



**HAL**  
open science

# A reduced stochastic model of core surface dynamics based on geodynamo simulations

Nicolas Gillet, Loic Huder, Julien Aubert

► **To cite this version:**

Nicolas Gillet, Loic Huder, Julien Aubert. A reduced stochastic model of core surface dynamics based on geodynamo simulations. *Geophysical Journal International*, 2019, 219 (1), pp.522-539. 10.1093/gji/ggz313 . hal-02191420

**HAL Id: hal-02191420**

**<https://hal.science/hal-02191420>**

Submitted on 23 Jul 2019

**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.

# A reduced stochastic model of core surface dynamics based on geodynamo simulations

Nicolas Gillet<sup>\*,1</sup>, Loic Huder<sup>1</sup>, and Julien Aubert<sup>2</sup>

<sup>1</sup>Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD,IFSTTAR, ISTERre, 38000 Grenoble, France.

<sup>2</sup>Institut de Physique du Globe de Paris, Sorbonne Paris Cité, Univ. Paris Diderot, CNRS, 1 rue Jussieu, F-75005 Paris, France.

Geophysical Journal International; Accepted 2019 July 9. Received 2019 June 20; in original form 2019 March 19

## Abstract

We make use of recent geodynamo simulations to propose a reduced stochastic model of the dynamics at the surface of Earth's core. On decadal and longer periods, this model replicates the most energetic eigen directions of the geodynamo computation. Towards shorter time-scales, it proposes a compensation for weaknesses of these simulations. This model furthermore accounts for the signature, in the geomagnetic secular variation, of errors of representativeness associated with unresolved processes. We incorporate the reduced stochastic model into a geomagnetic data assimilation algorithm – an augmented state ensemble Kalman filter – and apply it to re-analyze magnetic field changes over the period 1880–2015. Errors of representativeness appear to be responsible for an important fraction of the observed changes in the secular variation, as it is the case in the dynamo simulation.

Recovered core surface motions are primarily symmetric with respect to the equator. We observe the persistence of the eccentric westward gyre over the whole studied era, and vortices that partly follow isocontours of the radial magnetic field at the core surface. Our flow models provide a good fit to decadal changes in the length-of-day, and predict its interannual variations over 1940–2005. The largest core flow acceleration patterns are found in an equatorial belt below  $10^\circ$  in latitude, and are associated with non-axisymmetric features. No systematic longitudinal drift of acceleration patterns is found, even over the past decades where satellite data are available. The acceleration of the high latitude westward jet in the Pacific hemisphere is, during the satellite era, a factor 5 smaller than previously reported, and its structure shows some evidence for equatorial asymmetry. The era

---

\*Corresponding author: [nicolas.gillet@univ-grenoble-alpes.fr](mailto:nicolas.gillet@univ-grenoble-alpes.fr)

of continuous satellite records provides enhanced contrast on the rapid core flow variations. The proposed assimilation algorithm offers the prospect of evaluating Earth-likeness of geodynamo simulations.

**Keywords**— GENERAL SUBJECTS: Core; GEOMAGNETISM and ELECTROMAGNETISM: rapid time variations; GEOMAGNETISM and ELECTROMAGNETISM: Dynamo: theories and simulations; GEOPHYSICAL METHODS: probabilistic forecasting; GEODESY and GRAVITY: Earth rotation variations.

## 1 Introduction

Thanks to recent advances in numerical computations, geodynamo models do not only produce Earth-like pictures of the static magnetic field, but also capture some of the physics responsible for interannual to decadal changes. For instance, torsional waves propagating outward, as seen with geophysical observations [Gillet et al., 2010], become ubiquitous in simulations [Wicht and Christensen, 2010, Teed et al., 2014, Schaeffer et al., 2017, Aubert, 2018]. At decadal periods, frequency spectra  $S(f)$  for the main field (MF) Gauss coefficients at the core surface in the most up-to-date computations present a spectral behaviour with frequency  $f$  that decays like  $S(f) \propto f^{-p}$ , with a spectral index  $p \approx 4$  [Bouligand et al., 2016, Aubert, 2018]. This is in agreement with spectra obtained from observatory geomagnetic data [Lesur et al., 2018]. Such a behaviour is characteristic of magnetic field evolution functions that are differentiable only once in time, and that present abrupt changes in the trend of their time derivative [Gillet et al., 2013]. The best-known illustration of such processes in the physical space is the occurrence of geomagnetic jerks in ground-based series [e.g. Manda et al., 2010].

Several possible physical explanations to the observed interannual magnetic field changes have been proposed, although with no definitive evidence. They may be the consequence of short period waves supported by locally strong magnetic fields and fast rotation, as witnessed by Aubert [2018] in numerical simulations. Another possibility is the presence of waves in a stratified layer at the top of the core, as advocated for by Chulliat et al. [2015]. We should nevertheless keep in mind that our perception of rapid changes is blurred, associated with the intrinsic limitations of the ill-posed geomagnetic inverse problem: towards small length-scales, only long period SV variations are accessible. As a consequence, resolved interannual SV fluctuations may simply reflect the spectral index  $p \simeq 4$  of the magnetic field [see Gillet, 2019]. In such a case, relating this particular behavior to the rotating magneto-hydro-dynamics (MHD) turbulence remains to be understood [see the discussion of the dynamics as a function of periods and length-scales by Nataf and Schaeffer, 2015]. To answer these questions, an interplay between geophysical observations and dynamical models is needed.

