

# COV-OBS.x2 and ISTerre candidate models for IGRF-13

Loïc Huder<sup>1</sup>, Nicolas Gillet<sup>\*1</sup>, Chris C. Finlay<sup>2</sup>, and Magnus D. Hammer<sup>2</sup>

<sup>1</sup>*Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, IFSTTAR, ISTerre, F-38000 Grenoble, France*

<sup>2</sup>*DTU Space, National Space Institute, Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark*

2019

## 1 Introduction

We present below the ISTerre candidate models for the IGRF-13. We propose a main field (MF) model for 2015 and 2020, and a time average of the secular variation (SV) prediction over 2020-2025, together with their associated uncertainties. These are obtained from the COV-OBS.x2 field model and its posterior model error covariance matrix. The prediction for MF and SV result from a best linear unbiased estimate (BLUE) on COVOBS-x2 Gauss coefficient data.

The data and method used to derive the model and its predictions are presented in Sec. 2, while the model properties are illustrated in Sec. 3. Finally, we list in Sec. 4 the files containing the IGRF and DGRF field model coefficients and their associated uncertainties.

## 2 Construction of the COV-OBS.x2 field model

### 2.1 Geomagnetic data

Apart from modern satellite data, our data selection process follows closely the one used to construct the COV-OBS.x1 field model. Full details can be found in Gillet et al. (2013, 2015). Here we only briefly describe the new or updated aspects of the datasets. This only concerns satellite and observatory records over the past two decades or so:

---

<sup>\*</sup>Contact author: nicolas.gillet@univ-grenoble-alpes.fr

- Satellite data (see Sec. 2.1.2) are incorporated by means of virtual observatories (VO) built from the low Earth orbiting CHAMP and Swarm missions. These replace pointwise records in previous COV-OBS models.
- Ground observatory (GO) data (see Sec. 2.1.1) are considered through annual differences of 4-monthly means (instead of annual means in previous COV-OBS models).

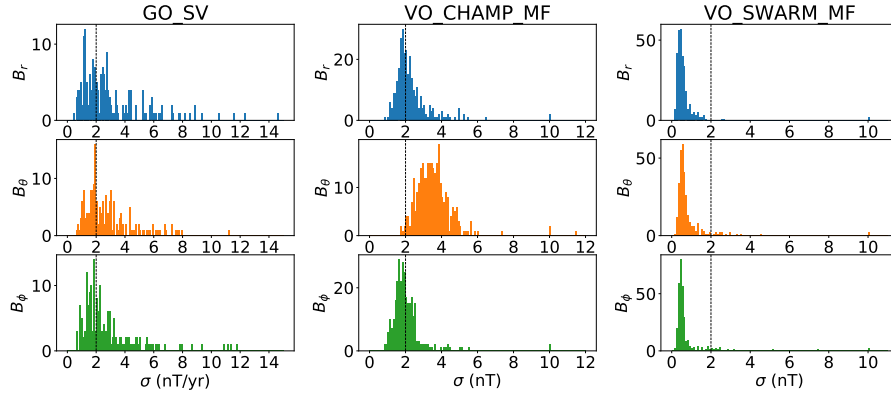
### 2.1.1 Ground observatories data

Up to June 1996 (included), we use the same dataset of ground observatories data as that used in COV-OBS.x1 (annual difference of annual means). For more recent epochs, we consider instead annual differences of 4-monthly means spanning August 1997 to March 2019. These revised means are computed from hourly mean values provided by the BGS database (Macmillan & Olsen, 2013), as described in Olsen et al. (2014). Being constructed upon data that sample all local times, the 4-monthly GO data are cleaned for the ionospheric field contribution (and its associated induced counterpart) using the CIY model (Sabaka et al., 2004). They are not cleaned for the magnetospheric contribution as this latter (and its induced counterpart) is co-estimated within the COV-OBS framework throughout the model time-span  $[t_s, t_e] = [1840, 2020]$  (see Sec. 2.2 and Sec. 2.4).

Standard deviations of error estimates ( $\sigma_{GO}^{obs}$ ) provided with the three components GO data are shown in Fig. 1. They are typically of the order of a few nT/yr (ranging from a fraction of nT/yr to above 10 nT/yr). We inflate these errors by an extra modelling error, of variance  $\sigma_{GO}^{mod2} = (2 \text{ nT/yr})^2$ , which accounts for the inability of COV-OBS.x2 (which is projected in time onto cubic B-splines) to fit SV changes at periods shorter than about 2 yrs (Pick et al., 2019). Considering that modelling and observation errors are independent, the resulting error budget is then  $\sigma_{GO}^2 = \sigma_{GO}^{mod2} + \sigma_{GO}^{obs2}$ .

### 2.1.2 Virtual observatories satellite data

Instead of the pointwise dataset used for COV-OBS.x1, we use  $N_o = 300$  VOs from the CHAMP and Swarm measurements. VOs consist of processed MF vector data distributed on an equal area grid at the altitude of 370 km for CHAMP and 490 km for Swarm. These are estimated every 4 months from March 2001 to November 2010 for CHAMP, and from November 2013 to July 2019 for Swarm. They are build from selected data (Sun at maximum  $10^\circ$  above the horizon and geomagnetically quiet conditions, see details in Barrois et al., 2018; Hammer, 2018). These data are cleaned from the crustal contribution, as estimated with the LCS model (Olsen et al., 2017), and from the ionospheric contributions as estimated from the CIY model (Sabaka et al., 2004). As for GO data, and contrary to what is done in Barrois et al. (2018), they are not cleaned for the magnetospheric contribution. We consider two sources of VO data error:



**Figure 1:** Histogram of observation errors  $\sigma_{GO}^{obs}$  and  $\sigma_{VO}^{obs}$  for GO and VO datasets, separated by spatial components. The dashed line represent the value of modeling error  $\sigma_{VO}^{mod}$  (see text for details).

- (i) ‘Observation’ uncertainties, of variance  $\sigma_{VO}^{obs^2}$ , as provided with the VO datasets. For each site, it is estimated as the magnitude of the residuals between VO series and the CHAOS-6 internal plus external predictions (Finlay et al., 2016). In practice  $\sigma_{VO}^{obs}$  are of the order of a few nT for CHAMP (and relatively larger on  $B_\theta$ ), and weaker than 1 nT for Swarm, as illustrated in Fig. 1.
- (ii) An extra error budget that covers unmodelled sources, of variance fixed to  $\sigma_{VO}^{mod^2} = (2 \text{ nT})^2$ . This aims at accounting for the inability of the COV-OBS models to capture signals of external source at periods shorter than 2 yrs (see Pick et al., 2019).

Considering these two error sources as independent, data error variances associated with VO data are thus  $\sigma_{VO}^2 = \sigma_{VO}^{obs^2} + \sigma_{VO}^{mod^2}$ . We acknowledge the fact that unmodelled sources certainly arise from spatially coherent structures, but accounting for spatial cross-covariances in unmodelled external field sources is out of the scope of the present study.

## 2.2 Parametrisation of the COV-OBS.x2 model

The construction of the COV-OBS.x2 field model is heavily based on the procedure described in Gillet et al. (2013). We recall here the main common points. In absence of electrical currents between observation points and the sources (here the Earth’s outer core, of radius  $c = 3485 \text{ km}$ , and the magnetosphere), the MF derives from a magnetic potential, or  $\mathbf{B} = -\nabla(V_i + V_e)$ , with  $V_i$  and  $V_e$  respectively the internal and external potentials.

