$$p\begin{pmatrix}\tau\\ \theta_m\\ \phi_m\end{pmatrix}\Big|o\Big) = p\begin{pmatrix}m_x = \sin\theta_m\cos\phi_m\\ m_y = \sin\theta_m\sin\phi_m\\ m_z = \cos\theta_m\end{pmatrix}\Big|o\Big)\,\tau^2\sin\theta_m, \tag{95}$$

where we made use of the change of variables theorem (Murphy 2012, Section 2.6). To obtain the distribution of the location we have to marginalize the magnitude from the mixture distribution

$$p\begin{pmatrix}\theta_m\\ \phi_m\end{pmatrix}\Big|o\Big) = \sum_i p(\vartheta_i|o)\int_0^\infty p\begin{pmatrix}\tau\\ \theta_m\\ \phi_m\end{pmatrix}\Big|o,\vartheta_i\Big)\,\mathrm{d}\tau. \tag{96}$$

Since individual dipole coefficients are normally distributed, we can analytically solve the integral. To do so, factor out m, complete the square and marginalize via the standard Gaussian integral equations (Owen 1980, eqs 10, 11 and 12). We deliberately skip the resulting expression as it is lengthy and of no particular interest.

However, to visualize the wander of geomagnetic north it is useful to numerically evaluate the resulting PDF. As can be seen in Fig. 15 the general movement is similar to the models we are comparing with. Again the 1500 epoch deviates and is ahead of the comparison path, which may be caused by outliers as stated previously. For comparison we only show COV-ARCH's sample mean because a scatter plot of the complete ensemble leads to an overstuffed picture. The ensemble covariance features similar magnitudes compared to what we find.

6 CONCLUSIONS AND PERSPECTIVES

The extensive theory of Sections 2–4 build the foundation of a new modelling strategy for archeo- and palaeomagnetic field models. The key advantage of this probabilistic approach is that realistic modelling uncertainty estimates are obtained, for example, via the standard deviation. The *a priori* distributions we choose have pros and cons. Taking the least subjective choice is an advantage as we *a priori* do not specify any preferred direction and, thus, our method is even well suited for time periods featuring reversals. However, a shortcoming—in particular concerning the prior dipole—is that we could not visualize a comparison of prior and posterior uncertainties.

Besides the *a priori* covariance structure and the weakly informative parameter priors, our modelling strategy depends only on a single parameter, the reference radius R. Conceptually, it is no problem to also integrate out this remaining parameter. Only a technicality arises if we want to integrate out R, Gauss coefficients require

scaling to a common radius—for example the Earth's surface—as they depend on R. The limiting factor is a 4-D parameter space with excessive computational demands.

To save the effort of implementing the exhaustive theory, a ready-to-use software suite called CORBASS (Schanner & Mauerberger 2019) was developed as part of the project. CORBASS is written in python and licensed under the GPLv3. A public GIT repository serves for development, maintenance and support. To facilitate first steps we provide usage examples in the form of a web based interactive environment (Jupyter notebooks), that further illustrate the modelling concept and the algorithm. To lighten system requirements we make use of the package and environment manager *conda*.

We carried out a case study to demonstrate the potential of our statistical model. All results presented in Section 5 are produced using CORBASS. Even though we use a rather small data set, computational costs are not negligible. At the time of writing, processing all epochs under consideration took about 30 hr on an ordinary workstation. The computational complexity is set by the number of observations and the parameter grid chosen. For a certain choice of model parameters, the computational complexity of GP regression is cubic with respect to number of observations. Performing a Riemann sum along one dimension scales quadratically according to the number of collocation points. Thus the complexity of the quadrature scheme for all three model parameters grows with the sixth power. Although numeric integration offers room for optimizations, the computational complexity cannot be lightened.