Data assimilation algorithms based on primitive (momentum, induction and heat) equations are possibly suited to re-analyse geomagnetic changes over long periods [Fournier et al., 2013, Sanchez, 2016]. Still, the Alfvén propagation speed remains too slow relative to the convective overturn velocity in computations, hindering a complete access to fast processes [Aubert, 2018]: indeed, the unrealistically low ratio of magnetic to kinetic energy found in simulations shifts the characteristic time-scale of Alfvén waves towards long periods [e.g. Schaeffer et al., 2017]. As a consequence, data assimilation tools that drive geodynamo models are currently unable to ingest rapid interannual changes such as those found with the past two decades of satellite records [Chulliat and Maus, 2014, Finlay et al., 2016b].

Alternative forward dynamical equations must then be sought. This is the purpose of quasi-geostrophic (QG) models that describe, in the equatorial plane, a coupled dynamics for the flow and quadratic quantities of the magnetic field [see Jault and Finlay, 2015]. By considering a two-dimensional core state, these have the advantage of significantly reducing the size of the unknown state vector (initially composed of the three-dimensional velocity, magnetic and codensity fields), in this severely ill-posed inverse problem. Such reduced deterministic models are currently under development [e.g. Maffei and Jackson, 2017] but have not yet reached an operational state.

We consider here an alternative route where we estimate, from geodynamo time series, the parameters of a reduced stochastic model of the core surface dynamics. Inference of core motion using spatial norms derived from geodynamo simulations has been introduced and iteratively refined in the recent years [Fournier et al., 2011, 2015, Aubert, 2013, 2014, 2015] but has been so far limited to the retrieval of snapshots in time that are decorrelated from each other. On the other hand, the introduction of a stochastic temporal description of core flows [Gillet et al., 2015b] proposes a framework to handle temporal correlations. In our previous work [Barrois et al., 2017, from now on referred to as BGA17] we have initiated an approach where the strengths of the spatial and temporal aspects are combined. BGA17 represent the electro-motive force (e.m.f.) involving unresolved magnetic and velocity fields (the so-called errors of representativeness) by means of an augmented state approach [e.g. Evensen, 2003]. They time-step the evolutions of both errors of representativeness and core surface motions using an order one auto-regressive (AR-1) process. We extend in the present study this approach by also considering statistics on flow increments between two epochs. This leads to two major changes in comparison with BGA17:

- (i) the eigenmodes of the AR-1 processes are no longer the spherical harmonic basis functions but a linear combination of these, which is based on the geodynamo physics.
- (ii) each eigenmode will be associated to a distinct time-scale. The need for a coexistence of short- and long-lived patterns to model the core surface evolution has indeed been recently put forward by Baerenzung et al. [2018].

We wish this way to avoid a model trajectory that relaxes too quickly towards an unlikely ‘most probable’ state (as occurred in BGA17 where the model relaxed to the geodynamo time average).

The possibility of considering time cross-covariances is offered by the outputs of advanced numerical dynamo simulations. Here, statistics are derived from the ‘Mid-Path’ dynamo [Aubert et al., 2017]. This simulation is run at high rotation rate, using hyper-diffusivity on the codensity and velocity fields at spherical harmonic degrees  $n \geq 30$ . This process allows it to preserve an Earth-like large-scale magnetic field geometry and to recover decadal field changes, while pushing parameters towards geophysical values [Ekman number  $E = 10^{-8}$ , magnetic Prandtl number  $P_m = 0.045$ , see Aubert, 2018]. For comparison, the ‘Coupled-Earth’ dynamo [Aubert et al., 2013] used in BGA17 was run at  $E = 3 \times 10^{-5}$  and  $P_m = 2.5$ , preventing any use of time-correlations at decadal periods in an operational framework.

The manuscript is organised as follow. In §2.1 we first illustrate the behaviour, in the Mid-Path dynamo, of the SV sources at large length scales. We next derive in §2.2 our stochastic forward model for the core surface dynamics, while the algorithm used to solve the inverse problem is detailed in §2.3. In §3 we apply our tool to geophysical Gauss coefficient data from the COV-OBS.x1 model [Gillet et al., 2015b] over the period 1880-2015. We present an analysis of the several sources of SV (§3.1), core flow fluctuations

around the gyre (§3.2), flow acceleration patterns (§3.3), and a focus on length-of-day variations and zonal motions (§3.4). Finally we propose in §4 some perspectives around this work.

## 2 Method

By construction, a stochastic model such as the one presented in BGA17 does not deterministically predict the SV trajectory. It instead proposes the evolution of a probability density function (PDF) for the core surface state, and by consequence for the SV. However, this PDF is centered on an average, which relaxes during any forecast period towards the stochastic model expectation. In BGA17, only decorrelated dynamo snapshots were available [decadal changes are not realistically modelled in the Coupled-Earth simulation by Aubert et al., 2013]. Therefore, a somewhat ad hoc restoring time scale was considered (the same for all flow components, for the sake of simplicity). As a consequence, the SV PDF was relaxing too fast towards the average flow (chosen to be the average of the dynamo run). We try here to avoid this drawback by allowing both short and long restoring time-scales depending on the spatial structure. It is made possible thanks to the rapid dynamics available with Mid-Path (see §2.1–2.2).

