



HAL
open science

Time-correlation-based regression of the geomagnetic field from archeological and sediment records

G. Hellio, N. Gillet

► **To cite this version:**

G. Hellio, N. Gillet. Time-correlation-based regression of the geomagnetic field from archeological and sediment records. *Geophysical Journal International*, 2018, 214, pp.1585-1607. 10.1093/gji/ggy214 . insu-03663863

HAL Id: insu-03663863

<https://insu.hal.science/insu-03663863>

Submitted on 10 May 2022

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

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

Time-correlation-based regression of the geomagnetic field from archeological and sediment records

G. Hellio^{1,2} and N. Gillet¹

¹Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTerre, 38000 Grenoble, France. E-mail: gabrielle.hellio@univ-nantes.fr

²Univ. de Nantes, Laboratoire de Planétologie et Géodynamique, CNRS UMR 6112, 44322 Nantes cedex 3, France

Accepted 2018 May 30. Received 2018 May 4; in original form 2017 December 26

SUMMARY

We propose two ensembles of geomagnetic field models spanning the last three millennia: COV-ARCH is calculated using all available archeological artefacts and volcanic lava flows; lake and marine sediment records are added to this data set to build COV-LAKE. Given the sparse distribution of archeomagnetic observations and their associated large uncertainties, the recovery of magnetic field changes from such data is an ill-posed inverse problem that requires assuming some prior information. This is usually performed by imposing arbitrary regularizations in space and time. Instead, we construct the prior knowledge entering the objective function to be minimized from spatial and temporal statistics of the geomagnetic field, as available from satellites, ground-based observatories and paleomagnetic measurements, and validated by numerical geodynamo simulations. Our approach relies on the projection of model coefficients onto temporal cross-covariance functions. We show with synthetic experiments that the dispersion within the ensemble of solutions provides a coherent measure of the model uncertainties. Gauss coefficients inverted from geophysical records compare satisfactorily with those deduced from the independent database built upon historical and observatory records. *A posteriori* model errors are reduced when incorporating sediment records; they logically increase towards decreasing length-scales, indicating that a partial information is available up to a spherical degree 4–5. Such models and their associated uncertainties are suited to be used as observations in geomagnetic data assimilation studies. Our results advocate for an approximately constant dipole decay since ≈ 1700 AD, preceded by an era (going back to 1000 AD) where the dipole trend is weak, possibly slightly positive. We observe in both hemispheres, at both low- and high-latitude, persistent patches over the past 3000 yr. We also confirm a westward drift of flux patches at the core–mantle boundary at a speed of about 0.20 to $0.25^\circ \text{ yr}^{-1}$. Despite the sparse data distribution in the southern hemisphere, the South Atlantic Anomaly appears in both ensembles of models around 1800 AD. A similar low-intensity event seems to have appeared below the Indian Ocean over 600–1400 AD. Both global models are in general good agreement with regional master curves, though filtering out some of the centennial oscillations.

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

1 INTRODUCTION

The recent era of satellite magnetic data, associated with the high temporal resolution of ground-based observatories, provide accurate knowledge of the magnetic field behaviour over the past decades. Combined with historical records, these contemporary data constrain the evolution of the large length-scales geomagnetic field for the past few centuries (Jackson *et al.* 2000). Several structures observed in the present field, such as the South Atlantic Anomaly, the axial dipole decay (recorded since about 1840 and the advent of

absolute measurements), or the westward drift captured in historical records of declination (and with flux patches isolated in magnetic models at the core surface, see Jackson & Finlay 2007, for a review), need to be studied over a longer time-span in order to strengthen the constraints on the core dynamics.

Having in mind the geomagnetic data assimilation applications (Fournier *et al.* 2010), one needs to characterize the field evolution over timescales significantly longer than the turn-over time in the Earth's liquid core – of the order of 200 yr. The only such available data are paleomagnetic records. Over the past millennia, they are

usually separated in two categories: (i) archeomagnetic measurements, obtained from heated archeological artefacts and volcanic lava flows, and (ii) lake and marine sediment records. Although considered less reliable, data from the latter group have better spatial and temporal coverage than the archeomagnetic data and therefore bring valuable information to improve the resolution of global models (i.e. spatiotemporal interpolations of surface observations).

Already several of these models exist. To our knowledge, there are four comparable, up-to-date, reference model families built from archeomagnetic and lava data: the ARCHxk series (Korte *et al.* 2009; Constable *et al.* 2016), the ensemble of AF_M realizations (Licht *et al.* 2013), SHA.DIF.14k (Pavón-Carrasco *et al.* 2014) and the AmR ensemble of models by Sanchez *et al.* (2016). We also know four families of Holocene models incorporating sediment data: the CALSxk series (Korte & Constable 2011; Constable *et al.* 2016), the HFM (Holocene Field Model) series (Panovska *et al.* 2015), the ensemble of ASDI.FM models (Licht *et al.* 2013) and pfm9k (Nilsson *et al.* 2014). All of them cover at least the past three millennia. Except for AmR (a series of snapshots estimates, built using a dynamo norm as prior information), they all rely on the same regularization procedure, first introduced in Bloxham & Jackson (1992): (i) a spatial norm minimizing the Ohmic dissipation (Gubbins 1975) at the core–mantle boundary (CMB), and (ii) a temporal norm minimizing the second time derivative of the radial field at the CMB. They all use spherical harmonic functions for the spatial description of the field, and a projection onto cubic B-splines for the temporal interpolation. If the ensemble of models by Licht *et al.* (2013) and Sanchez *et al.* (2016) are estimated up to spherical harmonic degree $l = 5$, all others models are calculated until $l = 10$. To the exception of the dynamo norm approach by Sanchez *et al.* (2016), the main differences mostly concern the way to handle the data and their uncertainties (bootstrap methods, error rescaling, etc.), not so much the inversion methodology.

If considering field coefficients as observations in assimilations tools, one needs realistic estimates of the global model uncertainties, which are hindered by damping techniques used to reduce the non-uniqueness issue. This motivates this study, where we aim at deriving ensembles of field models over the past three millennia using realistic priors, in order to produce reasonable posterior uncertainties. We follow a stochastic avenue, inspired from that first introduced by Gillet *et al.* (2013) for the construction of the COV-OBS field model from observatory data. We furthermore bring modifications to the usual algorithms by projecting Gauss coefficients series onto specifically chosen temporal cross-covariance functions.

We present here two ensembles of spherical harmonic models, COV-ARCH and COV-LAKE, spanning the past three millennia and truncated at degree 10. In the next section, we first present the data considered to build these models (Section 2.1), the way we handle their associated dating uncertainties (Section 2.2), the Gaussian process regression method we use and our choice of *a priori* covariances on the Gauss coefficients (Sections 2.3–2.4). Then in Section 3.1 we apply our algorithm to the inversion of synthetic data sets, in order to assess its reliability. Field models coefficients series and spatiotemporal structures derived from geophysical records are presented in Section 3.2. In a discussion section, we address issues about the azimuthal drift of field patches at the core surface (Section 4.1), low-intensity anomalies at the Earth's surface (Section 4.2), a comparison of global models predictions with regional master curves (Section 4.3) and the possibility of using our models and associated uncertainties as observations in geomagnetic data assimilation re-analyses (Section 4.4). This paper ends with concluding remarks (Section 5).

2 DATA AND METHOD

2.1 Archeomagnetic records and sediment series

The data sets used in this study consist of all archeomagnetic and volcanic data from 1200 BC included in the GEOMAGIA V3 database (Brown *et al.* 2015) by March 2017, and the sediment data set used by Panovska *et al.* (2015). We include additional archeological intensity data from a recent study in Greece (Genevey *et al.* 2017) and from a second one in France (Genevey *et al.* 2016). The latter data cover the early Middle Age, an era previously poorly sampled. Finally, to illustrate the necessity of an increasing number of data, we added the Mavora lake sediment core from New Zealand covering the full Holocene (Turner *et al.* 2015) that was not included yet in the sediment database.

Our archeomagnetic and volcanic data set includes 11 274 data, among which there are 4633 inclinations (I), 3217 declinations (D) and 3424 intensities (F). It represents approximately 800 and 1300 more data than in the data set used in the construction of, respectively, the A_FM (Licht *et al.* 2013) and CALS3k.4 (Korte & Constable 2011) field models, for the same studied time-span.

We follow Licht *et al.* (2013) in their attempt at recognizing the efforts accomplished by paleomagnetists to obtain accurate records, and choose to keep the errors published in the data set when provided. Some of the archeomagnetic and volcanic data are nevertheless published without error estimates (about 6 per cent of the intensities and 13 per cent of the directions present in the data set). In such a case, we assign $\alpha_{95} = 4.5^\circ$ for directional errors and $\sigma_F = 8.25\mu T$ for intensity errors. We then convert α_{95} values in errors on declination and inclination measurements using (Piper 1989):

$$\sigma_I = \frac{81}{140}\alpha_{95}, \quad \sigma_D = \frac{81}{140 \cos I}\alpha_{95}. \quad (1)$$

The temporal and spatial distributions of the data are very uneven, with fewer archeomagnetic and volcanic data back in time, and a much poorer coverage of all kinds of records in the southern hemisphere (see Fig. 1).

The sediment data include 73 sediment cores. It represents 27 084 data divided in 4710 relative paleointensities, 12 101 inclinations and 10 273 relative declinations. The spatial distribution is somewhat more homogeneous between the northern and southern hemispheres than for archeomagnetic data (see Fig. 1) and the temporal distribution is very homogeneous along the considered period. We follow the estimations of Panovska *et al.* (2012) of the sediment records uncertainties except for the more recent Mavora lake for which we assigned the uncertainties provided by the authors.

It is important to recall here that paleointensities and paleodeclinations are relative records that need to be calibrated when used for global modelling. For the Mavora lake records, we use the declination provided in the study as absolute records and use the calibration ratio provided by the authors to calibrate the relative paleointensities. All other records were already calibrated, with the declination set to zero-mean, from Constable *et al.* (2016). They apply the following methodology: (i) for each core, the predictions of a global model (ARCH10k.1) at the locations and times of the records are estimated, (ii) the median value over the considered time-span of the ratio between these predictions and the relative paleointensities are calculated and (iii) all data points within one sediment core are multiplied by this ratio. The obtained series are used as intensity and declination data when inverting for the field model.

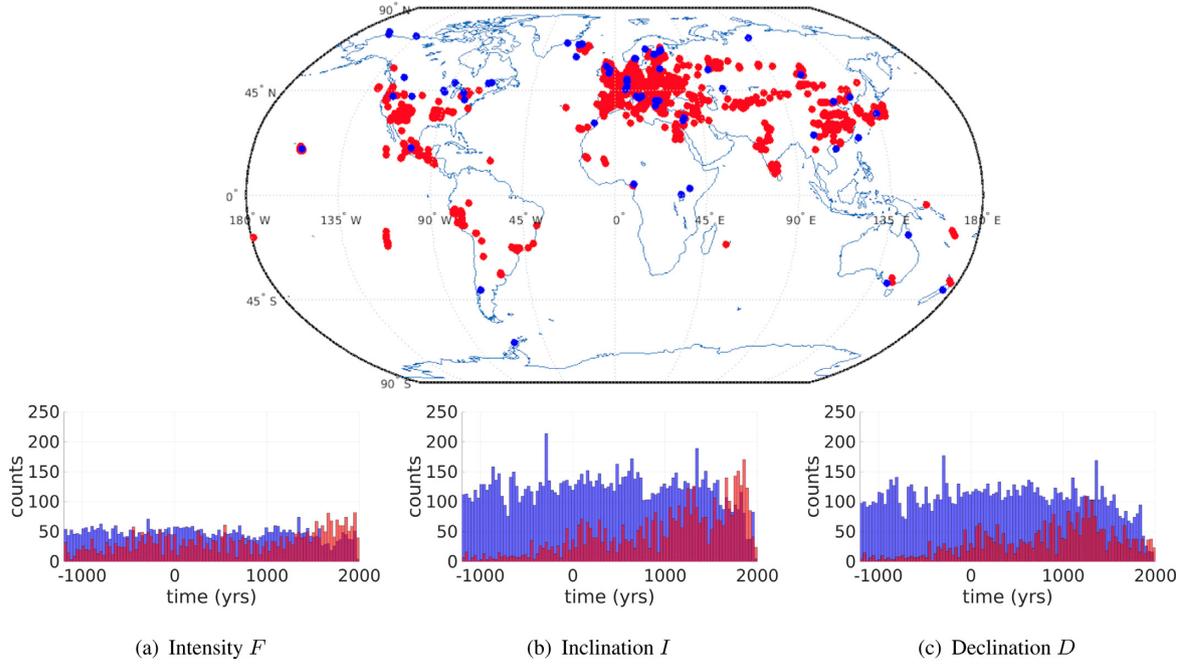


Figure 1. Geographical and temporal distributions of the archeomagnetic and volcanic records (red) and of the sediment cores (blue) used in this study.

2.2 Accounting for dating uncertainties

Age uncertainties associated with archeomagnetic data are important, and sometimes not provided in the database (for 13 per cent of the records in the data set). We follow Licht *et al.* (2013) and assign to data with missing dating errors an uncertainty 1.5 times the mean age uncertainty for every period of 500 yr. The resulting age uncertainties for the several epochs are respectively 160, 130, 120, 100, 138, 114 and 57 yr for periods spanning 1500–1000 BC, 1000–500 BC, 500 BC–0, 0–500 AD, 500–1000 AD, 1000–1500 AD and 1500–2000 AD.

We account for age uncertainties associated with archeomagnetic data using a bootstrap method on dating errors only, as opposed to Korte *et al.* (2009) or Licht *et al.* (2013), who simultaneously perform the bootstrap on both age and magnetic (D , I , F) values. Indeed, within our modelling strategy (see Section 2.3) the projection of measurement errors onto posterior model uncertainties is accounted for through the posterior model error covariance matrix. Age bootstrap consists in randomly drawing the date associated to each record following the provided age distribution, considered uniform in this study. We perform 50 independent bootstraps (enough to guarantee the convergence of the data misfit, ensemble average model and model dispersion). These provide 50 data sets, to be used with the Gaussian process regression described in the next section.

Sediment data require a more careful methodology because of the need to keep the stratigraphic order for all samples of one core. The method we follow differs from the one used in previous studies (e.g. Korte *et al.* 2009; Licht *et al.* 2013), where all ages from one core are shifted by the same timing error, estimated as a Gaussian random error drawn within the probability density function (PDF) for the date uncertainty of the eldest sample within this core. Instead, inspired by Nilsson *et al.* (2014), we stretch (instead of shift) the dates of each series. We draw a Gaussian random error (within the PDF for the dates uncertainties) for both the youngest and the eldest data of each core. Let's write $\langle t_j \rangle_{j \in [1, J]}$, the expectations of the ages provided with a sediment series (for a core containing J data), and t_1^* and t_J^* one random draw for the two ends of the series. For each

draw, all ages t_j^* of the core are obtained as

$$\forall j \in [1, J], t_j^* = t_1^* + \frac{t_j - t_1}{\langle t_j \rangle - \langle t_1 \rangle} (t_j^* - t_1^*). \quad (2)$$

This approach allows to increase or decrease the duration of each core while keeping the stratigraphic order. As for archeomagnetic data, we performed 50 bootstraps. Note that for each bootstrap, the same stretching of the dates is considered for all field components (D , I and F) within one sediment core.

2.3 Gaussian process regression for a nonlinear problem

Away from magnetic sources, the magnetic field \mathbf{B} derives from a scalar potential V , or $\mathbf{B} = -\nabla V$. For an internal source, V can be decomposed in spherical coordinates (r , θ , ϕ) as

$$V(r, \theta, \phi, t) = a \sum_{l=1}^{\infty} \left(\frac{a}{r}\right)^{l+1} \times \sum_{m=0}^l (g_l^m(t) \cos(m\phi) + h_l^m(t) \sin(m\phi)) P_l^m(\cos\theta), \quad (3)$$

with a the Earth's radius, l and m the degree and order of the spherical harmonics, respectively, g_l^m and h_l^m the Gauss coefficients and P_l^m the associated (Schmidt semi-normalized) Legendre polynomials. At a given epoch t , Gauss coefficients are stored in a vector $\mathbf{m}(t)$, of size $L(L+2)$ with $L = 10$, the spherical harmonics truncation degree. We consider N discrete epochs $[t_1 \dots t_N]$ to build a vector $\mathbf{x} = [\mathbf{m}_1(t_1) \dots \mathbf{m}_N(t_N)]$ of size $NL(L+2)$. In practice we sample the period spanning from 1200 BC to 2000 AD every 100 yr (or $N = 33$). Comparison with shorter sampling rates show no significant changes to the average model or the model dispersion, in link with the time resolution resulting from the data frequency and the dating errors.