$V_i$  is expanded on a spherical harmonic basis up to degree  $N_i = 14$ ,

$$V_i(r, \theta, \phi, t) = a \sum_{n=1}^{N_i} \left(\frac{a}{r}\right)^{n+1} \sum_{m=0}^n (g_n^m(t) \cos(m\phi) + h_n^m(t) \sin(m\phi)) P_n^m(\cos \theta), \quad (1)$$

with  $(g_n^m, h_n^m)$  the Schmidt semi-normalized internal Gauss coefficients of degree  $n$  and order  $m$ ,  $P_n^m$  the Legendre polynomials, and  $a = 6371.2$  km the Earth's radius.

The external potential  $V_e$  accounts for an external axial dipole field in dipole coordinates (plus its induced counterpart):

$$V_e(r, \theta, \phi, t) = r \sum_{m=0}^1 (\hat{q}_1^m(t) \cos(m\phi) + \hat{s}_1^m(t) \sin(m\phi)) P_1^m(\cos \theta). \quad (2)$$

We use the notations

$$\begin{cases} \hat{q}_1^m(t) = q_1^m(t) + \left(\frac{a}{r}\right)^3 g_1^{m\dagger}(t) \\ \hat{s}_1^m(t) = s_1^m(t) + \left(\frac{a}{r}\right)^3 h_1^{m\dagger}(t) \end{cases}, \quad (3)$$

with the Schmidt semi-normalized external  $(q_n^m, s_n^m)$  and induced  $(g_n^{m\dagger}, h_n^{m\dagger})$  Gauss coefficients, and

$$\begin{bmatrix} q_1^0 \\ q_1^1 \\ s_1^1 \end{bmatrix} (t) = q_{1d}^0(t) \mathbf{m}(t), \text{ with } \mathbf{m} = \frac{1}{\sqrt{g_1^{0^2} + g_1^{1^2} + h_1^{1^2}}} \begin{bmatrix} g_1^0 \\ g_1^1 \\ h_1^1 \end{bmatrix}. \quad (4)$$

A single coefficient,  $q_{1d}^0(t)$ , which describes the external dipole in dipole coordinates, is thus used to describe the evolution of the external field. The relation linking  $(g_n^{m\dagger}, h_n^{m\dagger})$  to  $(q_n^m, s_n^m)$  is detailed in Sec. 2.4.

All internal and external coefficients are projected in time onto order 4 cubic B-splines, with knots every 2 years, spanning the period 1838-2022. A L2 measure of the data misfit is employed, together with a  $3\sigma$  data rejection criterion. As a priori information in the inverse problem, we use cross-covariance associated with auto-regressive processes of order 2 (AR-2, see Sec. 2.3). Since ancient datasets contain some nonlinear data, and because of the relation (4), the model must be sought iteratively. This is processed using a Newton-Raphson algorithm.

Nevertheless, the parameterisation of the COV-OBS.x2 model differs from that of its predecessors:

- We consider an alternative AR-2 prior for the axial dipole (Sec. 2.3).
- We remove the contribution from the 20 nT background external dipole when estimating the field induced in the core (Sec. 2.4).

- Finally, the prediction over the period 2020-2025, where no data are available, is performed using a BLUE, considering as data sampled realisations of the COV-OBS.x2 Gauss coefficients (Sec. 2.5).

We further discuss these differences and their motivations below.

### 2.3 Stochastic prior for the axial dipole

In previous COV-OBS models, all internal field Gauss coefficients are considered as realisations of auto-regressive processes  $\varphi$  of order 2 (AR-2), governed by a stochastic differential equation of the form (e.g. Yaglom, 1962)

$$d\frac{d\varphi}{dt} + 2\omega d\varphi + \omega^2 \varphi dt = d\zeta(t), \quad (5)$$

with  $\zeta$  a Wiener process. Such two-parameter processes are characterized by auto-covariance functions of the form

$$C(\tau) = \sigma^2 (1 + \omega|\tau|) \exp(-\omega|\tau|), \quad (6)$$

with  $\sigma^2 = E((\varphi - \bar{\varphi})^2)$  the variance,  $\bar{\varphi} = E(\varphi)$  the average, and  $\omega^2 = E((\partial_t \varphi)^2) / \sigma^2$ .

For the sake of simplicity, we consider a variance of Gauss coefficients,  $\sigma_n^2 = E(g_n^{m2})$ , and a parameter  $\omega_n^2 = E((\partial_t g_n^m)^2) / \sigma_n^2$ , that depend only on the degree  $n$ . We use this formalism for all coefficients of degrees  $n \geq 2$ : for these we set parameters  $\sigma_n^2$  and  $\omega_n^2$  to the same values as in previous COV-OBS models (estimated from the MF and SV Lowes spectra obtained for a satellite field model in 2005, see Gillet et al., 2013). This description was found convenient as it is consistent with the -4 slope of the power spectrum density obtained for observatory series at periods from 5 to 70 years (De Santis et al., 2003), a feature confirmed later for Gauss coefficient series down to annual periods (Lesur et al., 2017). Indeed, the frequency spectrum of processes defined by Eq. (5),

$$P(f) = \frac{4\omega^3 \sigma^2}{[\omega^2 + (2\pi f)^2]^2}, \quad (7)$$

shows  $f^0$  dependence for low frequencies and  $f^{-4}$  dependence for high frequencies, with  $f$  the frequency. This concise description, based on only two parameters per harmonic degree, was validated by the analysis of geodynamo simulations for all coefficients but the axial dipole  $g_1^0$  (Bouligand et al., 2016).

Investigations on the frequency spectrum of this latter coefficient indeed show a  $f^{-2}$  dependence for intermediate frequencies from about  $10^5$  to  $10^2$  yrs periods. This was observed from both palaeomagnetic records (Constable & Johnson, 2005; Panovska et al., 2013) and dynamo calculations (Olson et al., 2012; Buffett & Matsui, 2015; Bouligand et al., 2016). To account for this effect, we modify the AR-2 prior for the axial dipole  $g_1^0$  in COV-OBS.x2, in comparison with previous COV-OBS models.

Following Hellio & Gillet (2018), we consider that the fluctuations of  $g_1^0$  (that is  $\tilde{g}_1^0(t) = g_1^0(t) - \bar{g}_1^0$  with  $\bar{g}_1^0$  the background axial dipole value) are governed by

a more general AR-2 process that obeys a three-parameter stochastic equation of the form

$$d\frac{d\varphi}{dt} + 2\chi d\varphi + \omega^2 \varphi dt = d\zeta(t), \quad (8)$$

where  $\chi$  and  $\omega$  are positive frequencies. The auto-covariance function for such a process is

$$C(\tau) = \frac{\sigma^2}{2\xi} \left( (\chi + \xi)e^{-(\chi - \xi)|\tau|} - (\chi - \xi)e^{-(\chi + \xi)|\tau|} \right), \quad (9)$$

with  $\xi^2 = \chi^2 - \omega^2$ . The associated frequency spectrum, given by

$$P(f) = \frac{4\chi\omega^2\sigma^2}{[\omega^2 - (2\pi f)^2]^2 + (4\pi\chi f)^2}, \quad (10)$$

indeed shows a  $f^{-2}$  dependence for intermediate frequencies (a domain depending on the values for  $\chi$  and  $\omega$ ). The  $f^0$  and  $f^{-4}$  dependences at respectively low and high frequencies are still present.