We use for any vectors  $\mathbf{x}$  and  $\mathbf{y}$  the following notations: the time average is  $\mathbf{x}_0 = E(\mathbf{x})$ , the perturbation to the average is  $\mathbf{x}' = \mathbf{x} - \mathbf{x}_0$ , and the cross-covariance matrix  $P_{xy} = E(\mathbf{x}'\mathbf{y}'^T)$  – with  $\mathbf{x}^T$  the transpose of  $\mathbf{x}$ . We also consider for any vector  $\mathbf{x}$  the auto-correlation matrix  $C_{xx} = \text{diag}(P_{xx})^{-1/2}P_{xx}\text{diag}(P_{xx})^{-1/2}$ , with  $\text{diag}(P_{xx})$  the diagonal matrix whose diagonal elements are identical to those of  $P_{xx}$ . Finally, we note  $P_{xy}^*$  empirical cross-covariances obtained from any Mid-Path geodynamo series  $\mathbf{x}^*(t)$  and  $\mathbf{y}^*(t)$ .

### 2.1 Contributions to the large-scale SV from numerical simulations

We consider the radial induction equation at the core–mantle boundary (CMB),

$$\frac{\partial B_r}{\partial t} = -\nabla_h \cdot (\mathbf{u}_h B_r) + \eta \nabla^2 B_r. \quad (1)$$

We use here the spherical coordinates  $(r, \theta, \phi)$ .  $B_r$  is the radial component of the magnetic field,  $\mathbf{u}_h$  the horizontal core motion, and  $\eta$  the magnetic diffusivity [e.g. Holme, 2015]. Since we have only access to the largest length-scales, we re-write equation (1) as

$$\frac{\partial \overline{B}_r}{\partial t} = -\overline{\nabla_h \cdot (\mathbf{u}_h \overline{B}_r)} + s_r + \eta \nabla^2 \overline{B}_r, \quad (2)$$

where overlines account for the projection onto large length-scales, and the term  $s_r = -\overline{\nabla_h \cdot (\mathbf{u}_h B_r)} + \overline{\nabla_h \cdot (\mathbf{u}_h \overline{B}_r)}$  stands for the nonlinear e.m.f. involving small-length-scale (subgrid) terms, projected onto large length-scales. We combine the latter two terms on the r.h.s. of (2) to construct errors of representativeness  $e_r = s_r + \eta \nabla^2 \overline{B}_r$ . We store in vectors  $\mathbf{b}$ ,  $\mathbf{u}$ ,  $\mathbf{s}$ ,  $\mathbf{e}$ ,  $\mathbf{d}$  and  $\mathbf{f}$  Schmidt semi-normalised spherical harmonic coefficients for respectively  $\overline{B}_r$ ,  $\overline{\mathbf{u}_h}$ ,  $s_r$ ,  $e_r$ , the diffusion term and the e.m.f. involving large length-scale fields only. In matrix form, equation (2) then becomes

$$\dot{\mathbf{b}} = \mathbf{A}(\mathbf{b})\mathbf{u} + \mathbf{s} + \mathbf{d} = \mathbf{f} + \mathbf{e}, \quad (3)$$

with  $\mathbf{e} = \mathbf{s} + \mathbf{d}$ . Matrix  $\mathbf{A}$  accounts for the Gaunt-Elsasser integrals [e.g. Whaler, 1986]. These are calculated using products of the magnetic field and the flow in the spatial domain, then back-projected onto the spherical harmonics basis [see Lloyd and Gubbins, 1990]. The several projections onto large length-scales are performed by truncation of the magnetic field, secular variation and flow vectors in the spectral domain. From now on, large length-scales MF and SV fields are truncated at spherical harmonic degree  $n_b = 13$ , while the core flow model is limited to degree  $n_u = 18$ . The corresponding vectors thus contain respectively  $N_b = n_b(n_b + 2)$  and  $N_u = 2n_u(n_u + 2)$  parameters. The choice for  $n_u$  corresponds to a trade-off, as  $n_u$  controls how much of the SV enters errors of representativeness. Having it too large (resp. low) means being over-optimistic (resp. pessimistic) on our ability to recover surface core flows. We want here  $n_u$  to at least encompass resolved core motions (which shortest wave-lengths correspond to  $n \approx 10$  to 12 in studies by BGA17 and GJF15, under respectively dynamo norm and QG spatial constraints), while avoiding aliasing issues that would show up with too low values.

We define the spatial spectrum for the kinetic energy at the CMB,

$$E_K(n) = \sum_{m=-n}^n \left[ T_n^{m2} + S_n^{m2} \right], \text{ with } (T_n^m, S_n^m) = \sqrt{\frac{n(n+1)}{2n+1}} (t_n^m, s_n^m) \quad (4)$$

and  $(t_n^m, s_n^m)$  the Schmidt semi-normalised surface toroidal and poloidal core flow coefficients respectively [Holme, 2015].  $(T_n^m, S_n^m)$  are stored in a vector  $\mathbf{v} = \mathbf{N}\mathbf{u}$ , of size  $N_u$ , with  $\mathbf{N}$  a diagonal normalization matrix filled according to equation (4). We define the SV Lowes spectrum at the core surface (of radius  $c = 3485$  km),

$$R(n) = (n+1) \left(\frac{a}{c}\right)^{2n+4} \sum_{m=0}^n \left( \frac{dg_n^m}{dt} \right)^2 + \left( \frac{dh_n^m}{dt} \right)^2, \quad (5)$$