We define the cross-covariance matrix $\mathbf{C}_{\mathbf{xx}} = E[\mathbf{xx}^T]$, fed with

$$E\left[g_l^m(t_j) g_{l'}^{m'}(t_{j'})\right] = \delta_{ll'} \delta_{mm'} C_l(t_j - t_{j'}). \quad (4)$$

We suppose here that Gauss coefficients are *a priori* independent one to the other, and that the covariance function $C_l(\tau)$ is independent of the order. C_l is further described in Section 2.4. We consider our variables in the framework of Gaussian processes (Rasmussen & Williams 2006), which are by definition entirely characterized by their mean and second-order statistics (covariances). Suppose we know an estimate of the discrete representation \mathbf{x} , entatched with uncertainties $\delta\mathbf{x}$; its associated error covariance matrix is $\mathbf{R}_{\mathbf{xx}} = E[\delta\mathbf{x}\delta\mathbf{x}^T]$. An estimate $\mathbf{m}(t)$ of Gauss coefficients at any epoch t is then given by

$$\mathbf{m}(t) = \mathbf{C}_{\mathbf{mx}}(t) [\mathbf{C}_{\mathbf{xx}} + \mathbf{R}_{\mathbf{xx}}]^{-1} \mathbf{x}, \quad (5)$$

where $\mathbf{C}_{\mathbf{mx}}(t) = E[\mathbf{m}(t)\mathbf{x}^T]$.

Archeomagnetic and sediment data and their associated errors are stored, respectively, in vectors \mathbf{y} and \mathbf{e} . They are nonlinearly related to Gauss coefficients (e.g. Bloxham *et al.* 1989). Each magnetic record y_k , recorded at epoch t_k , is then linked to the model vector through

$$y_k = h_k(\mathbf{m}(t_k)) + e_k = H_k(\mathbf{x}) + e_k. \quad (6)$$

Accordingly, we link \mathbf{x} and \mathbf{y} through a nonlinear observation operator \mathbf{H} (built from the H_k):

$$\mathbf{y} = \mathbf{H}(\mathbf{x}) + \mathbf{e}. \quad (7)$$

We follow a Bayesian approach to retrieve the model \mathbf{x} given observations \mathbf{y} . In a nonlinear case such as the one presented here, we need to proceed iteratively. This is done using a Gauss–Newton algorithm starting from an axial dipole background model $\bar{\mathbf{g}}_1^0$ (see Section 2.4). We checked that alternative initial conditions bring no significant changes to the ensemble average model or the model dispersion. At each iteration i , one has

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{K}_i(\mathbf{y} - \mathbf{H}(\mathbf{x}_i)), \quad (8)$$

with the Kalman gain matrix

$$\mathbf{K}_i = \mathbf{C}_{\mathbf{xx}} \nabla \mathbf{H}_i^T [\nabla \mathbf{H}_i \mathbf{C}_{\mathbf{xx}} \nabla \mathbf{H}_i^T + \mathbf{C}_{\mathbf{ee}}]^{-1}. \quad (9)$$

$\nabla \mathbf{H}_i$ is the Jacobian of \mathbf{H} in \mathbf{x}_i , and $\mathbf{C}_{\mathbf{ee}} = E(\mathbf{ee}^T)$ is the (diagonal) data error covariance matrix. To reduce the impact of outliers, we consider two ways that are confronted in the synthetic study (Section 3.1): either use an L2 measure of residuals $r_k(\mathbf{x}) = y_k - H_k(\mathbf{x})$ while rejecting data for which $|r_k| > 3\sigma_k$, or implement the Huber norm with an iteratively re-weighted least-squares algorithm (Farquharson & Oldenburg 1998). In this latter case $\mathbf{C}_{\mathbf{ee}}$ becomes a function of \mathbf{x} , with diagonal elements

$$\mathbf{C}_{\mathbf{ee}kk}(\mathbf{x}_i) = \tilde{\sigma}_k^2(\mathbf{x}_i) = \begin{cases} \sigma_k^2, & |r_k(\mathbf{x}_i)| < c \\ \frac{\sigma_k |r_k(\mathbf{x}_i)|}{c}, & |r_k(\mathbf{x}_i)| \geq c \end{cases}, \quad (10)$$

where $c = 1.5$ (e.g. Walker & Jackson 2000; Olsen 2002). The Huber measure thus comes down to an L1 (resp. L2) distribution of re-weighted outliers for large (resp. small) residuals. Using this norm, all records keep a non-zero weight in the inversion.

Using (9) one defines the *a posteriori* model error covariance matrix on the model \mathbf{x} at iteration i ,

$$\mathbf{R}_{\mathbf{xx}i} = \mathbf{C}_{\mathbf{xx}} - \mathbf{K}_i \nabla \mathbf{H}_i \mathbf{C}_{\mathbf{xx}}, \quad (11)$$

from which we build an ensemble of 20 models that statistically satisfy the observations given their associated errors (e.g. Gillet *et al.*

2013), for each data set obtained from the 50 bootstraps discussed in Section 2.2. The final ensemble includes all 20 realizations of the 50 data sets (i.e. a total of $N_e = 1000$ models). This study relies on two key methodological points: (i) the choice of *a priori* temporal cross-covariances that enter the matrix $\mathbf{C}_{\mathbf{xx}}$, and (ii) the projection in time of the field model using these covariance functions, through eq. (5), as performed by Hellio *et al.* (2014) for regional studies. This second point makes the method self-consistent, avoiding the use of support function such as the widely used splines introduced by Bloxham & Jackson (1992), and constitutes a major difference with the stochastic modelling approach proposed by Gillet *et al.* (2013).

2.4 Choice of *a priori* covariances

2.4.1 Spectral behaviour of Gauss coefficients

Studies of the spectra of geomagnetic series have been carried out in order to characterize the behaviour of the axial dipole at different frequencies. Constable & Johnson (2005) constructed a power spectrum $P(f)$ of geomagnetic dipole moment variations from centuries to tens of millions years. It decreases towards increasing frequencies f , and can be separated in several frequency ranges that present different spectral indices p , when approximating locally $P(f) \propto f^{-p}$. For very low frequencies ($f^{-1} > 10^5$ yr), the spectrum seems almost flat. One finds a spectral index $p \simeq 2$ for periods between 10^3 and 10^5 yr. To characterize $P(f)$ on shorter periods, one may resort to modern records. From ground observatory series, $p \simeq 4$ has been observed for periods between about 5 and 100 yr (Currie 1968; De Santis *et al.* 2003). The analysis of lake and marine sediments inclination, declination and intensity data by Panovska *et al.* (2013) indicated $p = 2.3 \pm 0.6$ for periods from 300 to 4000 yr. Olson *et al.* (2012) calculated a time spectrum of the virtual axial dipole moment from numerical simulations of the geodynamo. Their composite spectrum (relying on several computations with different parameters) is overall in a good agreement with the observations listed above. They found $p \simeq 1.8$ and $p \simeq 4$, for respectively $f^{-1} \in [400, 2 \times 10^5]$ yr and $f^{-1} \in [10, 400]$ yr. This is corroborated by the analysis of the axial dipole spectrum in simulations by Buffett (2015) and Bouligand *et al.* (2016), although one should keep in mind that (i) the time scaling of the computation is a crucial step, (ii) the considered simulations should be focused towards the targetted physics and (iii) current simulations still show too much diffusive processes, which hinder interpreting the spectra towards short periods.

Given the small amount of observations covering long periods, we can hardly deduce the spectral behaviour of non-dipole coefficients from observations on centennial and longer periods. The analysis of observatory data by Lesur *et al.* (2017) confirmed a spectral index of ≈ 4 for Gauss coefficients up to a degree about 5 over interannual and decadal periods. This finding agrees with the study by Bouligand *et al.* (2016) of Gauss coefficients series from high-resolution numerical simulations. For all coefficients but the axial dipole, they observed a transition from $p \simeq 4$ to $p \simeq 0$ from short to long periods. The cut-off period between the two frequency ranges decreases with the length-scale, and appears to be governed by the ratio of the main field and secular variation spatial spectra, as assumed by Gillet *et al.* (2013) when constructing the COV-OBS field model. In the analysed simulations, only the axial dipole spectrum presents an intermediate frequency range with $p \simeq 2$.

2.4.2 Mathematical representation of temporal cross-covariances

Following Gillet *et al.* (2013) and Helliou *et al.* (2014), we assume that each Gauss coefficient results from an order 2 autoregressive (AR-2), stationary, stochastic Gaussian process $\varphi(t)$, whose statistics are characterized by an average $\bar{\varphi} = E(\varphi)$, a variance $\sigma^2 = E(\varphi'^2)$ with $\varphi' = \varphi - \bar{\varphi}$, and an autocorrelation function $\rho(\tau) = E(\varphi'(t)\varphi'(t+\tau))/\sigma^2$ with τ the time lag. We assume that these only depend on the harmonic degree l , with no cross-covariances between coefficients of different degree and order (i.e. no spatial cross-correlations):

$$\forall t, C_l(\tau) = E \left[(g_l^m(t) - \bar{g}_l^m)(g_l^{m'}(t+\tau) - \bar{g}_l^{m'}) \right] = \delta_{ll'} \delta_{mm'} \sigma_l^2 \rho_l(\tau), \quad (12)$$

with a similar definition for h_l^m coefficients. We build the covariance matrix \mathbf{C}_{xx} (see Section 2.3) with the quantities of eq. (12). Here the \bar{g}_l^m represent the coefficients of the background model $\bar{\mathbf{m}} = E(\mathbf{m})$.

Helliou (2015) and Bouligand *et al.* (2016) prescribed an AR-2 process for the axial dipole, described through a stochastic equation for φ' of the form

$$d \frac{d\varphi'}{dt} + 2\chi d\varphi' + \omega^2 \varphi' dt = d\zeta(t), \quad (13)$$

where $\zeta(t)$ is a Brownian motion (or Wiener process). It is based on three parameters: the variance σ^2 and two timescales ω^{-1} and χ^{-1} . For $\chi > \omega$, its correlation function is

$$\rho(\tau) = \frac{1}{2\xi} \left[(\chi + \xi) e^{-(\chi - \xi)|\tau|} - (\chi - \xi) e^{-(\chi + \xi)|\tau|} \right], \quad (14)$$

with $\xi^2 = \chi^2 - \omega^2$. Its power spectral density (PSD),

$$P(f) = \frac{4\chi\omega^2\sigma^2}{(\omega^2 - (2\pi f)^2)^2 + (4\pi\chi f)^2}, \quad (15)$$

has spectral indices $p = 0, 2$ and 4 from low to high frequencies. The cut-off period between frequency ranges with 0 and 2 spectral indices is $T_s = 2\pi(\chi + \xi)/\omega^2$; that between $p = 2$ and $p = 4$ frequency ranges is $T_f = 2\pi(\chi - \xi)/\omega^2$ – see Bouligand *et al.* (2016). In the limit $\omega \ll \chi$, they become, respectively,

$$T_s = 4\pi\chi/\omega^2, \quad T_f = \pi/\chi. \quad (16)$$

For others coefficients than g_1^0 , Bouligand *et al.* (2016) considered the particular case $\chi = \omega$ that reduces eq. (13) to a two-parameter (the variance and a single timescale) AR-2 process, of correlation function

$$\rho(\tau) = (1 + \omega|\tau|) \exp(-\omega|\tau|), \quad (17)$$

and for which there is no frequency range where $P(f) \propto f^{-2}$.

For all coefficients but the axial dipole, we thus follow Gillet *et al.* (2013) and consider a two-parameter AR-2 process to define their temporal cross-covariances. We determine their *a priori* variances from the coefficients of the satellite field model CHAOS-6 (Finlay *et al.* 2016), estimated at a single epoch (in $t^* = 2005$):

$$\forall l \geq 2, \sigma_l^2 = \frac{1}{2l+1} \sum_{m=0}^l (g_l^{m2} + h_l^{m2})_{t=t^*}. \quad (18)$$

The variance for equatorial dipole is obtained as

$$\sigma_1^2 = \frac{1}{2} (g_1^{12} + h_1^{12})_{t=t^*}. \quad (19)$$

The time ω^{-1} entering eq. (17) will depend on the degree l . It is simply given by (Gillet *et al.* 2013; Bouligand *et al.* 2016)

$$\forall l, \omega_l^{-1} = \sqrt{\sigma_l^2 / \dot{\sigma}_l^2}, \quad (20)$$

where $\dot{\sigma}_l^2$ is obtained by replacing the coefficients (g_l^m, h_l^m) by their time derivatives ($\partial_t g_l^m, \partial_t h_l^m$) in eqs (18) and (19).

We proceed slightly differently for the axial dipole. We consider the anomaly to the background value, $g_1^{0'} = g_1^0 - \bar{g}_1^0$. The background value $\bar{g}_1^0 = -23\mu\text{T}$, and the standard deviation for the anomaly $\sqrt{E(g_1^{0'2})} = 6.0\mu\text{T}$ are estimated by translating paleomagnetic reconstructions of the virtual axial dipole moment (its time-averaged and standard deviation over the Brunhes chron, see Ziegler *et al.* 2011) into axial dipole values. We consider $\omega_1^{-1} = 400$ yr, as defined from eq. (20), to estimate the time entering eq. (14) for the axial dipole. Then, the second time χ^{-1} is fixed at 20 yr in order to reproduce the changes of spectral index in the observed and simulated spectra, with the two cut-off between $p = 0, 2$ and 4 frequency ranges at $T_s \simeq 100\,000$ yr and $T_f \simeq 60$ yr – see eq. (16). We show in Fig. 2 the PSD associated with the covariance functions at the different degrees used in this study. Temporal and spatial distributions of archeomagnetic and sediment data prevent to reach such a high resolution. However, truncating our models at degree $l = 10$ avoids to assign modelling errors to the error budget.

3 RESULTS

3.1 Validation with synthetic tests

We create a synthetic set of reference Gauss coefficients that satisfy the cross-covariance functions defined in Section 2.4. We evaluate the predictions from this model at the locations and epochs of the two data sets described in Section 2.1. We entach the value and age of these data of errors, within the dating and measurement uncertainties of the corresponding data set. These data sets, which at first do not contain outliers, are called Arch-Synth and Lake-Synth, the latter one containing archeological, volcanic and sediment data. For the sake of comparison, both data sets come from the same synthetic model, and archeological and volcanic data included in Lake-Synth are the same than in Arch-Synth. We name COV-ARCH-Synth and COV-LAKE-Synth the ensembles of models obtained in the two configurations.

Furthermore, in order to illustrate how outliers impact field models and guide our choice of measure for the residuals, we also consider in Section 3.1.3 a data set entached by outliers (Arch-Synth \dagger), from which we derive ensemble of models with either a Huber or an L2 measure of the residuals (named respectively COV-ARCH-Synth \ddagger and COV-ARCH-Synth \ddagger -L2). We should mention these synthetic tests represent a best-case scenarii, as we only address issues associated with the probability distributions of measurements and of records dates. The application to geophysical data incorporates obviously many more subtleties, in particular a temporal smoothing associated with sediment depositional processes (Pavón-Carrasco *et al.* 2014).

3.1.1 Data misfit and bias

We consider first inversions using a Huber measure of residuals. We report in Table 1 the normalized data misfit, defined as the ensemble average r.m.s prediction error

$$\mathcal{M} = \frac{1}{N_e} \sum_{j=1}^{N_e} \sqrt{\frac{1}{N_o} \sum_{k=1}^{N_o} \left(\frac{r_{k,j}}{\tilde{\sigma}_{k,j}} \right)^2}, \quad (21)$$

with $\tilde{\sigma}_{k,i}$ defined in eq. (10) and N_o , the number of data. We also give the normalized prediction bias, defined as the ensemble average

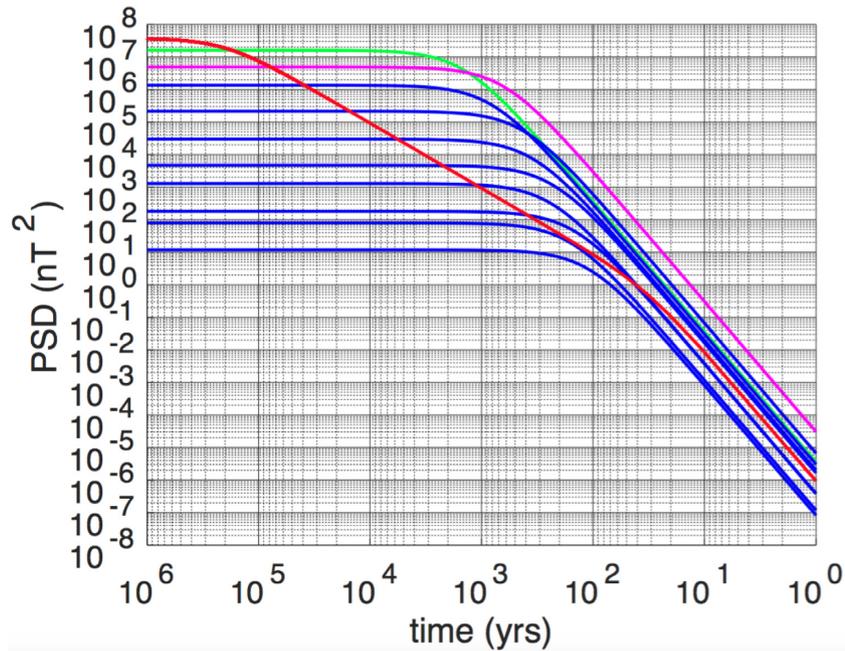


Figure 2. PSD of the covariance for the several spherical harmonic degrees l . In red the PSD for the axial dipole, in green for the equatorial dipole (g_1^1 and h_1^1), in magenta for degree $l = 2$ and in blue for degrees 3 to 10 (variance decreases with degree).