Our choice for the values of the three parameters that enter Eq. (9) slightly departs from that by Hellio & Gillet (2018) for the construction of the archaeomagnetic field models COV-ARCH and COV-LAKE (see Table 1). The background value  $\bar{g}_1^0 = -27000$  nT and the r.m.s.  $\sigma_{g_1^0} = 6000$  nT are estimated from the average and standard deviation of the axial dipole moment over the past 800 kyr (the Bruhnes chron), as estimated with the PADM2M (Ziegler et al., 2011) and SINT2000 (Valet et al., 2005) models (see also Buffett et al., 2013).

We choose  $\sigma_{g_1^0}^2 = E\left((\partial_t g_1^0)^2\right) = (10 \text{ nT/yr})^2$ , which associated with the above choice for  $\sigma_{g_1^0}$  comes down to  $\omega_{g_1^0}^{-1} = 600$  yr. The transition period between the spectrum ranges showing  $f^{-2}$  and  $f^0$  slopes is  $\tau_s \simeq 4\pi\chi/\omega^2$  (Hellio & Gillet, 2018). According to the analysis by Buffett & Matsui (2015), we set  $\tau_s = 29$  kyr (a more accurate estimate than the 100 kyr value proposed by Hellio & Gillet), so that the remaining parameter for the axial dipole prior is  $\chi_{g_1^0}^{-1} = 4\pi/(\tau_s\omega^2) \simeq 156$  yr.

For the two parameters that define the equatorial dipole statistics, governed by Eq. (5), we choose  $\sigma_1 = 4500$  nT and  $\omega_1^{-1} = 450$  yr, so that  $\sigma_{g_1^1}^2 = E\left((\partial_t g_1^1)^2\right) = (10 \text{ nT/yr})^2$  (with similar values for  $h_1^1$ ). This choice is consistent with the observations  $\sigma_{g_1^1} \simeq 0.75\sigma_{g_1^0}$  and  $\sigma_{g_1^1} \simeq \sigma_{g_1^0}$  made for instance with the low viscosity *Midpath* geodynamo simulation (Aubert et al., 2017).

In contrast with previous editions of the COV-OBS model, the prior variance and characteristic time for the equatorial dipole now depart from those used for the axial dipole parameters (see Table 1). Despite this change of configuration, we note that the a priori power authorized for dipole fluctuations at short periods is similar to that used for COV-OBS.x1, as illustrated in Fig. 2. Interestingly, this a priori power is significantly less than that used for constructing the COV-ARCH and COV-LAKE models at periods shorter than  $\approx 100$  yr (time-scales

Field model	$\sigma_{g_1^0}$	$\omega_{g_1^0}^{-1}$	$\chi_{g_1^0}^{-1}$	$\sigma_1$	$\omega_1^{-1}$
COV-ARCH	6000 nT	400 yr	20 yr	3800 nT	200 yr
COV-OBS.x2	6000 nT	600 yr	156 yr	4500 nT	450 yr
COV-OBS.x1	17300 nT	1014 yr	–	17300 nT	1014 yr

**Table 1:** Parameters used for the AR(2) process describing the statistics of axial and equatorial dipoles in the construction of COV-OBS.x2, compared with those used for the COV-ARCH model of Hellio & Gillet (2018), and the earlier editions COV-OBS.x1 (Gillet et al., 2015).

unconstrained with the archaeomagnetic database). This is the consequence of decreasing the cut-off period  $\tau_s$  while increasing  $\omega^{-1}$  (see above).

## 2.4 Parametrisation of the induced field

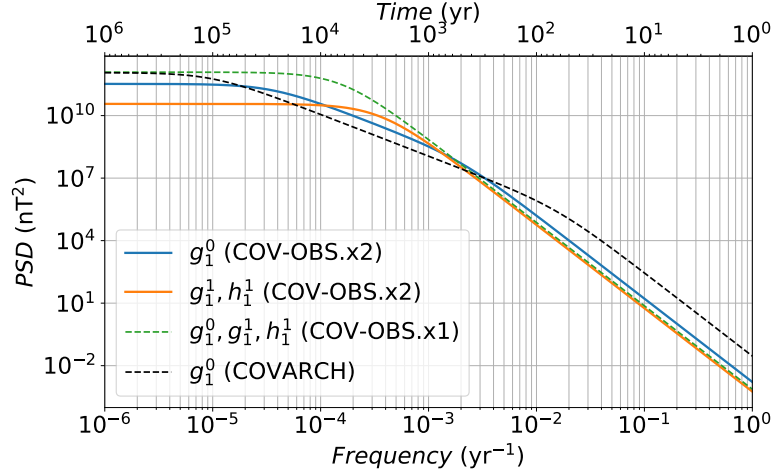
The induced field is anchored to the external field. We consider the core as a perfect conductor, an approximation reasonable since we model only field changes at periods longer than  $\approx 2$  years (see Fig. 1 in Olsen et al., 2005). In this framework, the induced field is simply computed by considering that the radial component of the induced field cancels that of the time-dependent inducing field at  $r = c$ . By differentiating Eq. (2) with respect to  $r$ , we obtain from Eq. (3)

$$\begin{bmatrix} g_1^{0\uparrow} \\ g_1^{1\uparrow} \\ g_1^{1\downarrow} \\ h_1^{1\uparrow} \end{bmatrix} (t) = Q_0 \begin{bmatrix} \tilde{q}_1^0 \\ \tilde{q}_1^1 \\ \tilde{q}_1^1 \\ \tilde{q}_1^1 \end{bmatrix} (t) = Q_0 \tilde{q}_{1d}^0 \mathbf{m}(t), \quad (11)$$

with  $Q_0 = \frac{1}{2} \left( \frac{c}{a} \right)^3 \simeq 0.082$ .  $\tilde{q}_{1d}^0 = q_{1d}^0(t) - \bar{q}_{1d}^0$  is the external field perturbation to the background value  $\bar{q}_{1d}^0 = 20$  nT. This latter approximately corresponds to the sum of Geocentric Solar Magnetospheric (GSM) and Solar Magnetic (SM) contributions to the magnetospheric dipole (see Maus & Lühr, 2005; Lühr & Maus, 2010; Olsen et al., 2014).

The above parametrisation slightly differs from previous COV-OBS models, where in Eq. (11)  $\tilde{q}_{1d}^0$  was replaced by  $q_{1d}^0$  (i.e. the inducing field contains the entire external dipole, even the stationary background). By reducing the core response to only the transient magnetospheric field, we shift in particular the induced axial dipole  $g_1^{0\uparrow}$  by  $\approx Q_0 \bar{q}_{1d}^0 \simeq 1.6$  nT (considering a dominant axial w.r.t. equatorial dipole), and consequently the core dipole  $g_1^0$  by the opposite value.