where  $(g_n^m, h_n^m)$  are Schmidt semi-normalised Gauss coefficients at the Earth's surface (of reference radius  $a = 6371.2$  km). We note  $R_f(n)$ ,  $R_e(n)$ ,  $R_d(n)$  and  $R_s(n)$  the Lowes spectra for the fields stored in vectors respectively  $\mathbf{f}$ ,  $\mathbf{e}$ ,  $\mathbf{d}$  and  $\mathbf{s}$ , from which we define the ratios

$$\xi_f(n) = R_f(n)/R(n), \quad \xi_e(n) = R_e(n)/R(n), \quad \xi_d(n) = R_d(n)/R(n), \quad \xi_s(n) = R_s(n)/R(n), \quad (6)$$

representing the SV power spectra of advection, errors of representativeness, diffusion and subgrid processes, normalized to that of the total SV. We show in Figure 1  $\xi_e$ ,  $\xi_d$ ,  $\xi_s$  and  $\xi_f$  for the Mid-Path simulation, for which  $e_r$  is calculated as the difference between the total SV (estimated from equation (1), where  $B_r$  and  $\mathbf{u}_h$  are considered up to  $n = 133$ ) and the e.m.f. involving large length-scale fields only. We recall that the Mid-Path dynamo (sampled every 0.2 yrs over 20,200 yrs) uses a stress-free condition at the outer boundary, so that all terms can be evaluated at the CMB [unlike with a no-slip condition where the SV contributions should be evaluated below the Ekman layer, e.g. Rau et al., 2000]. Figure 1 indicates that the relative contribution from  $\mathbf{f}$  is almost insensitive to  $n$ , with  $\xi_f \approx 60\%$ .  $e_r$  appears as a major source of SV, with  $\xi_e$  up to  $\approx 40$  to 60% of the total SV energy depending on the degree (a magnitude only slightly weaker than that of the SV generated by large scale fields). This estimate of the SV from  $e_r$  is slightly larger than the previous estimates by e.g. Pais and Jault [2008] or Gillet et al. [2015a], the latter from now on referred to as GJF15. In the construction of  $e_r$ , subgrid processes dominate over diffusion, with for the non-dipole SV  $\xi_d \simeq 8\%$  for all length-scales, and  $\xi_s$  gradually increasing from about 25 to 45% for degrees 2 to 13. The two contributions peak for the dipole, with

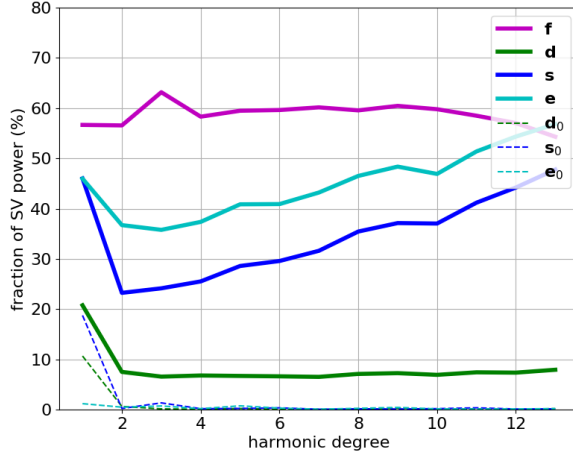


Figure 1: Fractions  $\xi_e$  (cyan),  $\xi_d$  (green),  $\xi_s$  (blue) and  $\xi_f$  (magenta) as defined in equation (6), for the Mid-Path dynamo. Lowes spectra have been averaged over time. Note that  $\xi_e + \xi_f$  can be larger than 1, because these two sources of SV may partly cancel each other. Thin lines (for  $e_0$ ,  $d_0$  and  $s_0$ ) correspond to the same quantities, but for the time-averaged fields.

$\xi_d(1) \simeq 20\%$  and  $\xi_s(1) \simeq 45\%$ , that partly cancel out each other as  $\xi_d(1) + \xi_s(1) > \xi_e(1)$ . It is worth noticing that time average contributions from diffusion, significantly non-zero for the dipole only, is almost entirely cancelled out by subgrid processes (compare the fractions for  $e_0$  and  $d_0$  in Figure 1).

Even small, the contribution from diffusion is much larger than what could be expected from estimates of the magnetic Reynolds number  $R_m = UL/\eta$  (ratio of diffusive to advective time-scales, resp.  $L^2/\eta$  and  $L/U$ , based on the large scale flow magnitude  $U$  and an integral length-scale  $L$ ), of the order of 1000 in both the Earth's core and the Mid-Path dynamo. This is most likely due to the presence of a magnetic boundary layer below the CMB, whose thickness  $L_\eta \ll L$  enhances the role of diffusion, resulting in an effective Reynolds number  $R_m^* = UL_\eta^2/(\eta L) \approx 10$  (from Figure 1), so that  $L_\eta \approx L/10$  [e.g. Amit and Christensen, 2008, Finlay and Amit, 2011]. Such an argument is pushed to an extreme end by Metman et al. [2019], who manage to fit SV changes over the past century with only diffusion modes, advocating against a too crude application of the frozen-flux hypothesis ( $\eta = 0$ ) often considered in the reconstruction of the core surface flow. In contrast, our above analysis of the Mid-Path dynamo gives a relatively minor role to diffusion, in comparison with subgrid processes, first put forward by Eymin and Hulot [2005] and popularized after Pais and Jault [2008].