Table 1. Ensemble average data misfit and bias in the synthetic case. \mathcal{B} and \mathcal{M} are defined in eqs (21) and (22). They are estimated both globally and separately for the different types of observations (D , I and F), for the COV-ARCH-Synth and COV-LAKE-Synth ensembles (COV-ARCH-Synth* is obtained with no dating errors and COV-ARCH-Synth** with no dating errors and without bootstrap on ages), using a Huber measure of the residuals. COV-ARCH-Synth† (resp. COV-ARCH-Synth†-L2) is obtained in presence of outliers with a Huber (resp. L2) norm of the residuals.

		Archeo & lava			Sediments			Global
		D	I	F	D	I	F	
COV-ARCH-Synth	N_o	3217	4633	3424	10 273	12 101	4 710	
	\mathcal{M}	1.53	1.54	1.22	–	–	–	1.45
COV-LAKE-Synth	\mathcal{B}	0.02	0.06	–0.03	–	–	–	0.02
	\mathcal{M}	1.52	1.52	1.19	1.07	0.96	0.91	1.14
COV-ARCH-Synth**	\mathcal{B}	0.01	0.04	–0.01	–0.01	0.02	0.03	0.01
	\mathcal{M}	1.21	1.20	1.20	–	–	–	1.20
COV-ARCH-Synth*	\mathcal{B}	0.09	0.04	–0.04	–	–	–	–0.03
	\mathcal{M}	1.55	1.53	1.25	–	–	–	1.46
COV-ARCH-Synth†	\mathcal{B}	0.04	0.08	–0.04	–	–	–	0.03
	\mathcal{M}	1.71	1.93	2.30	–	–	–	2.00
COV-ARCH-Synth†-L2	\mathcal{B}	–0.01	–0.09	0.12	–	–	–	–0.01
	\mathcal{M}	1.21	1.20	1.09	–	–	–	1.17
	\mathcal{B}	0.01	–0.01	–0.01	–	–	–	–0.01

prediction error,

$$\mathcal{B} = \frac{1}{N_e} \sum_{j=1}^{N_e} \frac{1}{N_o} \sum_{k=1}^{N_o} \frac{r_{k,j}}{\tilde{\sigma}_{k,j}}. \quad (22)$$

The two above quantities are given separately for D , I and F data. Biases are close to zero for both ensembles of models, for all three field components. Misfits for sediment data are close to unity. For archeomagnetic and lava records, misfits are similar for both ensembles of models. They are larger than one by ≈ 20 and 50 per cent for intensity and directional data, respectively. We suspect it to be related to incompatibilities between different (D , I and/or F) data, whose ages are drawn independently one to the other within each bootstrap. To check this hypothesis, we consider inversion of

Arch-Synth data with unperturbed dates. It leads to similar misfits when performing bootstraps (case COV-ARCH-Synth*), while misfits are significantly lower when no bootstrap is applied (case COV-ARCH-Synth**) – see Table 1. It confirms the incompatibility of the dates chosen with the bootstrap method for archeomagnetic data. This effect would be much reduced if considering either larger measurement errors (the avenue followed by e.g. Korte *et al.* 2009, at the expense of potentially losing precious information), or by selecting the most probable draws of dates (as performed by Hellio *et al.* 2014, a method unfortunately unaffordable with global data sets). This is less of a problem for sediment data (i) because of larger data errors that allow for more freedom in the model reconstitution, and (ii) since dates of the different field components in one core result from the same draw (see Section 2.2).

3.1.2 Model recovery and uncertainties

The two ensembles COV-ARCH-Synth and COV-LAKE-Synth are shown in Fig. 3, through the ensemble mean and the 3σ spread within the ensemble of models, for some dipole and quadrupole field Gauss coefficients. For all of them, the reference series is contained within the 3σ . Considering the sediment data, the dispersion within the ensemble is significantly decreased, especially for epochs before 1500 AD. Our modelling strategy allows to retrieve some of the centennial oscillations observed in the reference series (e.g. g_2^0). The ensemble average model obtained from Lake-Synth is closer to the reference series, in comparison with that inverted from Arch-Synth. This is particularly obvious for some coefficients such as h_2^1 , for which COV-ARCH-Synth, contrary to COV-LAKE-Synth, fails at recovering the amplitude of the strong oscillations seen in the reference curve.

The recovery of the reference model within the posterior model uncertainties is confirmed with Fig. 4, where we show the time-averaged normalized error as a function of l (measured as the difference between the ensemble average model $\langle \mathbf{x} \rangle$ and the reference model $\hat{\mathbf{x}}$, normalized by the *a priori* variance per degree),

$$\Psi^2(l) = \frac{1}{(2l+1)\sigma_l^2(t_e - t_s)} \times \int_{t_s}^{t_e} \sum_{m=0}^l ((g_l^m(t) - \hat{g}_l^m(t))^2 + (h_l^m(t) - \hat{h}_l^m(t))^2) dt, \quad (23)$$

and the time-averaged dispersion within the ensemble (normalized by the *a priori* variance per degree) as a function of l ,

$$\Phi^2(l) = \frac{1}{(2l+1)\sigma_l^2(t_e - t_s)N_e} \times \int_{t_s}^{t_e} \sum_{i=1}^{N_e} \sum_{m=0}^l (g_{l,i}^m(t) - \langle g_l^m(t) \rangle)^2 + (h_{l,i}^m(t) - \langle h_l^m(t) \rangle)^2 dt. \quad (24)$$

The curves $\Psi^2(l)$ show that taking into account sediment data, despite their large associated uncertainties, increases the resolution on Gauss coefficients: some partial information is retrieved up to degree $l = 5$ with COV-LAKE-Synth, instead of $l = 4$ with COV-ARCH-Synth. Furthermore, the curves Φ^2 are for all degrees very close to Ψ^2 : we thus conclude that in absence of outliers, and for realistic allocated data errors, our method provides a reasonable measure of the posterior model uncertainty (despite data misfit slightly larger than one, as shown in Table 1). In the spatial domain, we recover larger uncertainties in the southern hemisphere and generally where the data sampling is low (see in Fig. A1 the snapshot CMB maps of the dispersion within our ensemble of solutions). The dispersion is weaker when including sediment records, and we also witness the decrease of the data constraint towards the past.

3.1.3 Impact of outliers

An example of model predictions at the Earth's surface is given in Fig. 5 for COV-ARCH-Synth. In absence of outliers, the PDF of the prediction accounts well for the reference intensity curve, although some oscillations of period about 250 yr are partially smeared out. We now experiment a case with 12 per cent outliers (data set COV-ARCH-Synth†, chosen to be at the same dates and locations of the data identified as outliers in the geophysical application, for which $|r_k|/\sigma_k > 2$; to assign physically realistic intensity values to outliers, we replace the measurement values in Arch-Synth by

the corresponding values in the geophysical data set). In this latter case, error bars exclude the reference curve at some periods (see around 430 BC and after 1800 AD), and the 250 yr fluctuations are now poorly recovered. Despite so far, results were obtained with a Huber measure of residuals, our models are spoiled by the presence of outliers, with normalized data misfits significantly larger than unity (see Table 1). It also transpires into anomalously high values of the time-averaged spatial power spectrum at the top of the Earth's core (of radius $c = 3485$ km),

$$R(l) = \frac{1}{t_e - t_s} \int_{t_e}^{t_s} \left(\frac{a}{c}\right)^{2l+4} (l+1) \sum_{m=0}^l (g_l^m(t)^2 + h_l^m(t)^2) dt, \quad (25)$$

near the limit of resolution (see harmonic degrees 4 and 5 in Fig. 5). We thus consider the case where instead, an L2 norm is employed together with a 3σ rejection criterion (with no reweighting of the errors, i.e. $\tilde{\sigma}_k = \sigma_k$). Normalized misfits (resp. biases) obtained for COV-ARCH-synth†-L2, given in Table 1, are now reasonably close to 1 (resp. 0). The prediction in Paris is now very close to the case without outliers for both the mean model and the dispersion. Finally, spectra for both the ensemble average model and the dispersion within the ensemble of models are also very close to the case obtained in absence of outliers—the overestimation at degrees 4–5 has disappeared. A similar behavior (reduction of the misfit to the true curve, and of the too large power at degrees 4 and 5, while using an L2 instead of a Huber norm) is observed when including sediment data. For the geophysical application below, we thus prefer employing an L2 measure of normalized residuals rather than a Huber norm.

3.2 COV-ARCH and COV-LAKE

3.2.1 Data misfits and biases

We now apply the method described in Section 2 to the archeomagnetic and sediment data sets presented in Section 2.1. Data misfits and biases, as defined by eqs (21) and (22), are detailed in Table 2 for COV-ARCH and COV-LAKE, separately for each data type. We recall that we use for the geophysical application an L2-norm for the data misfit. For archeomagnetic and lava data, we obtain for both COV-ARCH and COV-LAKE misfits about 10 per cent larger than the ones found with the synthetic study in presence of outliers. As for the synthetic study, misfits much closer to unity are found for sediment cores. These tend to lower the global misfit, contrary to what was observed by Licht *et al.* (2013), where the misfit for ASDI.FM is higher than that for A.FM. The origin of this difference is related to the errors assigned to sediment series, since we use the (larger) values proposed by Panovska *et al.* (2013) for our sediment measurement uncertainties. The distribution of normalized data residuals (see Fig. B1) is close to Gaussian for directional data, while it is a bit more peaked towards zero for intensities.

3.2.2 Time evolution of Gauss coefficients and their dispersion

We show in Fig. 6 the temporal behaviour of the dipole and quadrupole coefficients for COV-ARCH and COV-LAKE. All coefficients are compared with the corresponding series obtained by Licht *et al.* (2013), and with the historical model *gufm1* (Jackson *et al.* 2000). First of all, *gufm1* series always lay within the COV-ARCH and COV-LAKE ensembles, which is very encouraging because our models are independent of *gufm1* (we do not account for any historical data), unlike Licht *et al.* (2013) who anchored

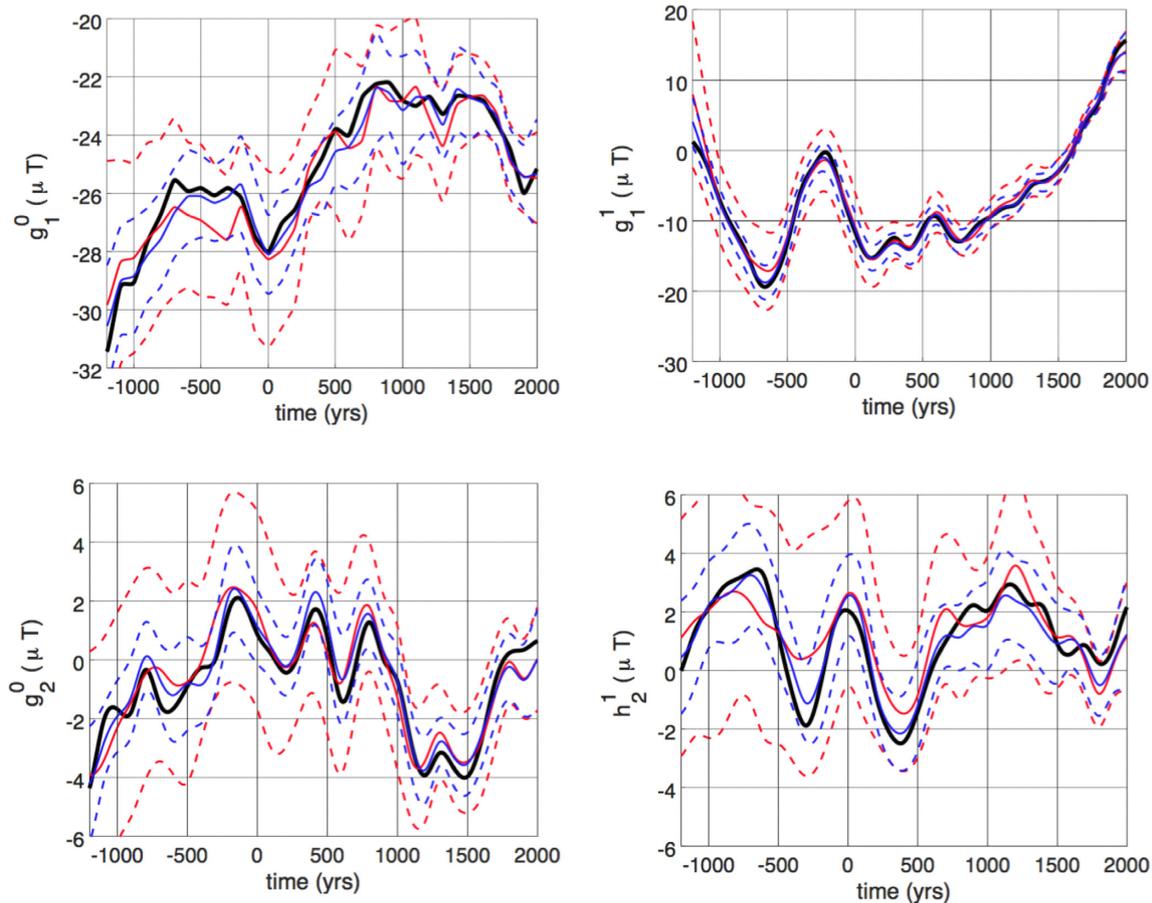


Figure 3. Synthetic Gauss coefficients series (reference series in black) for some coefficients of the dipole and quadrupole field, compared with the mean model retrieved from the Arch-Synth (red) and Lake-Synth (blue) data sets, together with their corresponding uncertainties, as measured by $\pm 3\sigma$ (dotted lines).

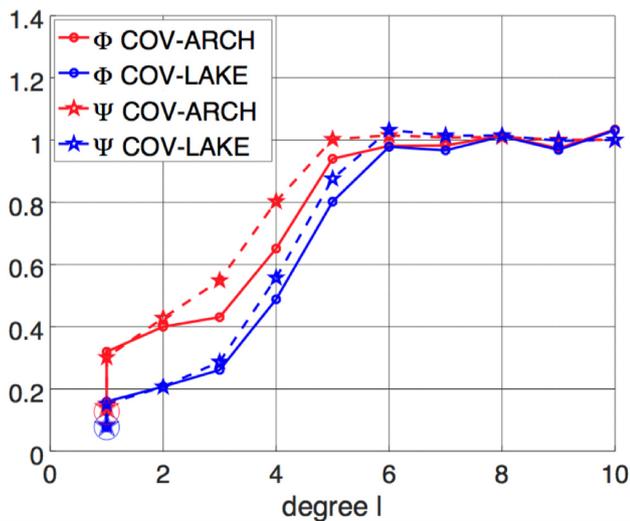


Figure 4. Diagnostics Ψ (circled plain lines) and Φ (dotted lines with stars) as a function of harmonic degree l , as defined in eqs (23) and (24), for the field models inverted from Arch-Synth (red) and Lake-Synth (blue) data sets. Diagnostics for the axial and equatorial dipoles are shown separately (with larger markers for g_1^0).

their models to data generated from *gufm1* predictions in 1990. We also note that including sediment data generally improves the fit of the ensemble average model to *gufm1*, whereas it is generally better for A_FM than for ASDI_FM. As observed with the synthetic

study in Section 3.1, the dispersion within the ensemble is larger for COV-ARCH than for COV-LAKE. The ensemble spread is almost constant during the whole period considered when accounting for sediment data, due to their more homogeneous temporal distribution (see Fig. 1).