A shift of  $\approx 4$  nT was observed between  $g_1^0$  in COVOBS and most other IGRF candidates (Thébault et al., 2015a,b). The above change to the parametrisation of induced fields for the present model COV-OBS.x2 should reduce this shift down to  $\approx 2.4$  nT. Our investigations suggest that the remaining difference is most likely associated with the data selection embedded with the satellite data in the COVOBS framework (see Sec. 2.1.2). In models like CHAOS-6 (Finlay



**Figure 2:** PSD for the processes defining the a priori information used, when constructing the COV-OBS.x2, for the axial dipole (blue) and the equatorial dipole (orange) coefficients. Comparison with those used by (Hellio & Gillet, 2018) for the axial dipole (dashed black), and by (Gillet et al., 2013, 2015) for all dipole coefficients (dashed green).

et al., 2016), the SM external field is anchored to indices (such as the Ring Current index RC, see Olsen et al., 2014) that consider both calm and stormy magnetic conditions. The induced field is related to the external field through complex Q-factors (that depend on a mantle conductivity profile). It is thus estimated in the frequency domain before it is transformed back to the time-domain (Maus & Weidelt, 2004; Olsen et al., 2005). Constructed as such, it has a zero mean when averaged over all epochs (as it should be if the external signal has stationary properties, but see Velínský & Finlay, 2011). Our present external model being computed only from data selected over calm periods, the above 2.4 nT shift cannot be reduced within the COV-OBS framework.

## 2.5 Spline-free stochastic forecast of the geomagnetic field

The stochastic 5-year forecast performed with COV-OBS.x1, candidate model to IGRF-12, was performed by expanding the model time-span (and the support B-spline functions) up to 2020, that is 5 yr after the last available data at that time (Gillet et al., 2015). We see two drawbacks to this procedure. First, there exists a potential for instabilities close to endpoints, associated with the use of splines together with an uneven data coverage (e.g., Gillet et al., 2010b). Second, it involves a relatively heavy procedure to validate the prediction over past periods where the behaviour of the field is (to some extent) known. We proceed differently, and perform a BLUE, using



- as data the Gauss coefficients of the COV-OBS.x2 model sampled at a set of epochs,
- as data errors the posterior uncertainties as provided with this model (see Gillet et al., 2013, for the method),
- as prior information, cross-covariances associated with the stochastic processes presented in Sec. 2.3.

The prior information in the COV-OBS framework being independent from one coefficient to the other, predictions are operated separately for all Gauss coefficients (i.e. we neglect the cross-covariances between data errors). In details, the procedure is the following. For each  $(n, m)$  we generate from the COV-OBS.x2 spline model coefficients a vector  $\mathbf{y}^o$  that contains  $g_n^m$  values at  $N^o$  epochs  $t_j^o$  spanning  $[t_s^o, t_e^o]$  every  $\Delta t^o$ . Observation error variances  $\sigma_{g_n^m}^{o2}(t_j^o)$  for this coefficient are extracted from the COV-OBS.x2 posterior covariance matrix, and stored into a diagonal matrix  $\mathbf{R}_{yy}$ .

We wish to estimate a vector  $\mathbf{x}$  that contains analysed Gauss coefficients  $g_n^{ma}$  at  $N^a$  epochs  $t_j^a$  spanning  $[t_s^a, t_e^a]$  with  $t_e^a = t_e^o + 5$  yr, together with its associated uncertainties. To this purpose, we construct cross-covariance matrices  $\mathbf{C}_{xy}$ ,  $\mathbf{C}_{xx}$  and  $\mathbf{C}_{yy}$ , of sizes respectively  $N^a \times N^o$  and  $N^o \times N^o$ , whose elements are

$$\begin{cases} \mathbf{C}_{xy}(i, j) = E \left( (g_n^m(t_i^a) - \bar{g}_n^m) (g_n^m(t_j^o) - \bar{g}_n^m) \right) \\ \mathbf{C}_{xx}(i, j) = E \left( (g_n^m(t_i^a) - \bar{g}_n^m) (g_n^m(t_j^a) - \bar{g}_n^m) \right) \\ \mathbf{C}_{yy}(i, j) = E \left( (g_n^m(t_i^o) - \bar{g}_n^m) (g_n^m(t_j^o) - \bar{g}_n^m) \right) \end{cases} \quad (12)$$

In the above definitions, the background value, denoted by overlines, is non-zero for the axial dipole  $g_1^0$  only (see Sec. 2.3).

The model  $\mathbf{x}$  thus results from the BLUE as

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{C}_{xy}(\mathbf{C}_{yy} + \mathbf{R}_{yy})^{-1}(\mathbf{y}^o - \bar{\mathbf{y}}) = \bar{\mathbf{x}} + \mathbf{K}_{xy}(\mathbf{y}^o - \bar{\mathbf{y}}), \quad (13)$$

where  $\mathbf{K}_{xy}$  is the Kalman gain matrix, and  $\bar{\mathbf{x}}$  (resp.  $\bar{\mathbf{y}}$ ) is a vector of size  $N^a$  (resp.  $N^o$ ) filled with the background value  $\bar{g}_n^m$ . Cross-covariances of the uncertainties on the analysed vector  $\mathbf{x}$  are then given by the posterior covariance matrix

$$\mathbf{R}_{xx} = \mathbf{C}_{xx} - \mathbf{K}_{xy}\mathbf{C}_{xy}^T. \quad (14)$$

We refer to, for instance, Rasmussen & Williams (2006) for a general framework on the above parameter estimation. To sample the dispersion of  $\mathbf{x}$ , an ensemble of  $k$  realisations is generated from the Choleski decomposition of  $\mathbf{R}_{xx}$  (see Gillet et al., 2013).

### 3 Analysis of the COV-OBS.x2 model

#### 3.1 Statistics on the COV-OBS.x2 prediction errors

We provide in Table 2 some statistics concerning the COV-OBS.x2 misfits and biases to the new (GO and VO) data sets, separately for all three components.

Dataset	$N_o$	$\mathcal{M}^*$			$\mu^*$			$\mathcal{M}$			$\mu$		
		$X$	$Y$	$Z$	$X$	$Y$	$Z$	$X$	$Y$	$Z$	$X$	$Y$	$Z$
GO_SV	21060	1.27	0.79	1.12	-0.01	-0.00	-0.03	4.54	2.86	4.26	-0.01	-0.01	-0.14
VO_CHAMP_MF	19771	1.29	1.06	1.03	0.01	0.05	0.08	3.36	2.62	2.52	0.07	0.12	0.23
VO_SWARM_MF	15122	1.30	0.87	0.93	0.05	-0.06	0.10	3.04	2.08	2.07	0.17	-0.09	0.23
Total	55953	1.28	0.91	1.04	0.01	0.00	0.04	3.78	2.59	3.19	0.07	0.01	0.09

**Table 2:** Errors statistics for the GO and VO datasets integrated in COV-OBS.x2: accepted number of data  $N_o$ , dimensionless L2 data misfit  $\mathcal{M}^*$  and bias  $\mu^*$ , and dimensional L2 data misfit  $\mathcal{M}$  and bias  $\mu$ . Dimensional misfits and biases are in units of nT for VOs, and nT/yr for GO.