We show in Figures 2a, b two examples of SV Gauss coefficient series, and the respective contributions from  $\mathbf{f}$  and  $\mathbf{e}$ , for the Mid-Path dynamo. We recover, in series of  $dg_n^m/dt$  sampled at decadal periods, abrupt changes in both sources of SV on the r.h.s. of equation (3). The SV correlation time for  $\mathbf{e}$  (Figure 2e, see legend for its definition) is significantly shorter than that for  $\mathbf{f}$ , as already suggested by GJF15. It decreases from about 40 to 15 yrs from harmonic degree 1 to 13 (larger at low degrees than estimated by GJF15). The SV correlation time for  $\mathbf{f}$  spans centennial to decadal periods from low to high degrees.

As a consequence, a major part of the slow (centennial and longer) SV is carried by the induction from large length-scale fields. For degrees  $2 \leq n \leq 10$ , significantly longer correlation times are found for coefficients of order  $m = n - 1$ , and to a lesser extent also for other equatorially anti-symmetric coefficients of order  $m = n - 3$ , etc. We recover this signature in the correlation time for the total SV, a result reminiscent of the findings by Amit et al. [2018]. They argue that such a feature cannot be explained by the interaction of the long-lived equatorially symmetric westward gyre [Pais and Jault, 2008] with the equatorially anti-symmetric dipole field: in simulations, because of field-flow alignment at mid- to high latitudes (see §3.2), this contribution to the SV is much smaller than what their respective amplitude could suggest. Following Amit et al. [2018], a lower alignment at low latitudes would favor stronger equatorially symmetric SV [Finlay and Amit, 2011], and consequently longer time-scales for the anti-symmetric SV.

Interestingly, the temporal spectrum of SV series for the axial dipole (see Figure 2c) indicates a power at decadal to centennial periods significantly larger for  $\mathbf{e}$  than for  $\mathbf{f}$  (see also Figure 2a). This is not in contradiction with Figure 1 since the SV energy is dominated by long periods (SV spectra are red), where the relative roles of both SV sources are more ballanced. The Gauss coefficients SV power spectra deflect from a spectral index  $p \simeq 2$  (corresponding to  $p \simeq 4$  for the MF) at periods shorter than 20 yrs. Towards high frequencies, the simulation data features a steeper spectral decay than that suggested by ground-based observations. From observatories series indeed,  $p \simeq 2$  is recovered down to about  $\approx 1$  yr periods [Lesur et al., 2018]. We conclude that computations, even as extreme as Mid-Path, do not capture interannual field changes, because rapid and turbulent MHD processes are not fully rendered. The use of hyperdiffusivity in Mid-Path alters nonlinear interactions. However, Aubert [2018] observes that the range of frequencies over which the scaled simulations replicate the observed geophysical spectrum increases both with and without hyperdiffusivity. We presume that the limited forcings accessible numerically (dimensionless numbers are still away from Earth-like values) are the main limitation to the recovery of interannual SV variations. We detail in the next section §2.2 our approach to complement the SV power spectrum towards short periods.

## 2.2 Derivation of the forward stochastic model from geodynamo simulations

We consider a reduced rank basis for the flow vector. We apply a principal component analysis (PCA) on the matrix  $\mathbf{P}_{vv}^*$  obtained from the Mid-path model, which eigen vectors are sorted by decreasing variance. With our definition of  $\mathbf{v}$  (see equation 4), the variance is equivalent to a core surface kinetic energy. The motivation for working in a sub-space for the flow is two-folded:

- (i) the length of the available geodynamo series (about 20,000 yrs) and the value for the overturn time [about 125 yrs in Mid-Path, see Aubert, 2018] result in  $\approx 160$  independent snapshots, while the flow vector  $\mathbf{v}$ , for a truncation at  $n_u = 18$ , is made of  $N_u = 720$  coefficients;
- (ii) the symmetry imposed by the fast rotation reduces by a factor about 4 the number of independent parameters [see Appendix D in Gillet et al., 2011].

We thus consider, conservatively,  $N_{\bar{v}} = 200$  principal components (PC), which account for up to  $\approx 96\%$  of the core surface kinetic energy. As we see no such physical motivation and numerical need for considering a subspace for  $\mathbf{e}$ , we describe its evolution with the original spherical harmonic basis functions.