Although a proof of concept rather than a fully featured EMF model, our case study already supports the findings of existing studies. In comparison with models using traditional methods this is useful since it is another source that quantifies what was described qualitatively. As an example, early studies questioned the reconstruction of the EMF in the southern hemisphere from archeomagnetic and volcanic data only due to poor data coverage (e.g. Korte *et al.* 2009; Constable & Korte 2015). That statement is quantified by our findings: The EMF's posterior standard deviation is small in areas of good data coverage, such as Europe, while uncertainties are large on the southern hemisphere (see top panel of Fig. 8). Even though our uninformative prior assumptions are significantly weaker, we find uncertainties similar to for example Licht *et al.* (2013), while noting that bootstrap ensemble methods tend to underestimate uncertainties in regions where there are no data to draw from. Surprisingly, our uncertainty estimates are on a par with more elaborate modelling concepts (e.g. Hellio & Gillet 2018) whereas our approach does not yet account for the temporal evolution. Future work will show if this is a coincidence, and if taking temporal dynamics into account yields different results.

In a general context, caution is advised when performing any further processing with posterior Gauss coefficients. Whenever possible, use the posterior EMF instead. At densely covered regions a vast amount of Gauss coefficients are necessary to represent all the information that is contained in the posterior EMF at the surface. However, this is irrelevant when studying the core field which cannot be retrieved beyond degrees around 14 anyway, due to the distance from the source and dominance of the lithospheric field at higher degrees.

The spatial correlation structure we employ makes truncating the SH decomposition obsolete and the model resolves according to the availability of data. This raises the obvious question of how our global, non-parametric model compares to higher resolving studies. To investigate this question one would have to include other sources of data such as historical logs and observatory data for recent times. Although the modelling strategy inherently works with magnetic field components B_N, B_E and B_Z we deliberately left out recent observatory data to put attention on non-linear observables. If we focus on areas of small modelling errors, precise locations must not be neglected. Considering elevation and the coordinate conversion from geodetic to geocentric is straightforward. The difficulty with historical logs is that the direction of travel is affected by large inaccuracies (Jackson *et al.* 2000; Jonkers *et al.* 2003). To account for location uncertainties, our data model requires an extension. Another difficulty is computational costs related to the large amount of historical and observatory data. Therefore one has to introduce a data selection and reduction process as the interest is in time spans of years but not days or less. Furthermore, our proxy Gaussian error model is intolerant of outliers. With only a few records that strongly deviate, the Gaussian error model causes a highly distorted reconstruction. Therefore, it is important to perform outlier analysis and select data with care. The work by Khokhlov *et al.* (2006) appears to be well suited to discriminate data that are incompatible with our modelling approach.

For the time increments of interest—$\Delta t \geq 1$ yr—we know with confidence when observatory data and historical records were acquired. For volcanic and archeomagnetic records, the average dating uncertainties amount to about a hundred years (Licht *et al.* 2013, section 2.2), and they tend to increase further back in time. For this study, we assume that our rather long bin width of 100 yr balances temporal errors, although it results in poor temporal resolution. Arguably, our model is still overly confident as we did not consider dating errors. In order to apply our method more generally and to longer times the inclusion of sediment records has to be implemented. Sediment records are affected by large dating uncertainties which require a data model that also accounts for temporal errors and preserves the stratification (e.g. Nilsson *et al.* 2014). Moreover, a strategy to deal with the scaling or relative intensity has to be developed and our two-step approach might become a challenge if the number of available full vector records with absolute intensity information is small.

Nonetheless, the presented snapshot model should be considered a first step in the direction of a time continuous correlation based Holocene magnetic field model, and, more generally, a new modelling method for the palaeomagnetic field on various timescales. We regard the time stationary binning an interim solution as it does not capture the dynamics of the EMF well. In our opinion a temporally continuous model also considering dating errors is needed to fulfil the needs of palaeomagnetic field modelling. For the extension of our modelling concept we are currently working on an empirical continuous time correlation kernel, similar to Bouligand *et al.* (2016); Hellio & Gillet (2018). Combining a temporal and a spatial kernel to a space–time kernel will make any binning obsolete. Within this space–time GP setting, existing techniques may be used to address dating uncertainties (McHutchon & Rasmussen 2011).

We conclude by emphasizing again that this is initial work towards a new palaeomagnetic field modelling strategy and an improved full Holocene model, and that it is destined to receive many improvements in the future. The open source modelling concept offers vast flexibility and a allows for a variety of refinements.

ACKNOWLEDGEMENTS

This work would not have been possible without the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), grant 388291411. The conceptional and theoretical work

was conducted by S. Mauerberger and M. Schanner, with significant assistance by M. Holschneider. M. Korte accounted for data selection as well as interpretation and embedding of the case study. S. Mauerberger prepared the manuscript with support of all authors. Software development and data processing were performed by M. Schanner with major contributions by S. Mauerberger. M. Korte and M. Holschneider supervised the findings of this work.