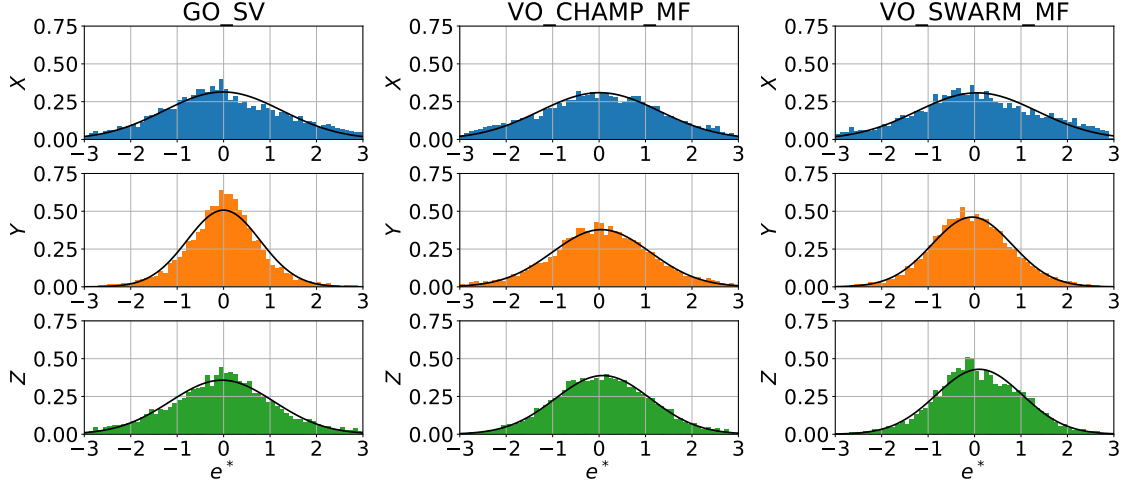
We consider the normalized L2 data misfit  $\mathcal{M}^*$  and bias  $\mu^*$ , as well as the dimensional L2 data misfit  $\mathcal{M}$  and bias  $\mu$ . For all three datasets and all components, no significant bias is found, as all normalized biases  $\mu^*$  remain relatively close to zero. Normalized L2 misfits are reasonably close to unity (in average slightly weaker on  $Y$  for GOs and Swarm data, and bit larger on  $X$  for all three data sources). Dimensional misfits, typically a few nT (or nT/yr) on all three components, are a bit larger on  $X$  (and on  $Z$  for GOs). Dimensional averaged biases are for all components of all datasets less than 0.25 nT. We present in Fig. 3 the distribution of the normalised residuals (VO and GO) for all three components. These are reasonably close to Gaussian (although sometimes slightly more peaked, see for instance on  $Y$  for the GO SV data).

In order to illustrate biases and misfits given in Table 2, we show our model SV predictions on ground with three components series at Niemegk, and MF predictions at Swarm’s altitude. We highlight the fact that the spread within the ensemble of COV-OBS.x2 realisations represent uncertainties on Gauss coefficients filtered at periods longer than  $\approx 2$  yrs. By construction our model is unable to recover time changes on shorter time-scales. This should be accounted for when using our SV model and uncertainties in e.g. core flow reconstructions (e.g. Gillet et al., 2019), for instance by adding, upon COV-OVS.x2 uncertainties, errors associated with AR-2 model at unmodelled high frequencies. Their contributions will be important (relative to COV-OBS.x2 uncertainties) mainly for large length-scales SV patterns over the well documented recent epochs.

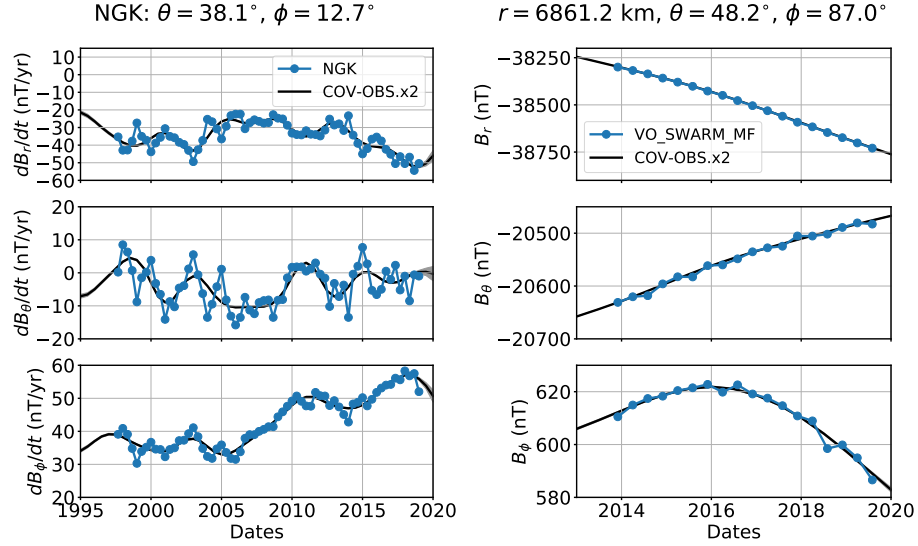
## 3.2 COV-OBS.x2 and its uncertainties

### 3.2.1 Validation of the 5-yr forecast

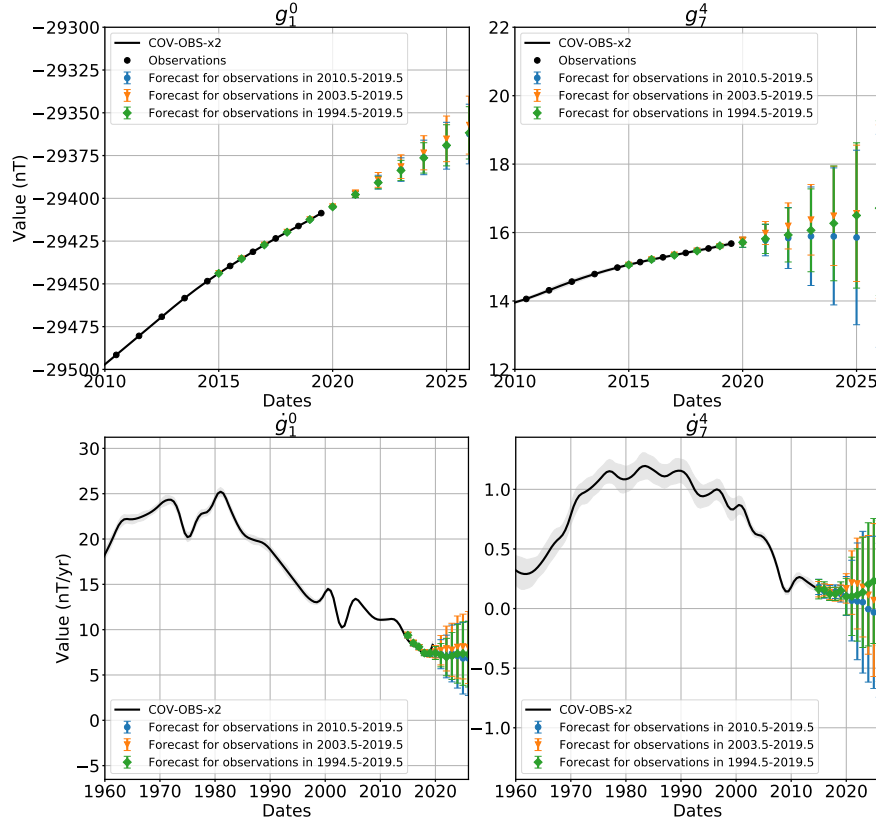
We now proceed to the model evaluation using the method described in Sec. 2.5. We first test the effect of the observation period  $[t_s^o, t_e^o]$  on the Gauss coefficient forecast. We perform several forecasts, with varying  $t_s^o - t_e^o$  from 6 to 75 yr with  $\Delta t^o = 1$  yr. Each time the forecast is evaluated between 2015 and 2025 (every  $\Delta t^p = 1$  year), and with  $t_e^p = 2019.5$ . The SV forecast is obtained by first differentiating MF prediction series. We give in Fig. 5 the MF and SV forecasts



**Figure 3:** Histograms of the normalised data misfit for the GO and VO datasets (separated all three components). In black is the normalised Gaussian curve  $\mathcal{N}(\mu, \sigma)$  with  $\mu$  the mean and  $\sigma$  the standard deviation of the normalised misfits (see Table 2).