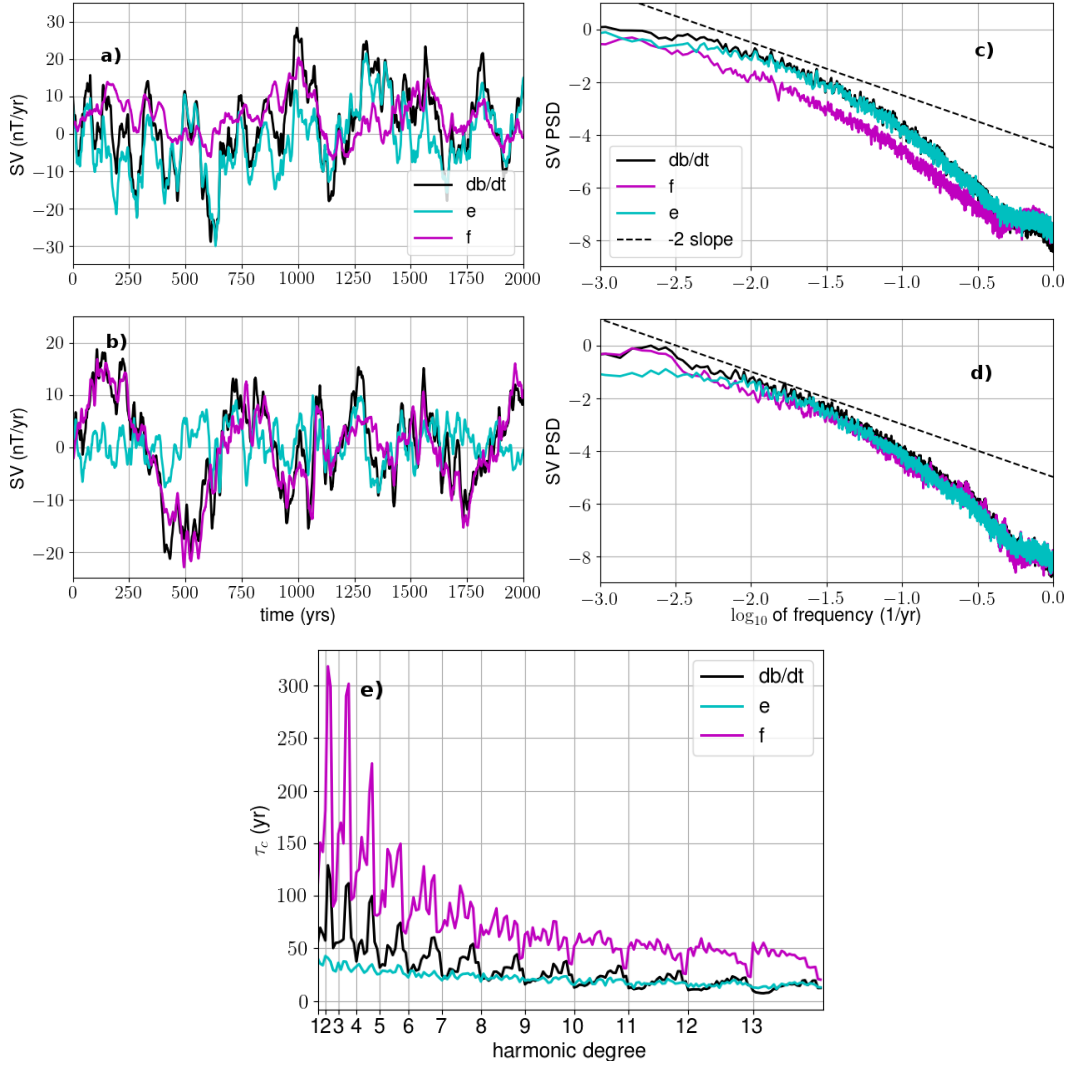


Figure 2: Gauss coefficients SV time series for respectively  $dg_1^0/dt$  (a) and  $dg_3^2/dt$  (b) for the Mid-Path dynamo, with the associated PSD (c and d). In black the total SV, with isolated contributions from **f** (magenta) and **e** (cyan). The dotted line corresponds to a spectral index  $p = 2$ . The PSD is performed with a multi-taper analysis, using 10 tapers over the whole time-span of Mid-Path, imposing a Hanning window on each taper after removing the end-to-end line. e) SV correlation times for series from the Mid-Path dynamo as a function of  $(n, m)$  and estimated (assuming a Laplacian time-correlation) as  $\tau_c(n, m) = -\Delta t / \log(\rho_n^m(\Delta t))$  with  $\rho_n^m(\Delta t)$  the cross-correlation for the series  $dg_n^m/dt(t)$  at lag  $\Delta t = 20$  yr. In black for the total SV, and for isolated contributions **f** (magenta) and **e** (cyan). Spherical harmonic components on the  $x$ -axis are lexicographically ordered as follows:  $(n, m) = (1, 0), (1, -1), (2, 0), (2, 1), (2, -1), (2, 2), (2, -2), (3, 0) \dots$  with positive (negative)  $m$  corresponding to  $g_n^m$  ( $h_n^m$ ).

The vector  $\mathbf{v}$  containing spherical harmonic coefficients is related to the vector  $\tilde{\mathbf{v}}$  (of size  $N_{\tilde{v}} < N_u$ ) that describes the flow in the PC reduced basis (constructed with the first  $N_{\tilde{v}}$  eigen vectors of  $\mathbf{P}_{vv}^*$ ), through

$$\mathbf{v} = \mathbf{v}_0 + \mathbf{S}_v \tilde{\mathbf{v}}, \quad (7)$$

with  $E(\tilde{\mathbf{v}}) = \mathbf{0}$  and  $\mathbf{v}_0$  the geodynamo time average flow vector. The operator  $\mathbf{S}_v$ , of size  $N_u \times N_{\tilde{v}}$ , is built from eigen vectors of  $\mathbf{P}_{vv}^*$  [details about the projection onto the PC can be found in e.g. Pais et al., 2015]. We concatenate  $\tilde{\mathbf{v}}$  with  $\mathbf{e}$  to constitute the augmented state vector  $\mathbf{z}^T = [\tilde{\mathbf{v}}^T \mathbf{e}^T]$ . We note that entries of the cross-covariance matrix  $\mathbf{P}_{\tilde{v}e}^*$ , which fills the two off-diagonal blocks of the matrix  $\mathbf{P}_{zz}^*$ , are very weak. For the sake of simplicity, we thus treat  $\tilde{\mathbf{v}}$  and  $\mathbf{e}$  as independent vectors.