Special thanks go to H. Matuschek for providing FieldTools (Matuschek & Mauerberger 2019), assistance and support. We also thank two anonymous Reviewers who contributed to improving the manuscript. Last but not least we thank C. Heckman for final proof reading.

REFERENCES

Backus, G., Parker, R. & Constable, C.G., 1996. *Foundations of Geomagnetism,* Cambridge Univ. Press.

Bausch, J., 2013. On the efficient calculation of a linear combination of chi-square random variables with an application in counting string vacua, *J. Phys. A,* **46**(50), 505202.

Bouligand, C., Gillet, N., Jault, D., Schaeffer, N., Fournier, A. & Aubert, J., 2016. Frequency spectrum of the geomagnetic field harmonic coefficients from dynamo simulations, *Geophys. J. Int.,* **207,** 1142–1157.

Bouligand, C., Hulot, G., Khokhlov, A. & Glatzmaier, G., 2005. Statistical palaeomagnetic field modeling and dynamo numerical simulation, *Geophys. J. Int.,* **161,** 603–626.

Brown, M.C. *et al.*, 2015. Geomagia50.v3: 2. a new palaeomagnetic database for lake and marine sediments, *Earth, Planets Space,* **67**(1), 70.

Constable, C.G., Johnson, C.L. & Lund, S.P., 2000. Global geomagnetic field models for the past 3000 years: transient or permanent flux lobes? *Phil. Trans. R. Soc. Lond., A,* **358,** 991–1008.

Constable, C.G. & Korte, M., 2015. Centennial- to millennial-scale geomagnetic field variations. In *Treatise on Geophysics,* 2nd ed., Vol. **5,** pp. 309–341, ed, Schubert , G., Elsevier.

Constable, C.G., Korte, M. & Panovska, S., 2016. Persistent high palaeosecular variation activity in southern hemisphere for at least 10 000 years, *Earth planet. Sci. Lett.,* **453,** 78–86.

Constable, C.G. & Parker, R.L., 1988. Statistics of the geomagnetic secular variation for the past 5 m.y., *J. geophys. Res.,* **93**(B10), 11569–11581.

Deserno, M., 2004, How to generate equidistributed points on the surface of a sphere, p. 99, ed., Polymerforshung, P.-If.

Farquharson, C.G. & Oldenburg, D.W., 1998. Non-linear inversion using general measures of data misfit and model structure, *Geophys. J. Int.,* **134**(1), 213–227.

Gillet, N., 2019. Spatial and temporal changes of the geomagnetic field: insights from forward and inverse core field models. In *Geomagnetism, Aeronomy and Space Weather: A Journey From the Earth'S Core to the Sun,* pp 115–132, Mandea M., Korte M., Yau A, eds, Cambridge Univ. Press.

Gubbins, D. & Bloxham, J., 1985. Geomagnetic field analysis–III. Magnetic fields on the core-mantle boundary, *Geophys. J. R. astr. Soc.,* **80**(3), 695–713.

Hartmann, G.A. & Pacca, I.G., 2009. Time evolution of the south atlantic magnetic anomaly, *Anais da Academia Brasileira de Ciências,* **81**(2), 243–255.

Hellio, G. & Gillet, N., 2018. Time-correlation-based regression of the geomagnetic field from archeological and sediment records, *Geophys. J. Int.,* **214**(3), 1585–1607.

Hellio, G., Gillet, N., Bouligand, C. & Jault, D., 2014. Stochastic modelling of regional archaeomagnetic series, *Geophys. J. Int.,* **199,** 931–943.

Holschneider, M., Lesur, V., Mauerberger, S. & Baerenzung, J., 2016. Correlation-based modeling and separation of geomagnetic field components, *J. geophys. Res.,* **121**(5), 3142–3160.

Jackson, A., Jonkers, A. & Walker, M., 2000. Four centuries of geomagnetic secular variation from historical records, *Phil. Trans. R. Soc. Lond., A,* **358**(1768), 957–990.