**Figure 4:** Left: SV predictions from COV-OBS.x2 for the three geocentric components at the Niemegk (NGK) observatory. Right: MF predictions from COV-OBS.x2 for the three geocentric components at one example of Swarm VO.



**Figure 5:** 5-years forecasts using  $\Delta t^o = 1$  yr, for different observation periods, compared with COV-OBS.x2 in black (with in grey shaded area the associated  $\pm\sigma$  uncertainties), for the MF (top) and the associated SV (bottom) of Gauss coefficients  $g_1^0$  (left) and  $g_7^4$  (right). Errorbars represent the dispersion ( $\pm\sigma$ ) with the ensemble of forecasts.

of  $g_1^0$  and  $g_7^4$ , together with their associated dispersion. Our investigations show that the length of the observation period has a relative little impact on the resulting forecast and its associated spread. This is certainly due to the nature of the employed AR-2 stochastic processes, which have memory over only two successive epochs: the correlation functions that enters matrices  $C_{xy}$  and  $C_{yy}$  play a major role on our forecast estimates and their associated dispersions, which then evolves  $\propto (t - t_e^o)^2$ .

To test the validity of the forecast as presented in Sec. 2.5, we test it over ancient periods covered by COV-OBS.x2 datasets. Drawing upon the above conclusion,  $N^o = 17$  observation epochs are used, with a timestep  $\Delta t^o = 1$  yr. This results in observations covering a timespan of 17 yrs, used to build 5-yr forecasts. We show in Fig. 6 the obtained MF and SV predictions for Gauss

coefficients  $g_1^0$  and  $g_7^4$ , and for three observation periods ending at  $t_e^o = 2000$ , 1931 and 1913, periods characterized by different behaviors in particular of the axial dipole. Again, the SV forecast is obtained by first differentiating MF prediction series. For all Gauss coefficients but the axial dipole, MF and SV COV-OBS.x2 average model stay within  $\pm\sigma$  of the forecast spread whatever the observation period chosen. For  $g_1^0$  the forecast spread must be extended to about  $\pm 2\sigma$  for epochs showing intense and monotonous SV trends. We thus consider that the stochastic forecasts are consistent with the COV-OBS.x2 past evolution. This validates the forecast by BLUE as set-up in Sec. 2.5.

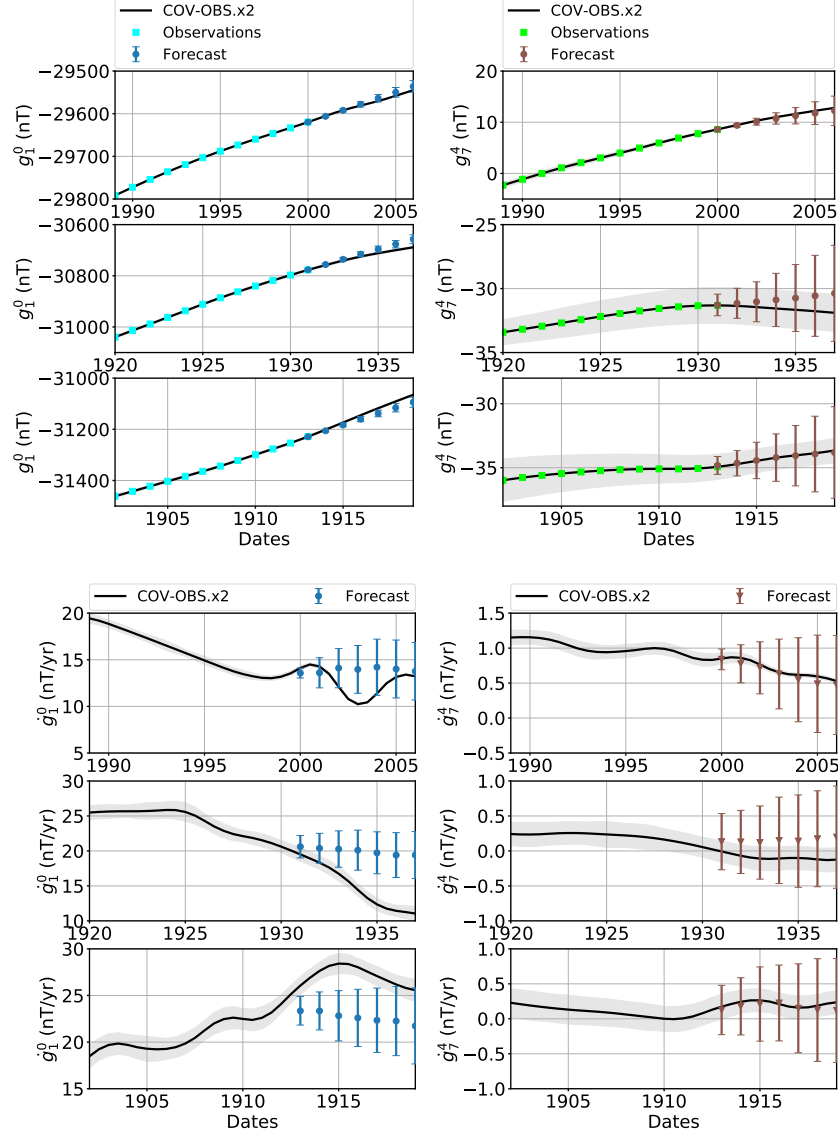
### 3.2.2 COV-OBS.x2 diagnostics

We present in Fig. 7 the MF and SV Lowes spectra obtained from the predictions (see Sec. 2.5) at the three epochs 2015, 2020, and 2025 of interest for the IGRF and DGRF model candidates (see Sec. 4). The MF dispersion spectrum is weaker in 2015 than in 2020 (at this latter epoch, data constraints are only from past epochs). It significantly increases for the prediction after 5 yr without observations, to reach values only slightly above those documented for the IGRF-12 candidate model based on COV-OBS.x1 (Gillet et al., 2015).

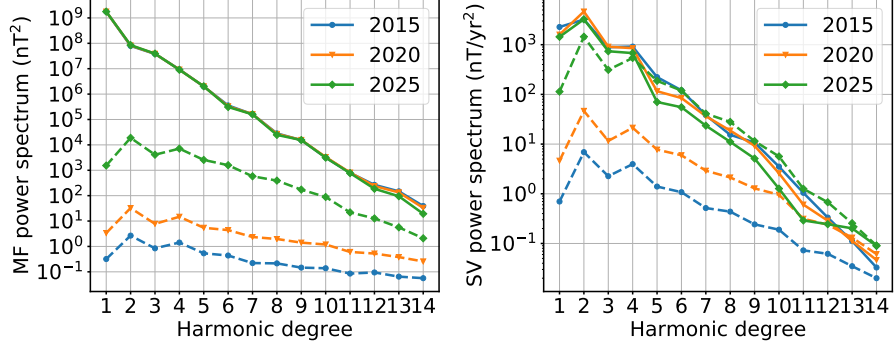
The SV spectrum for the ensemble average forecast decreases over time when no data is available. This reflects the fact that in average the stochastic prior drives the model back to the prior expectation. The spectrum of the SV spread gradually increases over time. Our present estimate of SV uncertainties is significantly larger than that documented in Gillet et al. (2015). After 5 yrs without observations, it reaches amplitudes similar to those of the average model for degrees  $n \geq 4$ , illustrating the inability (by construction) of the stochastic model to deterministically predict the magnetic field evolution.