Following Gillet et al. [2015b] and BGA17, we model the behaviour of the state vector  $\mathbf{z}(t)$  with a multi-variate AR-1 process, of the form

$$\begin{cases} d\tilde{\mathbf{v}} + \mathbf{D}_{\tilde{v}} \tilde{\mathbf{v}} dt &= d\zeta_{\tilde{v}} \\ d\mathbf{e}' + \mathbf{D}_e \mathbf{e}' dt &= d\zeta_e \end{cases}, \quad (8)$$

where  $\zeta_{\tilde{v},e}$  are Wiener (or Brownian motion) processes. The choice for this family of models is motivated by the loss of power towards short periods seen in Figure 2 for both  $\mathbf{e}$  and  $\mathbf{f}$  (and consequently for  $\mathbf{u}$  that governs the spectrum of  $\mathbf{f}$ ). By definition AR-1 equations will generate a spectral index  $p = 2$  for the SV PSD at high frequencies, as suggested by observatory time series [Lesur et al., 2018].

Compared with BGA17, we generalize here to the case where the drift matrices  $\mathbf{D}_{v,e}$ , corresponding to a ‘restoring force’ in equation (8), are possibly dense instead of diagonal. Discretizing equations (8) with an Euler-Maruyama scheme, we obtain for finite time increments  $\Delta t$  separating two epochs  $t_i$  and  $t_{i+1}$ ,

$$\begin{cases} \tilde{\mathbf{v}}_{i+1} &= (\mathbf{I}_{\tilde{v}} - \Delta t \mathbf{D}_{\tilde{v}}) \tilde{\mathbf{v}}_i + \sqrt{\Delta t} \mathbf{r}_{\tilde{v}i} \\ \mathbf{e}'_{i+1} &= (\mathbf{I}_b - \Delta t \mathbf{D}_e) \mathbf{e}'_i + \sqrt{\Delta t} \mathbf{r}_{ei} \end{cases}, \quad (9)$$

where  $\mathbf{I}_b$  (resp.  $\mathbf{I}_{\tilde{v}}$ ) is the identity matrix of rank  $N_b$  (resp.  $N_{\tilde{v}}$ ). We compute the zero mean normal random vectors  $\mathbf{r}_{ei}$  and  $\mathbf{r}_{\tilde{v}i}$  (of sizes respectively  $N_b$  and  $N_{\tilde{v}}$ ) as  $\mathbf{r}_{ei} = \mathbf{L}_e \mathbf{w}_{ei}$  and  $\mathbf{r}_{\tilde{v}i} = \mathbf{L}_{\tilde{v}} \mathbf{w}_{\tilde{v}i}$  where  $\mathbf{w}_{ei}$  and  $\mathbf{w}_{\tilde{v}i}$  are centered, unit variance normal random vectors. As such, their cross-covariance matrices are  $\mathbf{P}_{rre} = E(\mathbf{r}_e \mathbf{r}_e^T) = \mathbf{L}_e \mathbf{L}_e^T$  and  $\mathbf{P}_{r\tilde{v}\tilde{v}} = E(\mathbf{r}_{\tilde{v}} \mathbf{r}_{\tilde{v}}^T) = \mathbf{L}_{\tilde{v}} \mathbf{L}_{\tilde{v}}^T$ . In other words,  $\mathbf{L}_{v,e}$  are the Choleski decomposition of  $\mathbf{P}_{rr\tilde{v},e}$ .

Our goal is to generate discrete stochastic processes  $\mathbf{u}(t)$  and  $\mathbf{e}(t)$ , with spatial and temporal covariance matrices matching those of the geodynamo time series  $\mathbf{u}^*(t)$  and  $\mathbf{e}^*(t)$ . To achieve this, we need to determine  $\mathbf{D}_{\tilde{v}}$ ,  $\mathbf{P}_{r\tilde{v}\tilde{v}}$ ,  $\mathbf{D}_e$  and  $\mathbf{P}_{rre}$  from covariance matrices established upon the geodynamo series. The general context of the derivation may be found in Neumaier and Schneider [2001]. For each AR-1 process of equation (9), we need two linear constraints to estimate two matrices. These constraints are obtained from (i) the auto-covariance matrix of the dynamo processes, and (ii) the time cross-covariances of the same process estimated for a given lag (noted  $\Delta t^*$  below).

On top of  $\mathbf{P}_{\tilde{v}\tilde{v}}^*$ , we define several empirical cross-covariance matrices (in the PC basis) characteristic of the geodynamo free run sampled every  $\Delta t^*$ ,

$$\mathbf{P}_{\tilde{v}\tilde{v}+}^* = E(\tilde{\mathbf{v}}^*(t) \tilde{\mathbf{v}}^*(t + \Delta t^*)^T) \text{ and } \mathbf{P}_{\Delta\tilde{v}\Delta\tilde{v}}^* = E(\Delta\tilde{\mathbf{v}}^* \Delta\tilde{\mathbf{v}}^{*T}), \quad (10)$$

with the increment between two successive time-steps  $\Delta\tilde{\mathbf{v}}^*(t) = \tilde{\mathbf{v}}^*(t + \Delta t^*) - \tilde{\mathbf{v}}^*(t)$  – of zero average since the process  $\tilde{\mathbf{v}}^*$  is considered as stationary. We similarly define, for the process  $\tilde{\mathbf{v}}$  sampled by equation (9) every  $\Delta t \neq \Delta t^*$ , the matrices