Having a closer look at individual coefficients, the axial dipole for COV-ARCH (resp. COV-LAKE) is smoother than that of A_FM (resp. ASDI_FM). Our estimates for g_1^0 are also in average weaker by $\approx 3 \mu\text{T}$ prior to 1500 AD in comparison with those of Licht *et al.* (2013). Although this could be due to the particular prior we chose for g_1^0 , the corresponding coefficient for CALS3k.4 is even less energetic, particularly for periods before 500 AD. COV-LAKE (and to a lesser extent COV-ARCH) models agree well with the constant axial dipole decay predicted by *gufm1*. We recall that *gufm1* is only constrained by directional data before 1832 and the first absolute intensity measurements: as a consequence, Jackson *et al.* (2000) had to scale the dipole prior to this date (see Suttie *et al.* 2011) and assumed a constant dipole decay $\text{eqn}_{10d4a-10807} \text{d}g_1^0/\text{d}t = 15 \text{ nT}\cdot\text{yr}^{-1}$. Several studies discussed this issue. Gubbins *et al.* (2006) argued, on the basis of paleomagnetic data analysis, in favour of a linear decay much slower than the one considered by *gufm1*. Finlay (2008), using the database from CALS7K (Korte *et al.* 2005) to constrain the axial dipole between 1590 and 1832, found a constant axial dipole to be more likely over this period. Suttie *et al.* (2011) re-estimated the data errors of the data set used by Finlay (2008) and found a dipole decaying at an approximately constant rate of $12 \text{ nT}\cdot\text{yr}^{-1}$ over the historical period. We propose here a constant decay of the axial dipole of $16 \pm 2 \text{ nT}\cdot\text{yr}^{-1}$ over the past three

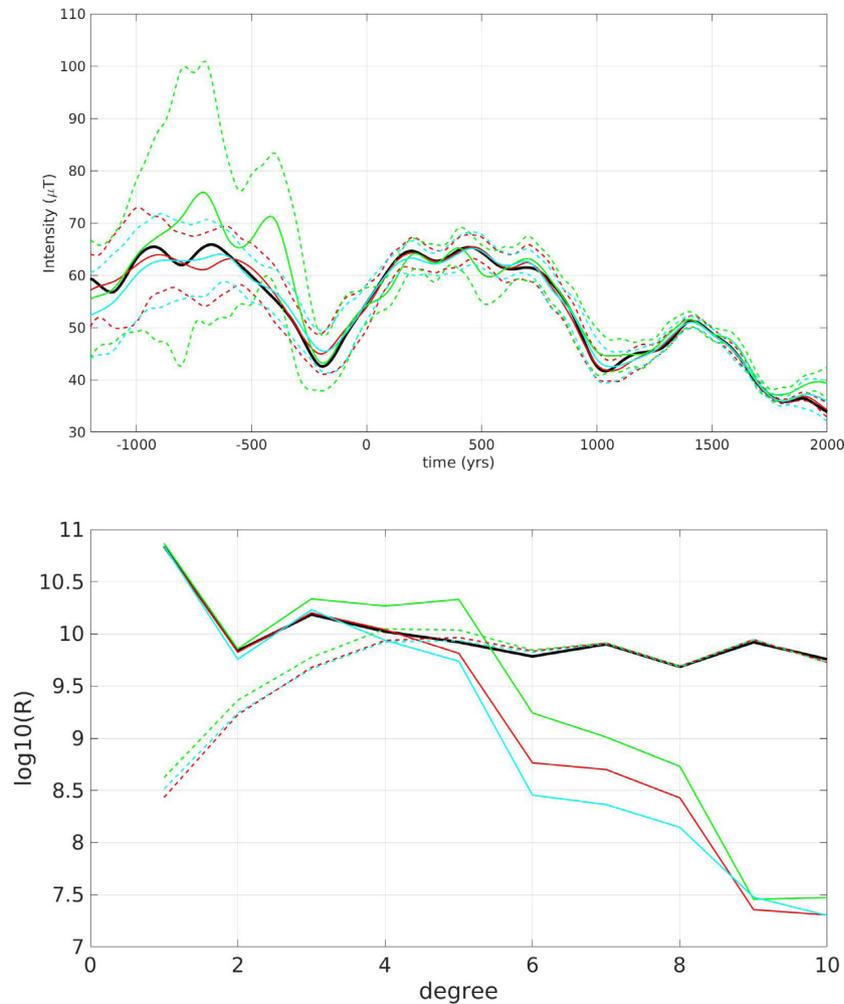


Figure 5. Top: predictions of the COV-ARCH-Synth model for the intensity in Paris (dotted lines: $\pm 3\sigma$), without outliers using a Huber norm (red), with 12 per cent outliers using a Huber norm (green) or an L2 norm and a 3σ rejection criterion (cyan), compared with the reference model prediction (black). Bottom: time-averaged spatial spectra $R(l)$ for the average solution (same colour code, in dotted lines the corresponding spectra for the dispersion within the ensembles) and the reference model (black).

Table 2. Same as Table 1, for COV-ARCH and COV-LAKE.

		Archeo & lava			Sediments			Global
		D	I	F	D	I	F	
COV-ARCH	\mathcal{M}	1.27	1.34	1.22	–	–	–	1.28
	\mathcal{B}	–0.04	–0.03	–0.01	–	–	–	–0.02
COV-LAKE	\mathcal{M}	1.30	1.35	1.24	0.98	0.99	0.81	1.05
	\mathcal{B}	0.02	0.03	0.01	–0.07	–0.10	0.05	–0.04

centuries for COV-LAKE and 11 ± 3 nT.yr $^{-1}$ for COV-ARCH. Prior to the historical era, both COV-LAKE and COV-ARCH are mostly constant: over the period spanning 1000 to 1700 AD, COV-LAKE shows a small growth of 2 ± 1 nT.yr $^{-1}$ (to be compared with 0 ± 1 nT.yr $^{-1}$ from COV-ARCH). This contrasts with recent findings by Poletti *et al.* (2018), who favour a linear decrease from 750 AD.

From ≈ 1700 AD, most model coefficients agree reasonably well with *gufm1*. In particular the sharp increase in g_2^2 , the intense maximum in h_2^2 around 1800 or the strong decrease in h_2^1 are nicely reproduced. This latter feature is better recovered when considering sediment data. The axial dipole put aside, g_2^0 is the coefficient showing the most significant non-zero average through time, with

a negative shift. We note the very good agreement between COV-LAKE and ASDL.FM for g_2^1 before 1500 AD. We also note that g_2^2 from COV-LAKE and COV-ARCH show similar time-averages, while before 1000 AD, g_2^2 from ASDL.FM is significantly shifted upwards (in contrast with A.FM); the same discrepancy holds between CALS3k.4 and ARCH3k.1.

Two clear oscillations appear in the g_1^1 series for the archeomagnetic models (COV-ARCH and A.FM), with maxima at epochs 100 AD and 1100 AD. Although slightly smoother, they also show up in the corresponding COV-LAKE coefficient, but not so clearly on ASDL.FM. Over the time window considered here, h_1^1 presents saw-tooth patterns on longer timescales, retrieved in all field models, including or not sediment data. Centennial oscillations are seen

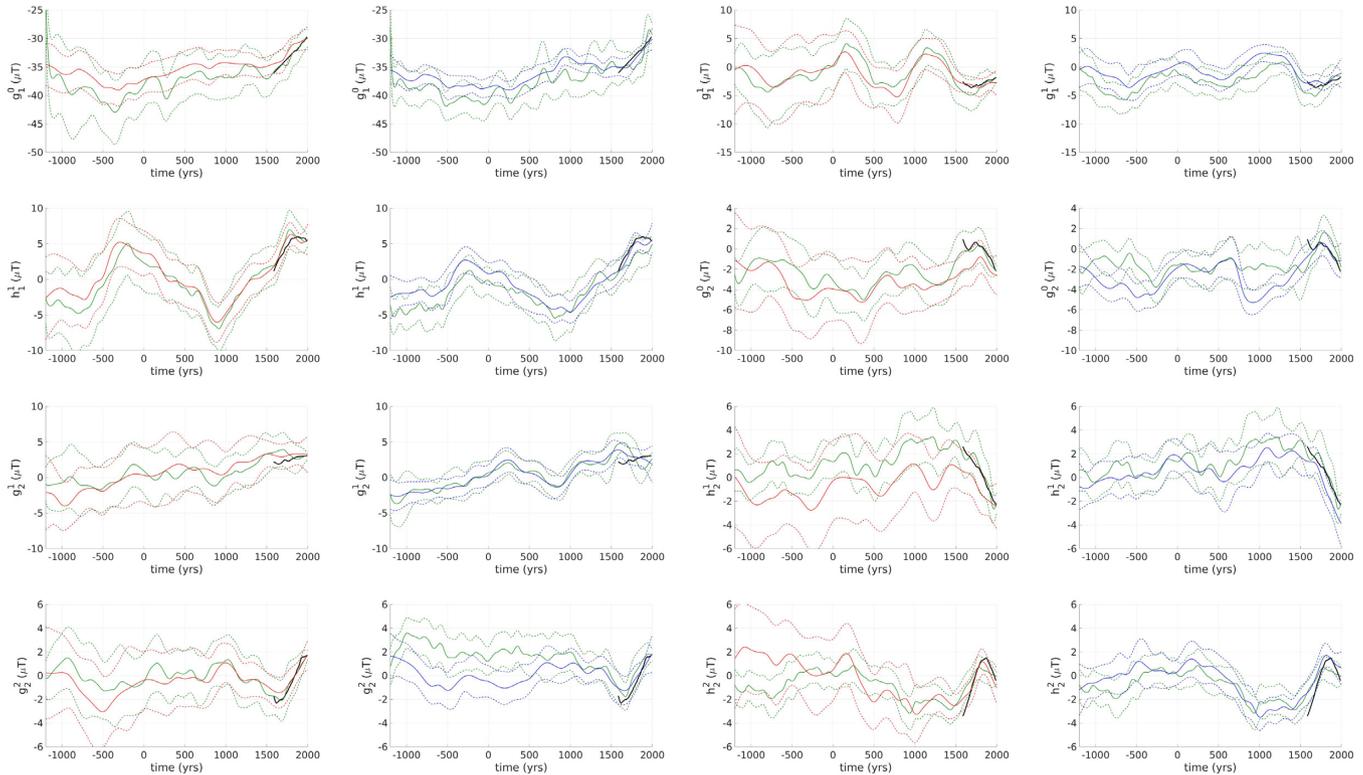


Figure 6. Dipole and quadrupole Gauss coefficients series for COV-ARCH (red), A.FM (green), COV-LAKE (blue) and ASDI.FM (green). For each model, the $\pm 3\sigma$ standard deviation is shown in dotted lines. The historical model *gufm1* is superimposed in black between 1600 and 1990 AD.

in h_2^2 and h_1^1 series, which local extrema occur around the same epoch for both COV-ARCH and A.FM. Such quasi-periodic patterns in particular for h_2^2 are also present in the AmR and ARCH3k archeomagnetic field models (see fig. 9 in Sanchez *et al.* 2016). Comparing with models including sediment data, such features, however, almost disappear.

3.2.3 On large-scale structures of the radial field

We now detail main field patterns at the CMB. We first focus on the time-averaged field over the past 3000 yr, whose ensemble averages for COV-ARCH and COV-LAKE are shown in Fig. 7, together with the dispersion within the ensemble of time-averaged models. Interestingly, non-axisymmetric patches appear persistent over the whole studied era. Several of these features are found in both COV-ARCH and COV-LAKE, as for instance the local extrema at mid to low latitudes around 100°E , 30°E and 100°W in the northern hemisphere. These patterns, in particular with COV-LAKE, look symmetric with respect to a magnetic equator slightly shifted northwards of the geographical equator, reminding the patches put forward by Jackson (2003) from modern observations. Already documented high-latitude lobes are also recovered, for example, under Northern America, and below the Eastern Antarctica (Constable 2007). Comparing our time-averaged fields with the corresponding time-averaged fields of ARCH3k.1 and CALS3k.4 (Figs 7a and b), we see some similar structures, for example, the high-latitude lobes in the southern hemisphere for CALS3k.4 and COV-LAKE. The mid-latitude patches symmetric to the equator present in COV-LAKE and COV-ARCH appear in the northern hemisphere for ARCH3k.1 and CALS3k.4, but these smoother models show no corresponding southern patterns.

We note discrepancies between our two time-averaged models: for instance, the areas of weak B_r , located under South Atlantic in COV-ARCH, and south of Madagascar in COV-LAKE. However, the uncertainty on these two low B_r areas are relatively large. Indeed, for the time-averaged COV-ARCH, below almost all oceans of the southern hemisphere the dispersion within the ensemble of time-averaged models is as high as $70 \mu\text{T}$. The resolution is better for COV-LAKE, but the error under the Indian Ocean still reaches some $50 \mu\text{T}$. This calls for some caution regarding the low B_r areas of the southern hemisphere seen in Fig. 7.

Examples of CMB maps of B_r at specific epochs (200 BC and 1400 AD) are shown in Fig. 8, together with their associated uncertainties. Posterior model errors are much reduced under Europe (and generally the northern hemisphere), where a higher density of data is available. The dispersion is also significantly lower for COV-LAKE than for COV-ARCH, at any given epoch. Flux patches discussed above for the time-averaged maps are retrieved on the snapshot field maps, with a stronger contrast. Local extrema are generally higher for COV-LAKE than for COV-ARCH. Most differences between the two ensemble average models show up in the southern hemisphere, and at older epochs when fewer archeomagnetic data are available. On the contrary, at more recent times (Figs 8 c and d), we retrieve in both models most radial field patterns positioned at similar locations, and this in both hemispheres – see in particular the low- to mid-latitude patches in the southern hemisphere that show up in 1400 AD, although some of them lay in areas of relatively high uncertainties.

We show in Fig. 9 the time-averaged spatial spectra $R(l)$ at the CMB for several models including ours. The power stored in model uncertainties (dotted lines) takes over that of the ensemble average model above spherical harmonic degree about 5 for both COV-ARCH and COV-LAKE, to then reach the prior power spectrum of

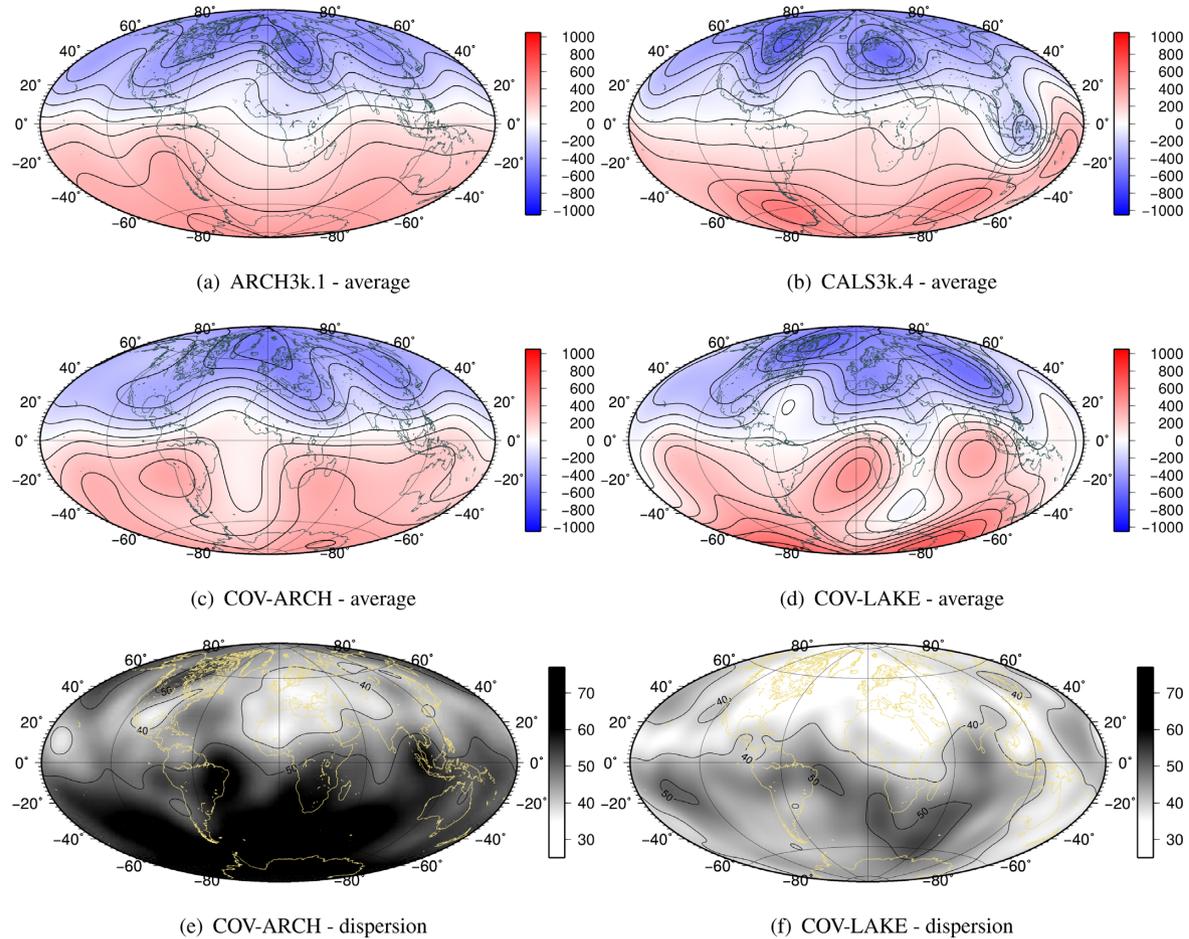


Figure 7. Time-averaged, over the past three millennia, of the ensemble average radial magnetic field B_r at the CMB (a–d), and the associated dispersion within the ensemble of time-averaged models (e and f). Contours are every $100 \mu\text{T}$ for the average and $10 \mu\text{T}$ for the dispersion.