We illustrate in Fig. 8 the time evolution of MF and SV Gauss coefficients for COV-OBS.x2. As observed by Gillet et al. (2015) with COV-OBS.x1, the evolutions are overall coherent with that of the CHAOS-6 model. Still we notice that

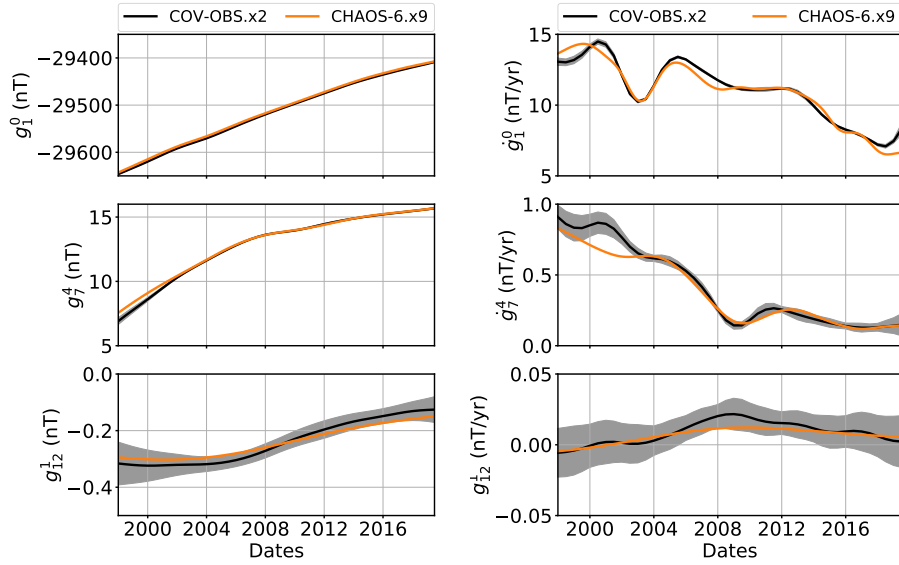
- (i) We do not capture some of the rapid changes (of period less than  $\approx 2$  yr) for the larger length-scales, due to the use of 2 yr knot spacing for the splines basis;
- (ii) Consequence of the employed stochastic prior, our solutions are less smooth at intermediate length-scales;
- (iii) The COV-OBS.x2 uncertainty estimates most often encompass the difference with CHAOS-6 at intermediate and small length-scales;
- (iv) COV-OBS.x2 uncertainties only represent errors on Gauss coefficients low-pass filtered (at periods longer than  $\approx 2$  yrs). These should be complemented by an error estimate that accounts for unmodelled rapid field changes, which will supplement COV-OBS.x2 formal errors at the largest length-scales.



**Figure 6:** 5-years forecasts using  $\Delta t^o = 1$  yr, for different observation periods, compared with COV-OBS.x2 in black (and in shaded grey its  $\pm\sigma$  dispersion), for the MF (top panels) and the associated SV (bottom panels) of Gauss coefficients  $g_1^0$  (left) and  $g_7^4$  (right). Errorbars represent the dispersion ( $\pm\sigma$ ) within the ensemble of forecasts.



**Figure 7:** MF (left) and SV (right) Lowes spectra at the Earth's surface for epochs 2015, 2020 and 2025, estimated with the method developed in Sec. 2.5. In dashed lines the respective spectra for the  $\pm\sigma$  spread within the ensemble of models.



**Figure 8:** MF (left) and SV (right) timeseries of  $g_1^0$  (top),  $g_7^4$  (middle) and  $g_{12}^1$  (bottom) for COV-OBS.x2 (black), compared with CHAOS-6.x9 (orange) between 1998 and 2020. The gray shaded areas represent the  $\pm\sigma$  dispersion within the ensemble of COV-OBS.x2 models.

## 4 Files description

From the average COV-OBS.x2 spline field model and its associated posterior covariance matrix, we generate with the method detailed in Sec. 2.5 an ensemble of  $K = 100$  realisations of the MF and SV field coefficients,  $\{\mathbf{g}_k(t), \dot{\mathbf{g}}_k(t)\}_{k \in [1, K]}$ .

We denote by  $\langle X \rangle = \frac{1}{K} \sum_{k=1}^K X_k$  the ensemble average. We list below the files submitted for evaluation:

- **DGRF\_MF\_2015\_ISTerre.dat**

It contains our DGRF MF candidate. It is the ensemble average MF model in 2015, estimated as:

$$\mathbf{g}_{DGRF} = \langle \mathbf{g}(2015) \rangle . \quad (15)$$

The associated errorbars are estimated as the dispersion within the ensemble of  $K$  realisations:

$$\delta \mathbf{g}_{DGRF} = \sqrt{\frac{1}{K} \sum_{k=1}^K [\mathbf{g}_k(2015) - \langle \mathbf{g}(2015) \rangle]^2} . \quad (16)$$

- **IGRF\_MF\_2020\_ISTerre.dat**

It contains our IGRF MF candidate. It is the ensemble average MF model in 2020, estimated as:

$$\mathbf{g}_{IGRF} = \langle \mathbf{g}(2020) \rangle . \quad (17)$$

The associated errorbars are estimated as the dispersion within the ensemble of  $K$  realisations:

$$\delta \mathbf{g}_{IGRF} = \sqrt{\frac{1}{K-1} \sum_{k=1}^K [\mathbf{g}_k(2020) - \langle \mathbf{g}(2020) \rangle]^2} . \quad (18)$$

- **IGRF\_SV\_2020-2025\_ISTerre.dat**

It contains our IGRF SV candidate. It is the ensemble average SV model, averaged in time over 2020–2025:

$$\dot{\mathbf{g}}_{IGRF} = \frac{1}{5} \int_{2020}^{2025} \langle \dot{\mathbf{g}}(t) \rangle dt . \quad (19)$$

The associated errorbars are estimated as the dispersion within the ensemble of  $K$  realisations:

$$\delta \dot{\mathbf{g}}_{IGRF} = \sqrt{\frac{1}{K-1} \sum_{k=1}^K \left[ \frac{1}{5} \int_{2020}^{2025} (\dot{\mathbf{g}}_k(t) - \langle \dot{\mathbf{g}}(t) \rangle) dt \right]^2} . \quad (20)$$