Jonkers, A.R.T., Jackson, A. & Murray, A., 2003. Four centuries of geomagnetic data from historical records, *Rev. Geophys.,* **41**(2), doi:10.1029/2002RG000115.

Khokhlov, A., Hulot, G. & Bouligand, C., 2006. Testing statistical palaeomagnetic field models against directional data affected by measurement errors, *Geophys. J. Int.,* **167**(2), 635–648.

Korte, M. & Constable, C.G., 2003. Continuous global geomagnetic field models for the past 3000 years, *Phys. Earth planet. Inter.,* **140,** 73–89.

Korte, M., Donadini, F. & Constable, C.G., 2009. Geomagnetic field for 0-3ka: 2. A new series of time-varying global models, *Geochem. Geophys. Geosys.,* **10,** Q06008, doi:10.1029/2008GC002297.

Langel, R. & Hinze, W., 1998. *The Magnetic Field of the Earth's Lithosphere: The Satellite Perspective,* Cambridge Univ. Press.

Licht, A., Hulot, G., Gallet, Y. & Thébault, E., 2013. Ensembles of low degree archeomagnetic field models for the past three millennia, *Phys. Earth planet. Inter.,* **224,** 38–67.

Livermore, P.W., Fournier, A., Gallet, Y. & Bodin, T., 2018. Transdimensional inference of archeomagnetic intensity change, *Geophys. J. Int.,* **215**(3), 2008–2034.

Love, J.J. & Constable, C.G., 2003. Gaussian statistics for palaeomagnetic vectors, *Geophys. J. Int.,* **152**(3), 515–565.

Lowes, F.J., 1974. Spatial power spectrum of the main geomagnetic field, and extrapolation to the core, *Geophys. J. Int.,* **36**(3), 717–730.

Mandea, M., Korte, M., Mozzoni, D. & Kotzé, P., 2007. The magnetic field changing over the Southern African continent - a unique behaviour, *S. Afr. J. Geol.,* **110,** 193–202.

Matuschek, H. & Mauerberger, S., 2019. FieldTools: Toolbox for manipulating vector fields on the sphere, V. 1.0. GFZ Data Services, doi:10.5880/fidgeo.2019.033

McHutchon, A. & Rasmussen, C.E., 2011. Gaussian process training with input noise, in *Advances in Neural Information Processing Systems 24,* pp. 1341–1349, eds, Shawe-Taylor, J., Zemel, R.S., Bartlett, P.L., Pereira, F. & Weinberger, K.Q., Curran Associates, Inc.

Murphy, K.P., 2012. *Machine Learning: A Probabilistic Perspective,* The MIT Press.

Nilsson, A., Holme, R., Korte, M., Suttie, N. & Hill, M., 2014. Reconstructing Holocene geomagnetic field variation: new methods, models and implications, *Geophys. J. Int.,* **198**(1), 229–248.

Owen, D.B., 1980. A table of normal integrals, *Commun. Stat. - Simulat. Comput.,* **9**(4), 389–419.

Rasmussen, C. & Williams, C., 2006. *Gaussian Processes for Machine Learning,* MIT Press.

Sanchez, S., Fournier, A., Aubert, J., Cosme, E. & Gallet, Y., 2016. Modelling the archaeomagnetic field under spatial constraints from dynamo simulations: a resolution analysis, *Geophys. J. Int.,* **207**(2), 983–1002.

Schanner, M. & Mauerberger, S., 2019. CORBASS: CORrelation Based Archeomagnetic SnapShot model v.1.1, Potsdam: GFZ Data Services. doi:10.5880/GFZ.2.3.2019.008.

Scott, D.W., 1979. On optimal and data-based histograms, *Biometrika,* **66**(3), 605–610.

Senftleben, R., 2019, Earth's magnetic field over the last 1000 years. *PhD thesis,* University of Potsdam, Germany.

Suttie, N. & Nilsson, A., 2019. Archaeomagnetic data: the propagation of an error, *Phys. Earth planet. Inter.,* **289,** 73–74.

Thébault, E., *et al.*, 2015. International geomagnetic reference field: the 12th generation, *Earth, Planets Space,* **67**(1), 79.

Walker, M. & Jackson, A., 2008. Robust modeling of the Earth's magnetic field, *Geophys. J. Int.,* **143,** 799–808.