CHAOS-6. It confirms the result found with synthetic experiment in Section 3.1 that no information is to be expected at $l > 5$ (see also Sanchez *et al.* 2016). The decrease of $R(l)$ towards large degrees for the ensemble average model is steeper than (resp. similar to) that found for regularized models also truncated at degree 10 for COV-ARCH (resp. COV-LAKE). The power stored into degrees from 1 to 5 for COV-ARCH, very similar to that of SHA.DIF.14k, is significantly higher than that of AmR, based on a dynamo spatial norm, for $l \geq 3$. The power for the ensemble average COV-LAKE model is close to the reference spectrum of CHAOS-6 for $l \leq 3$. At degrees 4 and 5, it is significantly higher than $R(l)$ found for all others models, including the prior model. Although not incompatible with the large natural fluctuations of the spatial spectrum found in geodynamo simulations (see Fig 3 in Bouligand *et al.* 2016), it indicates some possible remaining impact of outliers onto COV-LAKE (despite the use of a rejection criterion, see section 3.1.3).

4 DISCUSSION

4.1 Westward drift of field patches at the CMB

We revisit the quest for azimuthal drift of field patches. Following Dumberry and Finlay (2007), we show in Fig. 10 some examples of time–longitude diagrams for radial field anomalies at the CMB (once removing the axisymmetric time-averaged, and filtering out periods longer than 2000 yr). We focus on the northern hemisphere

for COV-LAKE, which presents the weaker uncertainties. Similar conclusions are found with COV-ARCH, and as in Dumberry & Finlay no clear drift shows up in the southern hemisphere. The most robust feature is a westward drift, at a speed of ≈ 0.20 to $0.25^\circ \text{yr}^{-1}$ (i.e. it takes about 1500 to 1800 yr to circulate around the globe). Dominated by large length-scales, it is reproducible from one realization to the other. The drift estimate is nevertheless subject to important uncertainties in link with the too-short-studied era (as illustrated with fig. 2 in Dumberry & Finlay 2007).

We detect no significant variation with latitude of the travelling speed of these structures, which are carried by harmonic degrees $l \leq 5$. As in Dumberry & Finlay (2007), they appear more obvious at mid to high latitudes. In comparison with their results, we obtain a comparable (although slightly weaker) drift speed. Eastward motion of field patterns appear less obvious (though still potentially there), and some westward patterns conversely can be followed here over almost all longitudes. We cannot perform a direct comparison with the stronger westward drift of equatorial field patches, highlighted by Finlay & Jackson (2003) from historical records, due to the larger wavelengths/longer periods considered here.

4.2 Low-intensity anomalies at the Earth's surface

A frequently raised issue in the geomagnetic community concerns the persistence over time of the weak-intensity area observed in modern times above the South Atlantic (SAA, or South Atlantic

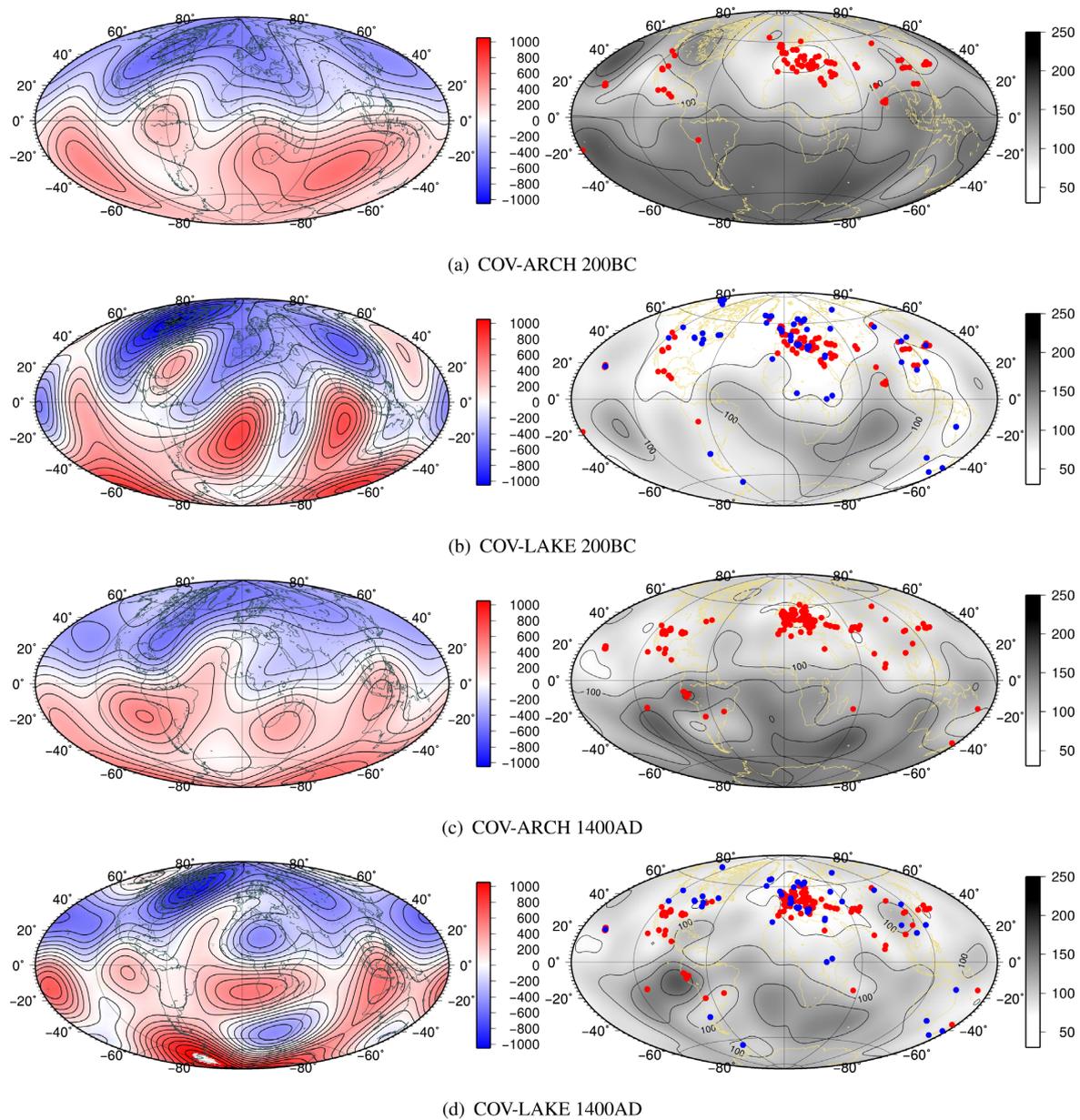


Figure 8. Snapshots of the radial magnetic field B_r at the CMB. left: ensemble average COV-ARCH and COV-LAKE models in 200BC and 1400 AD. Contours are every $100 \mu\text{T}$. right: Corresponding dispersion within the ensembles. Data present at each epoch ($\pm 50 \text{ yr}$) are indicated with red and blue dots. Contours are every $25 \mu\text{T}$.

Anomaly). Archeomagnetic and sediment data allow us to follow backwards the evolution of the intensity at the Earth's surface. We recover well in both COV-ARCH and COV-LAKE the SAA (see Figs 11c and d), and trace it only back to $\approx 1800 \text{ AD}$ (contrary to e.g. CALS3k.4b where it is persistent over a longer era, see Fig. C1).

Towards earlier epochs, we observe several of these low-intensity patches. The most obvious anomaly is located above the Indian Ocean between ≈ 600 and 1400 AD (a less intense one is also found above the Western Pacific in the same time interval) as illustrated in Figs 11(a) and (b). Present, although weaker, when using archeomagnetic data only, the Indian Ocean anomaly of the Middle Age is overtaken in COV-ARCH by a poorly resolved low-intensity area above the Southern Atlantic (see the corresponding dispersion maps), absent from COV-LAKE. This feature does not show up in alternative models such as CALS3k.4b, where it is replaced

by a low-intensity area wider in longitude and closer to the equator (see Fig. C1). This configuration suggests, given the relatively large value of the dipole at the considered period, the existence of relatively strong reversed flux patches in the southern hemisphere around 70°E (Terra-Nova *et al.* 2017), as that of Fig. 8(d).

Gallet *et al.* (2009) propose a link between geomagnetic jerks (abrupt change in the field orientation and extremum of its intensity, see Gallet *et al.* 2003) and epochs characterized by a large fraction of the quadrupolar field (i.e. periods when reducing the field to an eccentric dipole requires strong eccentricities). From this, they speculate that several low-intensity events may have occurred in the past. The rationale behind this is that, for instance, the SAA above the Earth's surface is well approximated by an eccentric dipole (see Domingos *et al.* 2017). Our above observation of an Indian Ocean anomaly thus corroborates the suggestion by Gallet *et al.* (2009).

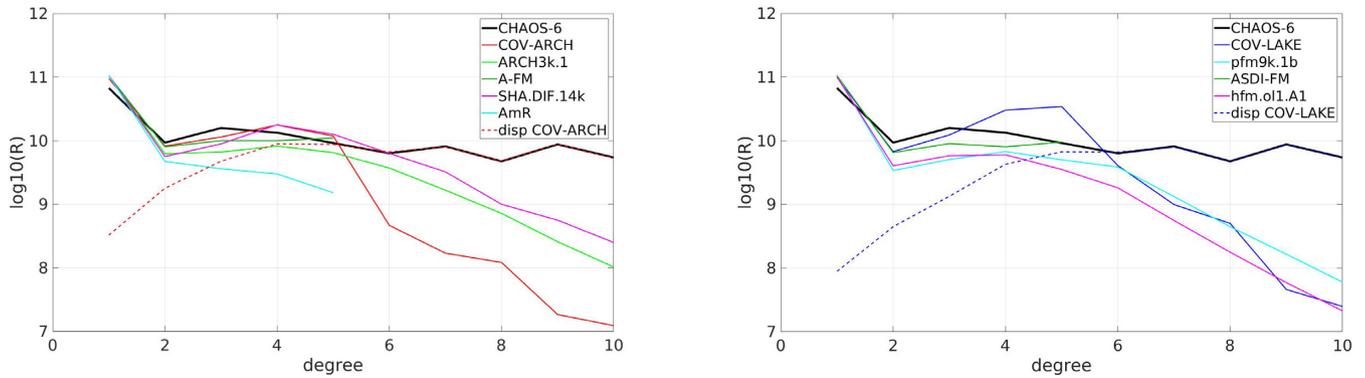


Figure 9. Time-averaged (over the past 3 millennia) power spectra of the field at the CMB for COV-ARCH (red) and COV-LAKE (blue), compared with spectra of alternative field models (see legend for details). That of CHAO-6 in 2005, used to define the *a priori* information (see Section 2.4.2), is shown in black for comparison. The time-averaged power spectra of the *a posteriori* dispersion within the ensemble are shown in dotted lines for COV-ARCH (red) and COV-LAKE (blue).

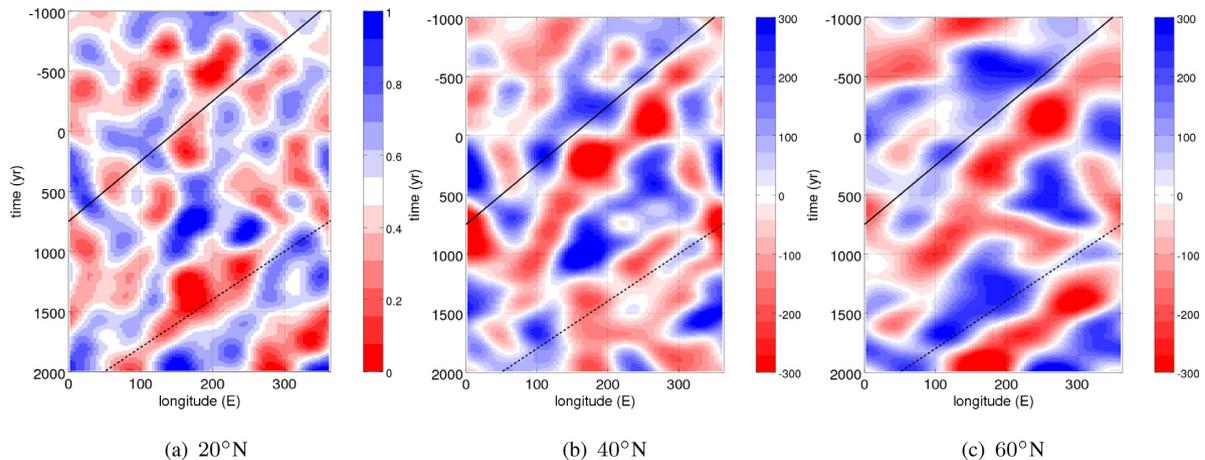


Figure 10. Time–longitude diagrams of the radial field at the CMB, for the COV-LAKE ensemble average model, at 20°N, 40°N and 60°N. The axisymmetric time-averaged has been removed, and periods longer than 2000 yr have been filtered out. The black full (resp. dotted) line corresponds to a westward drift of 0.20° yr⁻¹ (resp. 25° yr⁻¹).

4.3 Confronting regional series and global field models

4.3.1 Regional master curves

We now investigate how global model predictions at Earth’s surface compare with regional studies. Fig. 12 shows the regional master curves, represented as a PDF, obtained using the stochastic method introduced by Helliö *et al.* (2014) with archeomagnetic data from the COV-ARCH data set in a radius of 1250 km around Paris—the best documented area for centennial variations. We focus on the period from 400 AD. The regional curve obtained for the intensity shows six maxima over 400–2000 AD, five of them coinciding with extrema highlighted by Genevey *et al.* (2016), including data in a radius of 700 km around Paris—the missing one corresponding to a separation in two of their local maximum around 1400 AD.

Intensity predictions from COV-ARCH present instead only two local maxima: a broad peak centred around 800 AD, and one less pronounced around 1600 AD that coincides with that of the regional study. Inflections of the global model ensemble average prediction sign some of the missing maxima. Note that corresponding COV-LAKE predictions are very similar, as expected given the very good coverage of archeomagnetic data over this area. The fact that global models fail to reproduce some oscillations isolated in regional studies is intriguing, since the *a priori* information chosen for our global

study allows for such time variations. For completeness, predictions of COV-ARCH at others locations around the globe are given in Fig.D1. Three main differences between the regional curve and the global model predictions could explain such a discrepancy:

(1) *Different ways to handle dating errors:* Contrary to our strategy for building the global models, Helliö *et al.* (2014) did not apply a blind bootstrap on the dates of the samples. Instead they only selected the draws with the highest probabilities, following Markov Chain Monte Carlo rules. They showed that this process potentially increases the contrast of the recovered curve. However, around Paris where data contain relatively small dating uncertainties, we tested that performing or not the selection within the draws does not change much the shape of the regional PDF.

(2) *Difference of a priori information:* The prior on F series in Helliö *et al.* (2014) is derived from the same stochastic framework as the one employed in this paper, summing cross-covariance functions on Gauss coefficients. The only change concerns the axial dipole, defined as a Matérn AR-2 process in the study of Helliö *et al.* (2014) and as a damped oscillator in this study (see Section 2.4.2). It significantly differs from the smoothing constraint applied by Genevey *et al.* (2016). Since both regional approaches give reasonably close master curves (with at least five local maxima in Paris), this hypothesis appears unlikely.