## References

- Aubert, J. & Finlay, C. C., 2019. Geomagnetic jerks and rapid hydromagnetic waves focusing at earths core surface, *Nature Geoscience*, **12**(5), 393.
- Aubert, J., Gastine, T., & Fournier, A., 2017. Spherical convective dynamos in the rapidly rotating asymptotic regime, *J. Fluid Mech.*, **813**, 558–593.
- Barrois, O., Hammer, M. D., Finlay, C. C., Martin, Y., & Gillet, N., 2018. Assimilation of ground and satellite magnetic measurements: Inference of core surface magnetic and velocity field changes, *Geophys. J. Int.*, **215**(1), 695–712.
- 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**(2), 1142–1157.
- Buffett, B. & Matsui, H., 2015. A power spectrum for the geomagnetic dipole moment, *Earth Planet. Sci. Lett.*, **411**, 20–26.
- Buffett, B. A., Ziegler, L., & Constable, C. G., 2013. A stochastic model for palaeomagnetic field variations, *Geophys. J. Int.*, **195**(1), 86–97.
- Constable, C. & Korte, M., 2015. 5.09 - centennial- to millennial-scale geomagnetic field variations, *Treatise on Geophysics (Second Edition)*, **5**(09), 309–341.
- Constable, C. G. & Johnson, C., 2005. A paleomagnetic power spectrum, *Phys. Earth Planet. Int.*, **153**, 61–73.
- De Santis, A., Barraclough, D., & Tozzi, R., 2003. Spatial and temporal spectra of the geomagnetic field and their scaling properties, *Phys. Earth Planet. Int.*, **135**(2-3), 125–134.
- Finlay, C., Dumberry, M., Chulliat, A., & Pais, M., 2010. Short timescale core dynamics: theory and observations, *Space science reviews*, **155**(1-4), 177–218.
- Finlay, C. C., Olsen, N., Kotsiaros, S., Gillet, N., & Tøffner-Clausen, L., 2016. Recent geomagnetic secular variation from Swarm, *Earth, Planets and Space*, **68**(1), 1–18.
- Fournier, A., Hulot, G., Jault, D., Kuang, W., Tangborn, A., Gillet, N., Canet, E., Aubert, J., & Lhuillier, F., 2010. An introduction to data assimilation and predictability in geomagnetism, *Space science reviews*, **155**(1-4), 247–291.
- 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*.
- Gillet, N., Jault, D., Canet, E., & Fournier, A., 2010. Fast torsional waves and strong magnetic field within the Earths core, *Nature*, **465**(7294), 74–77.

- Gillet, N., Lesur, V., & Olsen, N., 2010. Geomagnetic core field secular variation models, *Space science reviews*, **155**(1-4), 129–145.
- Gillet, N., Jault, D., Finlay, C. C., & Olsen, N., 2013. Stochastic modeling of the Earth’s magnetic field: Inversion for covariances over the observatory era, *Geochem., Geophys. Geosyst.*, **14**(4), 766–786.
- Gillet, N., Barrois, O., & Finlay, C. C., 2015. Stochastic forecasting of the geomagnetic field from the COV-OBS.x1 geomagnetic field model, and candidate models for IGRF-12, *Earth, Planets and Space*, **67**(1), 71.
- Gillet, N., Huder, L., & Aubert, J., 2019. A reduced stochastic model of core surface dynamics based on geodynamo simulations, *Geophys. J. Int.*, **219**(1), 522–539.
- Hammer, M. D., 2018. *Local Estimation of the Earth’s Core Magnetic Field*, Ph.D. thesis, Technical University of Denmark (DTU).
- 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.
- Jackson, A., Jonkers, A. R., & Walker, M. R., 2000. Four centuries of geomagnetic secular variation from historical records, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **358**(1768), 957–990.
- Lesur, V., Wardinski, I., Baerenzung, J., & Holschneider, M., 2017. On the frequency spectra of the core magnetic field gauss coefficients, *Phys. Earth Planet. Int.*.
- Lühr, H. & Maus, S., 2010. Solar cycle dependence of quiet-time magnetospheric currents and a model of their near-earth magnetic fields, *Earth, planets and space*, **62**(10), 14.
- Macmillan, S. & Olsen, N., 2013. Observatory data and the Swarm mission, *Earth, Planets and Space*, **65**(11), 15.
- Maus, S. & Lühr, H., 2005. Signature of the quiet-time magnetospheric magnetic field and its electromagnetic induction in the rotating earth, *Geophys. J. Int.*, **162**(3), 755–763.
- Maus, S. & Weidelt, P., 2004. Separating the magnetospheric disturbance magnetic field into external and transient internal contributions using a 1d conductivity model of the earth, *Geophys. res. lett.*, **31**(12).
- Nataf, H.-C. & Schaeffer, N., 2015. Turbulence in the core, in *Treatise on Geophysics, Core Dynamics, 2nd edition*, edited by G. Schubert & P. Olson, vol. 8, pp. 161–181, Elsevier, Oxford.

- Olsen, N., Sabaka, T. J., & Lowes, F., 2005. New parameterization of external and induced fields in geomagnetic field modeling, and a candidate model for igrf 2005, *Earth, planets and space*, **57**(12), 1141–1149.
- Olsen, N., Lühr, H., Finlay, C. C., Sabaka, T. J., Michaelis, I., Rauberg, J., & Tøffner-Clausen, L., 2014. The CHAOS-4 geomagnetic field model, *Geophys. J. Int.*, **197**(2), 815–827.
- Olsen, N., Ravat, D., Finlay, C. C., & Kother, L. K., 2017. Lcs-1: a high-resolution global model of the lithospheric magnetic field derived from champ and swarm satellite observations, *Geophys. J. Int.*, **211**(3), 1461–1477.
- 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.
- 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.
- Pick, L., Korte, M., Thomas, Y., Krivova, N., & Wu, C.-J., 2019. Evolution of large-scale magnetic fields from near-earth space during the last 11 solar cycles, *J. Geophys. Res.: Space Physics*, **124**(4), 2527–2540.
- Rasmussen, C. E. & Williams, C. K. I., 2006. *Gaussian processes for machine learning*, The MIT Press.
- Sabaka, T. J., Olsen, N., & Purucker, M. E., 2004. Extending comprehensive models of the earth’s magnetic field with ørsted and champ data, *Geophys. J. Int.*, **159**(2), 521–547.
- Sabaka, T. J., Olsen, N., Tyler, R. H., & Kuvshinov, A., 2015. CM5, a pre-Swarm comprehensive geomagnetic field model derived from over 12 yr of CHAMP, Ørsted, SAC-C and observatory data, *Geophys. J. Int.*, **200**(3), 1596–1626.
- Schaeffer, N., Jault, D., Nataf, H.-C., & Fournier, A., 2017. Turbulent geodynamo simulations: a leap towards Earth’s core, *Geophys. J. Int.*, **211**(1), 1–29.
- Thébault, E., Finlay, C. C., Alken, P., Beggan, C. D., Canet, E., Chulliat, A., Langlais, B., Lesur, V., Lowes, F. J., Manoj, C., et al., 2015. Evaluation of candidate geomagnetic field models for igrf-12, *Earth, Planets and Space*, **67**(1), 112.
- Thébault, E., Finlay, C. C., Beggan, C. D., Alken, P., Aubert, J., Barrois, O., Bertrand, F., Bondar, T., Boness, A., Brocco, L., et al., 2015. International geomagnetic reference field: the 12th generation, *Earth, Planets and Space*, **67**(1), 79.

- Valet, J.-P., Meynadier, L., & Guyodo, Y., 2005. Geomagnetic dipole strength and reversal rate over the past two million years, *Nature*, **435**(7043), 802–805.
- Velínský, J. & Finlay, C., 2011. Effect of a metallic core on transient geomagnetic induction, *Geochem., Geophys. Geosyst.*, **12**(5).
- Yaglom, A., 1962. An Introduction to the Theory of Stationary Random Functions. Engle-wood Cliffs, *N. J.: Prentice-Hall*, p. 84.
- 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.