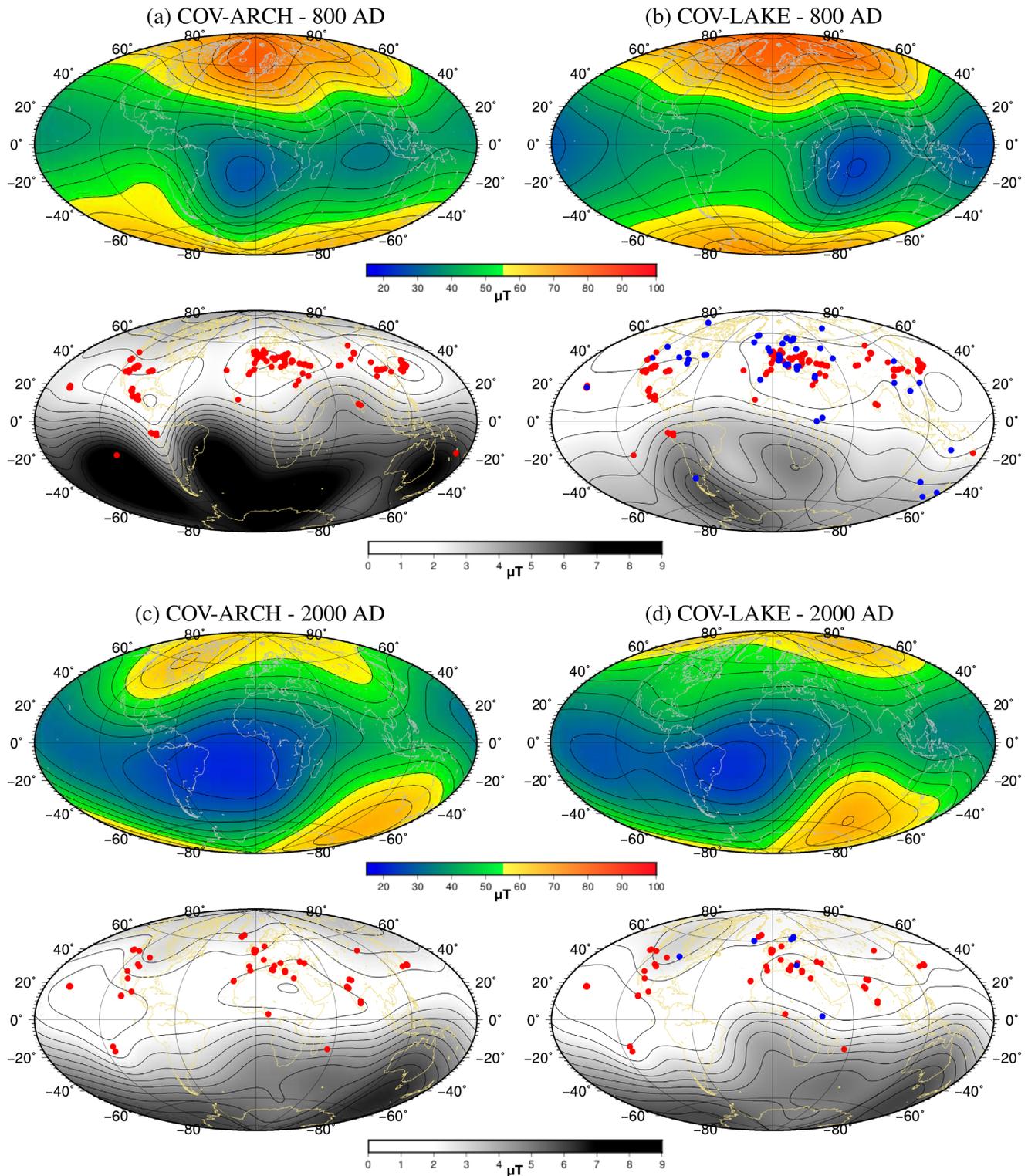


Figure 11. Intensity at the Earth's surface. Ensemble average model (in colour scale, contours are every $5 \mu\text{T}$) and dispersion within the ensembles (in grey scale, contours are every $0.5 \mu\text{T}$) at epochs 800 AD (top) and 2000 AD (bottom) for COV-ARCH (left) and COV-LAKE (right). Data present at each epoch (± 50 yr) are indicated with red and blue dots for archeomagnetic and sediment records, respectively.

(3) *Interdependence of data constraints*: However, a major difference between our global approach and regional studies concerns the (implicitly) assumed interdependence of the data constraints. Indeed, the prior for the regional PDF in Fig. 12 is obtained considering an axial dipole as a background, reducing much any potential

influence of directional data (see Hellio *et al.* 2014, appendix A). On the opposite, the iterative process involved in the global field modelling (see Section 2.3) correlates all data constraints together, since the Jacobian that enters the Kalman gain matrix is built upon the complex field of the current iteration. As such, directional data

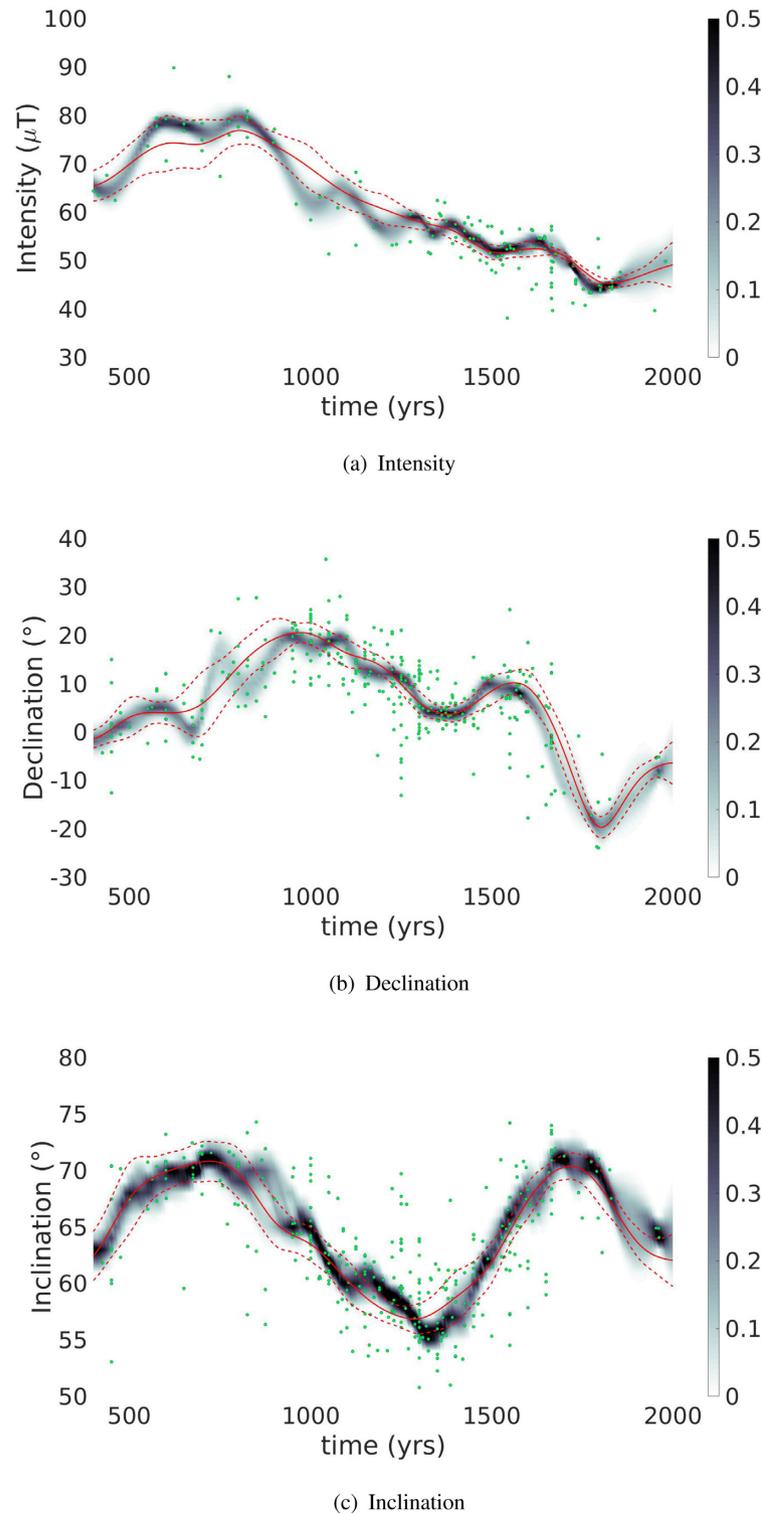


Figure 12. PDF of intensity, declination and inclination at Paris ($\lambda = 48.9^\circ\text{N}$, $\phi = 2.3^\circ\text{E}$), obtained with the regional method by Helliou *et al.* (2014) for directional and intensity data sets including all archeomagnetic data of the global data set, used to build COV-ARCH, in a radius of 1250 km around Paris. Model predictions from COV-ARCH are shown in red with their $\pm 3\sigma$ confidence interval.

around the considered region of the globe are likely to affect the global model (through the data kernels, e.g. fig. 13 in Constable 2007), and thus its local intensity predictions. For a given magnetic field model (especially with non-dipole contributions), these data are not independent. The corresponding curves for directions are shown in Figs 12(b) and (c). As for the intensity master curve, if the

general trend of global and regional PDFs correspond well, some centennial oscillations that appear in the regional curves are not fully recovered by the global model.

The above hypothesis calls for further work to either isolate incompatibilities between several archeomagnetic records (e.g.

through their dates), or best couple the information contained into directional and intensity records around a site (possibly accounting for some *a priori* magnetic field model). Nevertheless, it should not hide the fact that the presence of outliers tends to filter out oscillations at ≈ 250 yr periods, as witnessed with the synthetic tests in Fig. 5.

4.3.2 Importance of sediment series

We now illustrate with Fig. 13 the influence of sediment data on different global models at locations around which few (New Zealand) or no (Antarctica) archeomagnetic data are available. The closest records from Palmer Deep, Antarctica (Figs 13a and b) are more than 2700 km away. Over the past centuries, and despite the few data present at this location, our models recover well *gufm1* predictions, which lays within the model prediction uncertainties (of course no model do as well as CALS3k.4 that was constrained to reproduce *gufm1*). ASDI.FM (and to a lesser extent COV-LAKE) ensembles present more variations than the smoother CALS3k.4 and pfm9k.1b. Despite this, ensemble average predictions for directional data from all models agree rather well. The situation is quite different for intensity predictions, which are significantly stronger for pfm9k.1b in comparison with others models, prior to ≈ 1200 AD. Before 500 AD, COV-LAKE predictions are well within the observation uncertainties, slightly lower than intensities from all others models, and recovering the longer oscillations present in ASDI.FM. Nilsson *et al.* (2014) argued that CALS10k.1b underestimated the intensity for the last 2000 yr in South America, according too much importance to this specific marine record. This underestimation should also happen at the record site. However, we obtain a lower intensity prediction from COV-ARCH than from COV-LAKE, suggesting that this sediment core tends instead to increase F locally: another explanation must be found to explain the high values for F from pfm9k.1b.

Finally, we show in Fig. 13(c) the predictions of the models at the Mavora lake in New Zealand (Turner *et al.* 2015), contained in our data set but not accounted for when building ASDI.FM. It illustrates how one sediment core does change drastically the predictions, particularly in the southern hemisphere where the data distribution is sparse. This argue in favour of more sediment records in the southern hemisphere, or in areas poorly covered in the northern hemisphere, in order to significantly improve archeomagnetic models.

4.4 COV-ARCH and COV-LAKE, data sets for assimilation studies

We have seen that our method, provided reasonable data errors are assigned, produces realistic estimate of posterior uncertainties on Gauss coefficients. We show in Fig. 14, for some of the existing global models, the diagnostic $\Phi(l)$ defined in eq. (24). The relative dispersion for COV-ARCH and COV-LAKE increases monotonically with the degree to reach 100 per cent for $l = 5$ and $l = 6$, respectively. For the ensembles of A_FM and ASDI.FM models, Φ is much less sensitive to the spherical harmonic degree, providing significantly smaller uncertainties for $l \geq 3$. The relative dispersion within the ensemble of CALS3K.4b and pfm9k.1b models are rather similar, and for $l \leq 3$, it is about twice the one we obtain for COV-LAKE. With such regularized field models, Φ reaches a maximum at $l = 5$, then it starts decreasing to reach very low values at large degrees. This behaviour is due to the spatial damping used to ensure

the spectral convergence of these models, providing unrealistically small posterior errors towards small length-scales.

Field models Gauss coefficients may be used as data in archeomagnetic assimilation studies (e.g. Fournier *et al.* 2013). In this context, assessing realistic uncertainties on the g_l^m is crucial in order to obtain an unbiased estimate of the core state. Fig. 14 advocates in favour of using the ensemble of COV-ARCH and COV-LAKE models for re-analyses of centennial to millennial geomagnetic field changes.

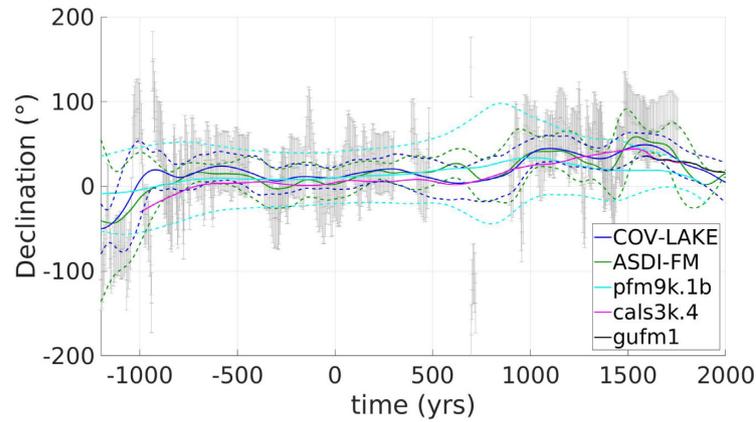
5 CONCLUDING REMARKS

In this study we introduce two ensembles of spherical harmonic magnetic field models spanning the past three millennia and truncated at degree 10: (i) COV-ARCH calculated from archeological and lava flow data and (ii) COV-LAKE calculated from the latter data set plus lake and marine sediment records. Instead of the usual spatio-temporal regularizations employed so far, we construct our prior information from spatial and temporal statistics of the geomagnetic field, as available from satellites, ground observatories and paleomagnetic measurements. This information is carried by temporal cross-covariance functions, onto which we project Gauss coefficients. We show that these efforts put on the modelling side provide statistically coherent posterior errors on the model parameters.

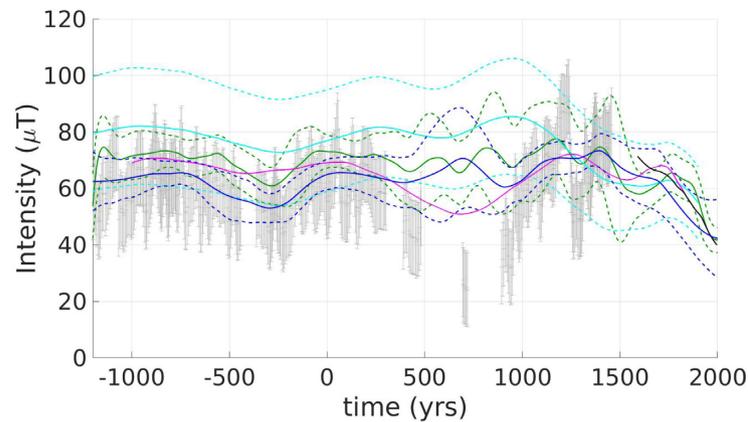
We observe a better agreement between COV-ARCH and COV-LAKE, in comparison with the corresponding A_FM and ASDI.FM ensembles, probably in link with the larger uncertainties reassigned to our sediment data set following Panovska *et al.* (2015). We also show a satisfying agreement between our models and models calculated from historical and contemporary data. It is worth noting that COV-LAKE Gauss coefficients better fit *gufm1* than COV-ARCH. This could be expected, since sediment data improve notably the spatial and temporal data distribution.

Our models suggest that the axial dipole (directly measured since the first absolute measurements in 1832) decays at a rate of $\approx 16 \pm 2$ nT.yr⁻¹ for COV-LAKE and $\approx 11 \pm 3$ nT.yr⁻¹ for COV-ARCH since 1700 AD. It is preceded by an era, from 1000 AD onwards, where the centennial trend of g_1^0 is weak. We find trace of the SAA back to 1800 AD in both COV-ARCH and COV-LAKE ensembles, preceded by another low-intensity patch above the Indian Ocean over 600–1400 AD. We observe several flux patches at low to mid-latitudes on the radial field at the CMB, which persist over the whole studied era, and look in COV-LAKE approximately symmetric with respect to a geomagnetic equator slightly shifted northwards. We confirm the existence of a westward drift of flux patches at a speed of ≈ 0.20 to 0.25° yr⁻¹ at the CMB. We show that predictions from our models at precise locations on ground agree well with the corresponding regional master curves, but that they fail at reproducing some of the centennial oscillations. This calls for a better consideration of the interplay between the constraints from the several field components.

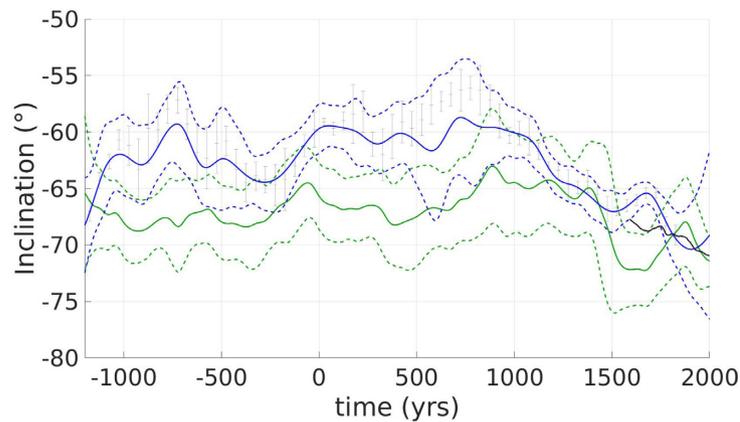
One further improvement to this work concerns the treatment of sediment records. Indeed, Nilsson *et al.* (2014) argue that significant uncertainties are to be linked with the estimation of age/depth models. To tackle this issue they perform timescale re-adjustments that we could account for in a next generation of model. Panovska *et al.* (2012) also pointed out systematic errors on declination, due to the recovery of cores. They suggest that only variations in declination should be used in global modelling. These issues may be



(a) Palmer Deep, Antarctica



(b) Palmer Deep, Antarctica



(c) Mavora lakes, New Zealand

Figure 13. (a and b) Model predictions in Antarctica ($\lambda = 64.9^\circ\text{S}$, $\phi = 295.8^\circ\text{E}$) for COV-LAKE (blue), ASDI-FM (green), pfm9k.1b (cyan), CALS3k.4 (magenta) and *gufm1* (black) with the 3σ uncertainties when provided (corresponding dotted lines). (c) Model predictions in New Zealand ($\lambda = 45.3^\circ\text{S}$, $\phi = 168.2^\circ\text{E}$) for COV-LAKE (blue) and ASDI-FM (green). Note that these latter data are used in the construction of COV-LAKE, but not in that of ASDI-FM.

overcome by adjusting the error model, or by modifying the observation operator—for example, involving scaling factors to be inverted for (Panovska *et al.* 2015), and/or considering sediment data as time-weighted averages. Having this in mind, one may consider extending backwards the global geomagnetic field reconstruction by applying the method proposed here to ancient sediment core. On

longer periods, some studies (Yiou *et al.* 1985) have also shown that geomagnetic fluctuations could lead to detectable variations in the production of cosmogenic isotopes, such as ^{14}C or ^{10}Be (Simon *et al.* 2016). These independent records may be used either as observations or to better constrain the temporal spectrum of the field at millennial timescales, and thus improve our prior information.

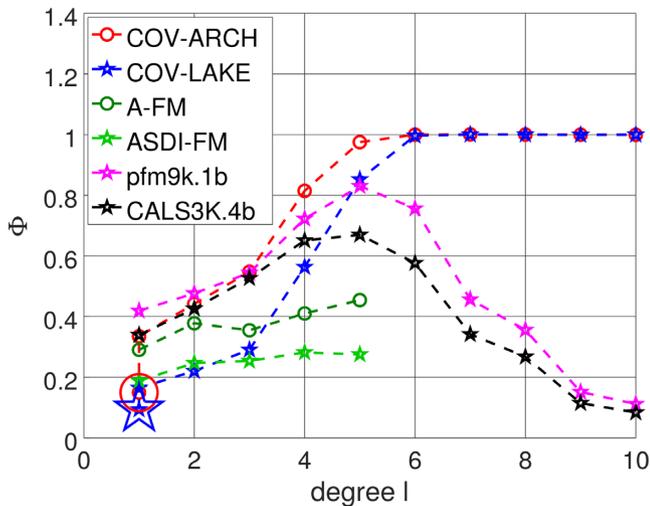


Figure 14. Comparison of the diagnostic Φ as a function of harmonic degree l , for the different models discussed in this study. Models calculated from archeological and lava data are represented with circles, and those including also sediment data with stars. We plot separately the diagnostic for the axial and equatorial dipoles of COV-ARCH and COV-LAKE (with larger markers for g_1^0).

Another extension of this work is to include more recent data. We choose here to use *gufm1*, and more generally all historical and contemporary data, as an external quality check. Instead, one could improve the resolution of models towards recent epochs by adapting the sampling of covariance functions and including historical data, or any measurements of the modern era. This would not only improve the knowledge of the axial dipole decay thanks to the constraint brought by archeomagnetic intensity data, but also propagate back in time the important information carried by historical data, through the employed cross-covariance functions.

Finally, on top of the potential use of alternative data types, we argue that more archeological and sediment data in poorly covered areas would improve archeomagnetic models. Some recent data not included yet in GEOMAGIA V.3 do not enter our models. Those concerning, in particular, the African continent (Osete *et al.* 2015; Tarduno *et al.* 2015; Kapper *et al.* 2017) or South America (Goguitchaichvili *et al.* 2012, 2015; Roperch *et al.* 2014, 2015; Poletti *et al.* 2016), once processed in GEOMAGIA, will considerably improve the constraint on archeomagnetic field models, and help target issues associated with the persistence over centuries of weak-intensity anomalies.

ACKNOWLEDGEMENTS

Dominique Jault, Claire Bouligand and Chris Finlay helped mature this project along discussions about the projection onto time cross-correlation functions. We thank Monika Korte and Sanja Panovska for their useful help and comments about the lake and marine sediment database, and Agnès Genevey for providing the recent archeointensity data from Greece and France. We thank Javier Pavón-Carrasco and Sanja Panovska for their thorough reviews that helped improve the quality of our manuscript. Numerical computations were performed at the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>) supported by the Rhône-Alpes region (GRANT CPER07 13 CIRA), the OSUG@2020 Labex (reference ANR10 LABX56) and the

Equip@Meso project (reference ANR-10-EQPX-29-01). ISTERre is part of Labex OSUG@2020 (ANR10 LABX56).

REFERENCES

- Bloxham, J. & Jackson, A., 1992. Time-dependent mapping of the magnetic field at the core–mantle boundary, *J. geophys. Res.*, **97**, 19 537–19 563.
- Bloxham, J., Gubbins, D. & Jackson, A., 1989. Geomagnetic secular variation, *Phil. Trans. R. Soc. Lond., A*, **329**, 415–502.
- 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.
- Brown, M.C., Donadini, F., Korte, M., Nilsson, A., Korhonen, K., Lodge, A., Lengyel, S.N. & Constable, C.G., 2015. Geomag50. v3: 1. general structure and modifications to the archeological and volcanic database, *Earth Planets Space*, **67**(1), 83, doi:10.1186/s40623-015-0232-0
- Buffett, B., 2015. Dipole fluctuations and the duration of geomagnetic polarity transitions, *Geophys. Res. Lett.*, **42**(18), 7444–7451.
- Constable, C., 2007. Centennial to millennial-scale geomagnetic field variations, eds Kono, M. & Schubert, G. in *Treatise on Geophysics*, Vol. 5, Elsevier, pp. 337–372.
- Constable, C., Korte, M. & Panovska, S., 2016. Persistent high paleosecular variation activity in southern hemisphere for at least 10 000 years, *Earth planet. Sci. Lett.*, **453**, 78–86.
- Constable, C.G. & Johnson, C., 2005. A paleomagnetic power spectrum, *Phys. Earth planet. Inter.*, **153**, 61–73.
- Currie, R.G., 1968. Geomagnetic spectrum of internal origin and lower mantle conductivity, *J. geophys. Res.*, **73**(8), 2779–2786.
- De Santis, A., Barraclough, D. & Tozzi, R., 2003. Spatial and temporal spectra of the geomagnetic field and their scaling properties, *Phys. Earth planet. Inter.*, **135**, 125–134.
- Domingos, J., Jault, D., Pais, M.A. & Manda, M., 2017. The South Atlantic Anomaly throughout the solar cycle, *Earth planet. Sci. Lett.*, **473**, 154–163.
- Dumberry, M. & Finlay, C.C., 2007. Eastward and westward drift of the Earth’s magnetic field for the last three millennia, *Earth planet. Sci. Lett.*, **254**(1), 146–157.
- Farquharson, C. & Oldenburg, D., 1998. Non-linear inversion using general measures of data misfit and model structure, *Geophys. J. Int.*, **134** 213–227.
- Finlay, C.C., 2008. Historical variation of the geomagnetic axial dipole, *Phys. Earth planet. Inter.*, **170**(1–2), 1–14.
- Finlay, C.C. & Jackson, A., 2003. Equatorially dominated magnetic field change at the surface of Earth’s core, *Science*, **300**(5628), 2084–2086.
- Finlay, C.C., Olsen, N., Kotsiaros, S., Gillet, N. & Toffner-Clausen, L., 2016. Recent geomagnetic secular variation from swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model, *Earth Planets Space*, **68**(1), 112, doi:10.1186/s40623-016-0486-1
- Fournier, A. *et al.*, 2010. An introduction to data assimilation and predictability in geomagnetism, *Space Sci. Rev.*, **155**(1–4), 247–291.
- Fournier, A., Nerger, L. & Aubert, J., 2013. An ensemble Kalman filter for the time-dependent analysis of the geomagnetic field, *Geophys. Geochem. Geosyst.*, **14**(10), 4035–4043.
- Gallet, Y., Genevey, A. & Courtillot, V., 2003. On the possible occurrence of ‘archaeomagnetic jerks’ in the geomagnetic field over the past three millennia, *Earth planet. Sci. Lett.*, **214**, 237–242.
- Gallet, Y., Hulot, G., Chulliat, A. & Genevey, A., 2009. Geomagnetic field hemispheric asymmetry and archeomagnetic jerks, *Earth planet. Sci. Lett.*, **284**(1), 179–186.
- Genevey, A., Gallet, Y., Jesset, S., Thébaud, E., Bouillon, J., Lefèvre, A. & Le Goff, M., 2016. New archeointensity data from French Early Medieval pottery production (6th–10th century AD). Tracing 1500 years of geomagnetic field intensity variations in Western Europe, *Phys. Earth planet. Inter.*, **257**, 205–219.
- Genevey, A., Kondopoulou, D., Pétridis, P., Aidona, E., Muller, A., Blondé, F. & Gros, J., 2017. New constraints on geomagnetic field intensity variations in the Balkans during the Early Byzantine period from ceramics unearthed at Thasos and Delphi, Greece, *J. Archaeological Sci.*

- Gillet, N., Jault, D., Finlay, C.C. & Olsen, N., 2013. Stochastic modelling of the Earth's magnetic field: inversion for covariances over the observatory era, *Geophys. Geochem. Geosyst.*, **14**, 766–786.
- Goguitchaichvili, A., Loponte, D., Morales, J. & Acosta, A., 2012. The archaeointensity of the Earth's magnetic field retrieved from Pampean ceramics (South America), *Archaeometry*, **54**(2), 388–400.
- Goguitchaichvili, A., Morales, J., Schavelzon, D., Vásquez, C., Gogorza, C.S., Loponte, D. & Rapalini, A., 2015. Variation of the Earth's magnetic field strength in South America during the last two millennia: new results from historical buildings of Buenos Aires and re-evaluation of regional data, *Phys. Earth planet. Inter.*, **245**, 15–25.
- Gubbins, D., 1975. Can the Earth's magnetic field be sustained by core oscillations? *Geophys. Res. Lett.*, **2**, 409–412.
- Gubbins, D., Jones, A.L. & Finlay, C.C., 2006. Fall in Earth's magnetic field is erratic, *Science*, **312**(5775), 900–902.
- Hellio, G., 2015. *Stochastic modeling of archeomagnetic measurements*. PhD thesis, Université Grenoble Alpes, France, Grenoble.
- Hellio, G., Gillet, N., Bouligand, C. & Jault, D., 2014. Stochastic modelling of regional archaeomagnetic series, *Geophys. J. Int.*, **199**(2), 931–943.
- Jackson, A., 2003. Intense equatorial flux spots on the surface of the Earth's core, *Nature*, **424**(6950), 760–763.
- Jackson, A. & Finlay, C., 2007. Geomagnetic secular variation and its applications to the core, eds Kanamori, H. & Schubert, G., in *Treatise on geophysics*, Vol. 5, Elsevier, 147–193.
- Jackson, A., Jonkers, A. R.T. & Walker, M.R., 2000. Four centuries of geomagnetic secular variation from historical records, *Phil. Trans. R. Soc. Lond. A*, **358**, 957–990.
- Kapper, L., Donadini, F., Serneels, V., Tema, E., Goguitchaichvili, A. & Morales, J.J., 2017. Reconstructing the geomagnetic field in West Africa: first absolute intensity results from Burkina Faso, *Sci. Rep.*, **7**, 45225, doi:10.1038/srep45225
- Korte, M. & Constable, C., 2011. Improving geomagnetic field reconstructions for 0–3 ka, *Phys. Earth planet. Inter.*, **188**, 247–259.
- Korte, M., Genevey, A., Constable, C.G., Frank, U. & Schnepf, E., 2005. Continuous geomagnetic field models for the past 7 millennia: 1. a new global data compilation, *Geophys. Geochem. Geosyst.*, **6**(2), doi:10.1029/2004GC000800
- Korte, M., Donadini, F. & Constable, C., 2009. Geomagnetic field for 0–3 ka: 2. a new series of time-varying global models, *Geophys. Geochem. Geosyst.*, **10**(6), doi:10.1029/2008GC002297
- Lesur, V., Wardinski, I., Baerenzung, J. & Holschneider, M., 2018. On the frequency spectra of the core magnetic field Gauss coefficients, *Phys. Earth planet. Inter.*, **276**, 145–158.
- 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.
- 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**, 229–248.
- Olsen, N., 2002. A model of the geomagnetic field and its secular variation for epoch 2000 estimated from ørsted data, *Geophys. J. Int.*, **149**(2), 454–462.
- Olson, P., Christensen, U. & Driscoll, P., 2012. From superchrons to secular variation: a broadband dynamo frequency spectrum for the geomagnetic dipole, *Earth planet. Sci. Lett.*, **319–320**, 75–82.
- Ossete, M.L., Catanzariti, G., Chauvin, A., Pavón-Carrasco, F.J., Roperch, P. & Fernández, V.M., 2015. First archaeomagnetic field intensity data from Ethiopia, Africa (1615±12 AD), *Phys. Earth planet. Inter.*, **242**, 24–35.
- Panovska, S., Finlay, C., Donadini, F. & Hirt, A., 2012. Spline analysis of Holocene sediment magnetic records: uncertainty estimates for field modeling, *J. geophys. Res.*, **117**(B2), doi:10.1029/2011JB008813
- Panovska, S., Finlay, C. & Hirt, A., 2013. Observed periodicities and the spectrum of field variations in Holocene magnetic records, *Earth planet. Sci. Lett.*, **379**, 88–94.
- Panovska, S., Korte, M., Finlay, C. & Constable, C., 2015. Limitations in paleomagnetic data and modelling techniques and their impact on Holocene geomagnetic field models, *Geophys. J. Int.*, **202**(1), 402–418.
- Pavón-Carrasco, F.J., Ossete, M.L., Torta, J.M. & De Santis, A., 2014. A geomagnetic field model for the Holocene based on archaeomagnetic and lava flow data, *Earth planet. Sci. Lett.*, **388**, 98–109.
- Piper, J., 1989. Paleomagnetism, in *Geomagnetism*, Vol. 3, pp. 31–61, ed. Jacobs, J.A., Academic press.
- Poletti, W., Trindade, R.I., Hartmann, G.A., Damiani, N. & Rech, R.M., 2016. Archeomagnetism of Jesuit missions in South Brazil (1657–1706 AD) and assessment of the South American database, *Earth planet. Sci. Lett.*, **445**, 36–47.
- Poletti, W., Biggin, A.J., Trindade, R.I., Hartmann, G.A. & Terra-Nova, F., 2018. Continuous millennial decrease of the Earth's magnetic axial dipole, *Phys. Earth planet. Inter.*, **274**, 72–86.
- Rasmussen, C.E. & Williams, C. K.I., 2006. *Gaussian Processes for Machine Learning*, The MIT Press.
- Roperch, P., Chauvin, A., Le Pennec, J.-L. & Lara, L.E., 2014. Paleomagnetic study of juvenile basaltic–andesite clasts from Andean pyroclastic density current deposits, *Phys. Earth planet. Inter.*, **227**, 20–29.
- Roperch, P., Chauvin, A., Lara, L.E. & Moreno, H., 2015. Secular variation of the Earth's magnetic field and application to paleomagnetic dating of historical lava flows in Chile, *Phys. Earth planet. Inter.*, **242**, 65–78.
- 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**, 983–1002.
- Simon, Q. *et al.*, 2016. Authigenic ¹⁰Be/⁹Be ratio signatures of the cosmogenic nuclide production linked to geomagnetic dipole moment variation since the Brunhes/Matuyama boundary, *J. geophys. Res.*, **121**(11), 7716–7741.
- Suttie, N., Holme, R., Hill, M.J. & Shaw, J., 2011. Consistent treatment of errors in archaeointensity implies rapid decay of the dipole prior to 1840, *Earth planet. Sci. Lett.*, **304**(1), 13–21.
- Tarduno, J.A., Watkeys, M.K., Huffman, T.N., Cottrell, R.D., Blackman, E.G., Wendt, A., Scribner, C.A. & Wagner, C.L., 2015. Antiquity of the South Atlantic Anomaly and evidence for top-down control on the geodynamo, *Nature commun.*, **6**, 7865, doi:10.1038/ncomms8865
- Terra-Nova, F., Amit, H., Hartmann, G.A., Trindade, R.I. & Pinheiro, K.J., 2017. Relating the South Atlantic Anomaly and geomagnetic flux patches, *Phys. Earth planet. Inter.*, **266**, 39–53.
- Turner, G., Howarth, J., de Gelder, G. & Fitzsimons, S., 2015. A new high-resolution record of Holocene geomagnetic secular variation from New Zealand, *Earth planet. Sci. Lett.*, **430**, 296–307.
- Walker, M.R. & Jackson, A., 2000. Robust modelling of the Earth's magnetic field, *Geophys. J. Int.*, **143**(3), 799–808.
- Yiou, F., Raisbeck, G., Bourles, D., Lorius, C. & Barkov, N., 1985. ¹⁰Be in ice at Vostok Antarctica during the last climatic cycle, *Nature*, **316**(6029), 616–617.
- Ziegler, L.B., Constable, C.G., Johnson, C.L. & Tauxe, L., 2011. PADM2M: a penalized maximum likelihood model of the 0–2 Ma palaeomagnetic axial dipole moment, *Geophys. J. Int.*, **184**(3), 1069–1089.

APPENDIX A: POSTERIOR UNCERTAINTY MAPS OF THE RADIAL FIELD AT THE CMB (SYNTHETIC EXPERIMENT)

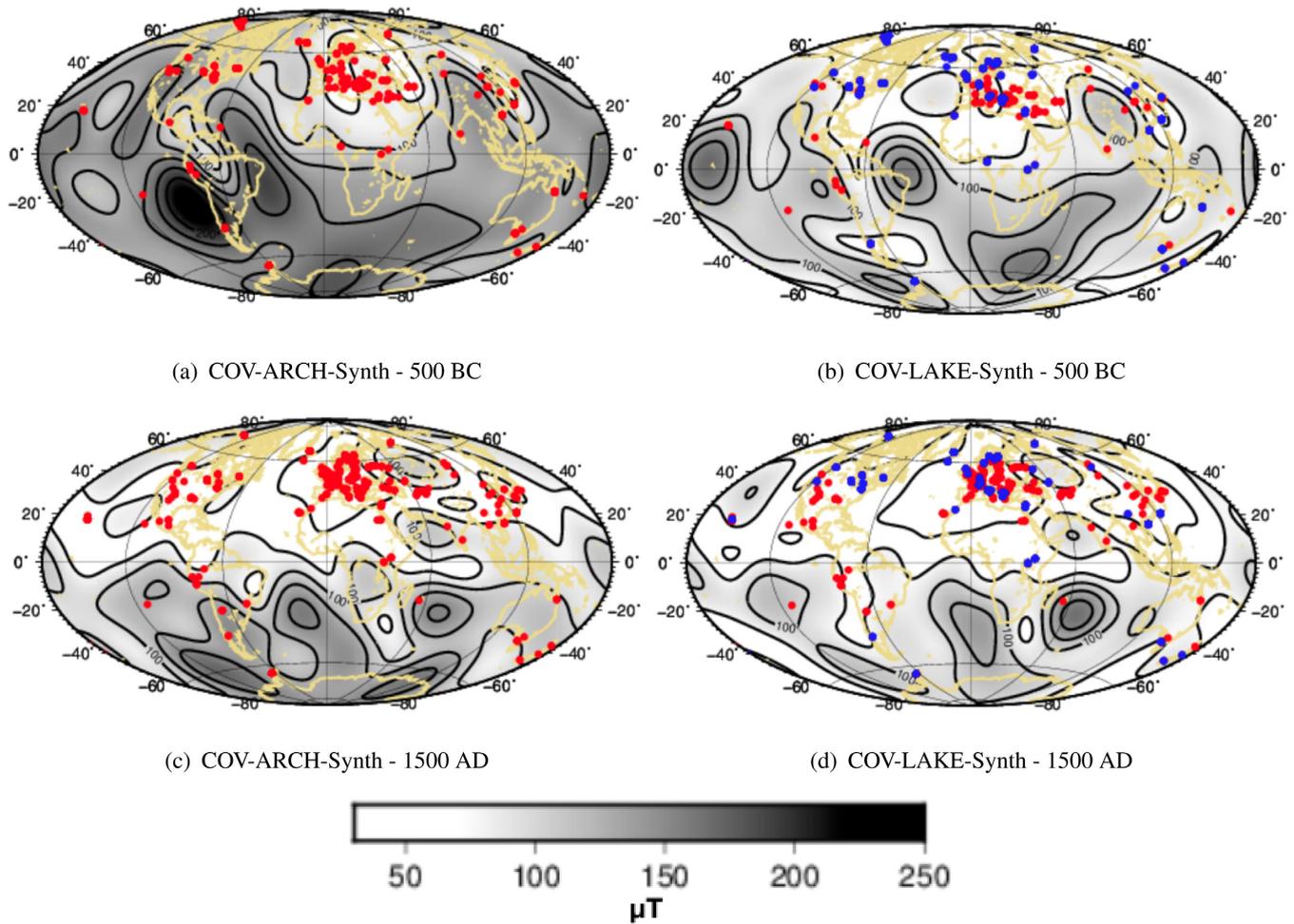


Figure A1. Standard deviation σ_{B_r} within the ensembles of COV-ARCH-Synth and COV-LAKE-Synth, for the radial component at the CMB, in 500 BC and 1500 AD. Contours are plotted every 25 μT . Red (resp. blue) dots: location of archeomagnetic and lava (resp. sediment) data at the corresponding epoch (± 50 yr).

APPENDIX B: RE-WEIGHTED DATA RESIDUALS

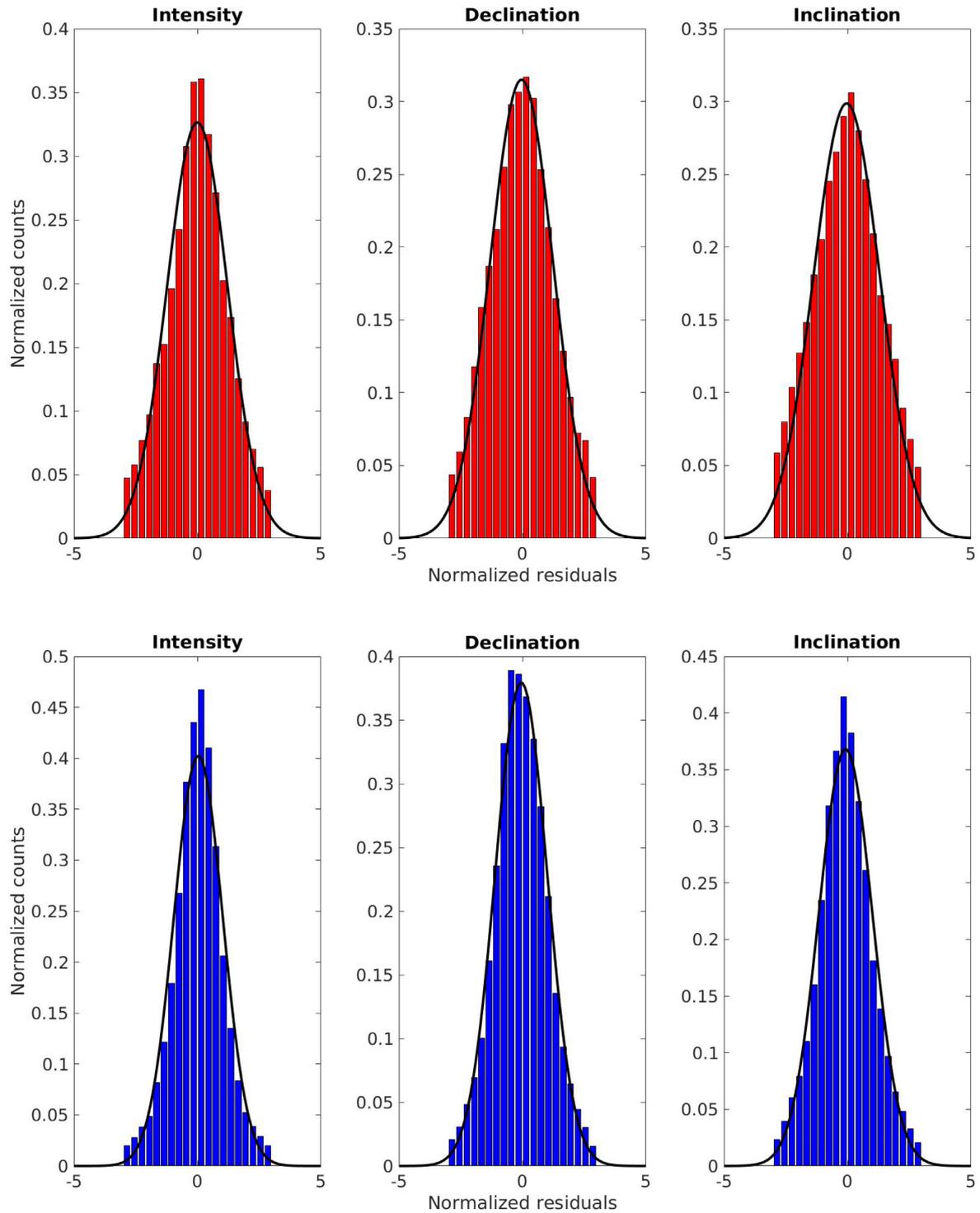


Figure B1. Distribution of normalized residuals for the COV-ARCH and COV-LAKE ensemble of field models. Superimposed in black, the Gaussian PDF determined by the mean and standard deviation of the residuals (see values in Table 2).

APPENDIX C: COMPARISON BETWEEN COV-LAKE AND CALS3K.4B SURFACE INTENSITY MAPS

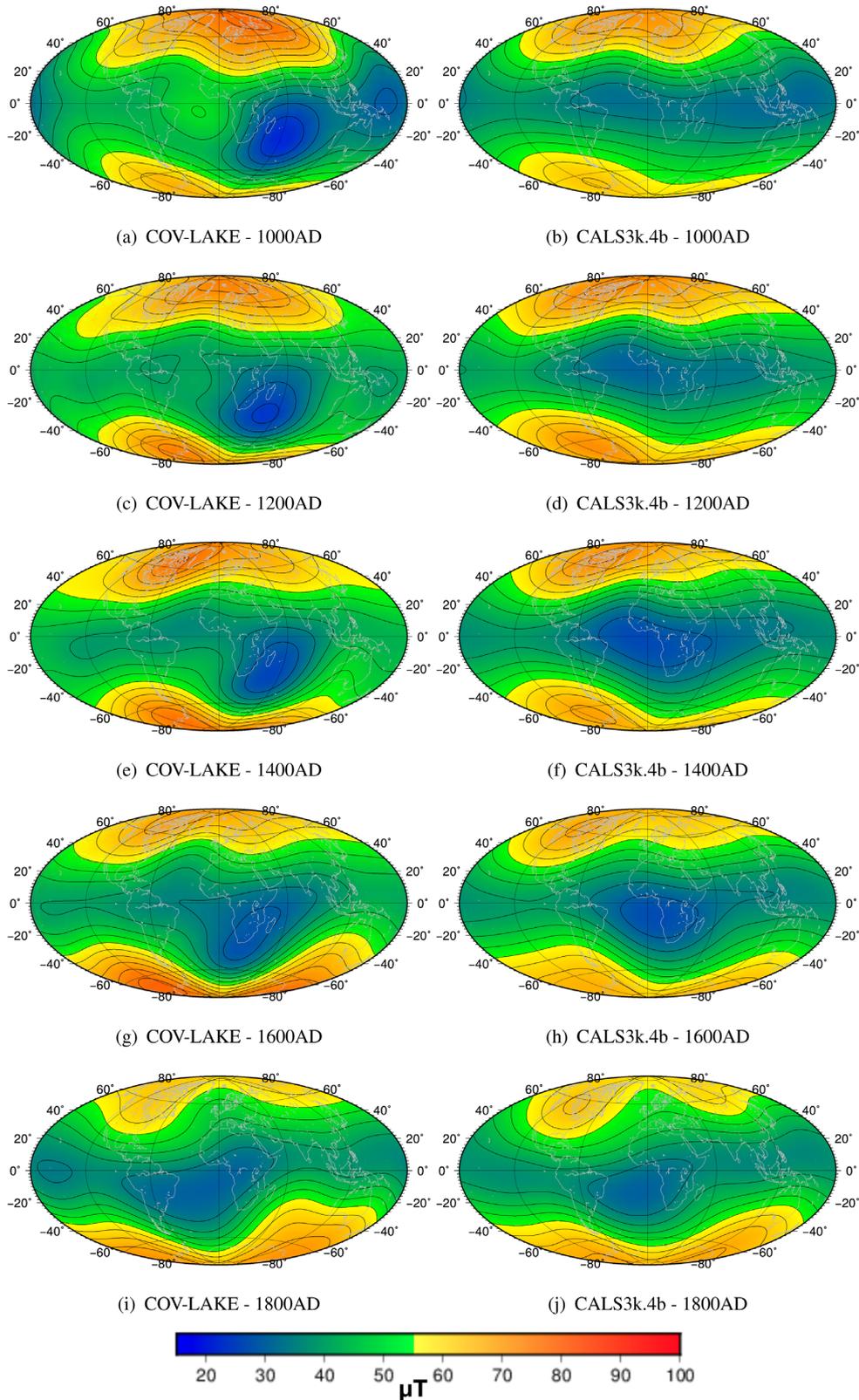


Figure C1. Intensity at the Earth's surface for COV-LAKE (left) and CALS3k.4b (right) every 200 yr from 1000 to 1800 AD (from top to bottom).

APPENDIX D: PREDICTIONS OF COV-ARCH AT WELL-SAMPLED LOCATIONS

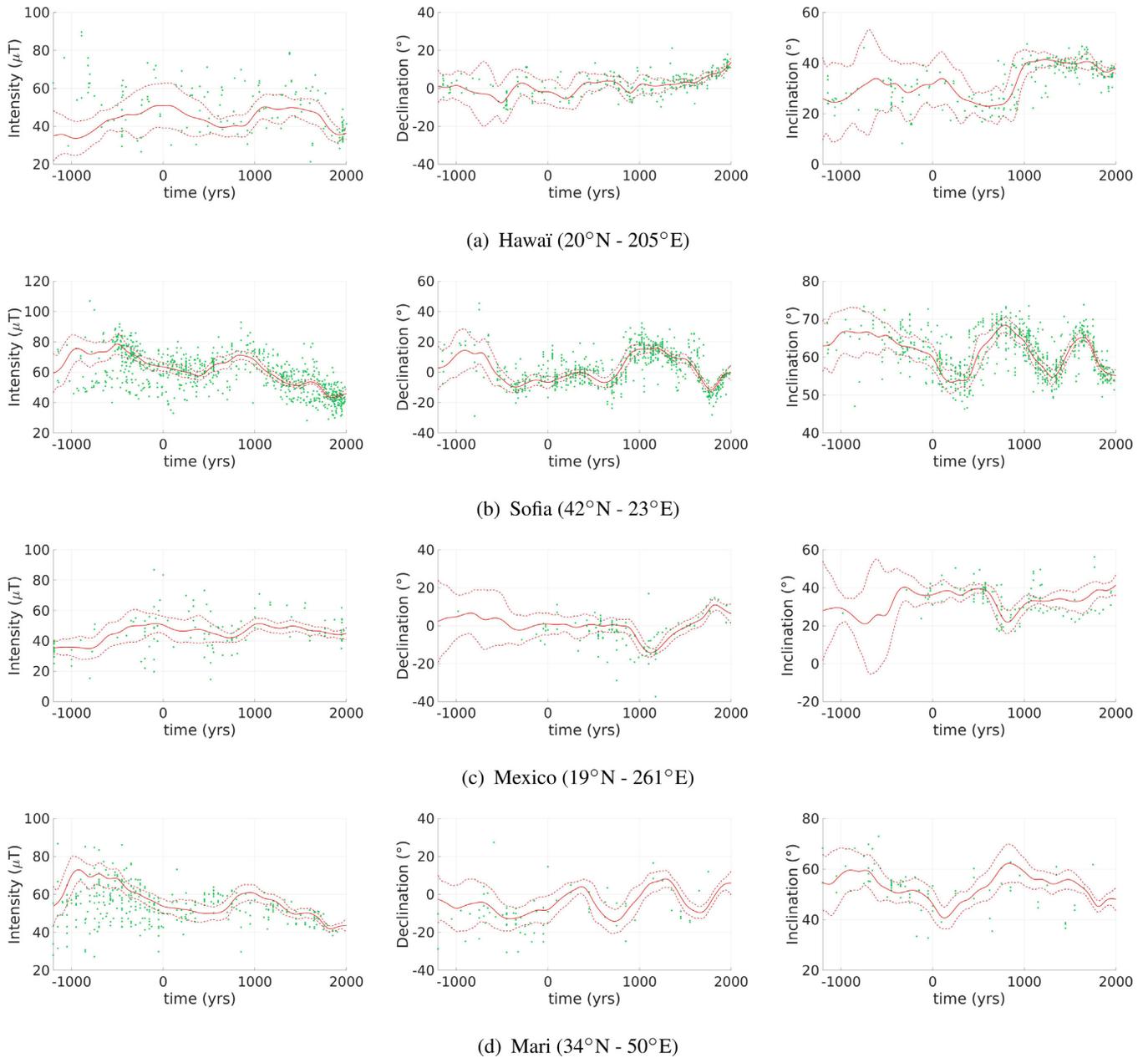


Figure D1. Comparison of predictions from COV-ARCH (average model in red) and archeomagnetic observations (green dots) at several locations. In red dotted lines the $\pm 3\sigma$ model predictions.