$$\mathbf{P}_{\tilde{v}\tilde{v}+} = E(\tilde{\mathbf{v}}_i \tilde{\mathbf{v}}_{i+1}^T), \mathbf{P}_{\tilde{v}\Delta\tilde{v}} = E(\tilde{\mathbf{v}}_i \Delta\tilde{\mathbf{v}}_i^T) \text{ and } \mathbf{P}_{\Delta\tilde{v}\Delta\tilde{v}} = E(\Delta\tilde{\mathbf{v}}_i \Delta\tilde{\mathbf{v}}_i^T), \quad (11)$$

with  $\Delta\tilde{\mathbf{v}}_i = \tilde{\mathbf{v}}_{i+1} - \tilde{\mathbf{v}}_i$ . Estimating  $E(\tilde{\mathbf{v}}_i\tilde{\mathbf{v}}_{i+1}^T)$  from equation (9) we first get, since  $E(\tilde{\mathbf{v}}_i\mathbf{r}_i^T) = 0$ ,

$$\mathbf{P}_{\tilde{v}\tilde{v}+} = \mathbf{P}_{\tilde{v}\tilde{v}} (\mathbf{I}_{\tilde{v}} - \Delta t \mathbf{D}_{\tilde{v}}^T) \Rightarrow \mathbf{D}_{\tilde{v}} = \frac{\mathbf{I}_{\tilde{v}} - (\mathbf{P}_{\tilde{v}\tilde{v}}^{-1} \mathbf{P}_{\tilde{v}\tilde{v}+})^T}{\Delta t}, \quad (12)$$

while from  $E(\Delta\tilde{\mathbf{v}}_i\Delta\tilde{\mathbf{v}}_i^T)$  we derive

$$\mathbf{P}_{rr\tilde{v}} = \frac{\mathbf{P}_{\Delta\tilde{v}\Delta\tilde{v}}}{\Delta t} - \Delta t \mathbf{D}_{\tilde{v}} \mathbf{P}_{\tilde{v}\tilde{v}} \mathbf{D}_{\tilde{v}}^T. \quad (13)$$

Then by choosing for equations in (9)

$$\begin{cases} \mathbf{D}_{\tilde{v}}^* = \frac{\mathbf{I}_{\tilde{v}} - (\mathbf{P}_{\tilde{v}\tilde{v}}^{*-1} \mathbf{P}_{\tilde{v}\tilde{v}+}^*)^T}{\Delta t^*} \\ \mathbf{P}_{rr\tilde{v}}^* = \frac{\mathbf{P}_{\Delta\tilde{v}\Delta\tilde{v}}}{\Delta t^*} - \Delta t^* \mathbf{D}_{\tilde{v}}^* \mathbf{P}_{\tilde{v}\tilde{v}}^* \mathbf{D}_{\tilde{v}}^{*T} \end{cases}, \quad (14)$$

and with equivalent expressions for  $\mathbf{D}_e^*$  and  $\mathbf{P}_{rre}^*$ , we design with the system (9) a process that in principle resembles the considered geodynamo series – provided  $\Delta t \ll \Delta t^*$ . Equations of the system (14) may be recovered in section §3.1 of Neumaier and Schneider [2001]. The projection back to the spherical harmonic coefficients basis is performed using equation (7). It is worth pointing out that the statistics obtained from the geodynamo depend on the considered sampling period  $\Delta t^*$ . This period should be neither too small (to avoid the integrated stochastic system being imprinted by possibly non-geophysical dynamics in the geodynamo at short periods, see Figures 2c, d), nor too long to avoid losing the decadal dynamics of interest in the simulation.

We show in Figure 3 the eigenvalues of matrices  $\mathbf{D}_{\tilde{v},e}^*$ , decomposed as  $\lambda_{\tilde{v},e} = \lambda_{\tilde{v},e}^r + i\lambda_{\tilde{v},e}^i$ . These are composed of pairs of complex conjugates. Their associated decay times  $1/\lambda_{\tilde{v},e}^r$  indicate that unmodelled sources act on a much shorter time-scale than some of the flow constituents: while  $e_r$  is mainly responsible for decadal SV changes, part of the memory on the core surface flow indeed holds over millennial time-scale. This could already be suspected from Figure 2e, where the correlation time is shorter for  $\mathbf{e}$  than for  $\mathbf{f}$ , consistent with the westward gyre persistence over more than a century [Aubert, 2014, Pais et al., 2015, Baerenzung et al., 2018]. We also notice long pulsations associated with these modes, with a (general but non systematic) trend for the longest periods to be associated with long decay times. Slightly shorter decay times are found while increasing the sampling period  $\Delta t^*$  of geodynamo series.

To assess how much our stochastic model fulfills our goal stated above, we compare in Figure 4 the temporal spectra of individual core flow Gauss coefficients for the stochastic model series and the geodynamo series (see Appendix A for a comparison of spatial cross-correlations and spatial spectra). Our stochastic model replicates the geodynamo spectrum over centennial and longer periods, even showing comparable oscillations on millennial time-scales (see for instance the PSD for  $s_2^{1,s}$  and  $t_2^{1,c}$ ). The crucial point is that at short periods the model extends the frequency range where  $p = 2$  in the core flow spectrum down to the Nyquist period  $2\Delta t$ . Despite a small loss of core flow magnitude (see Appendix A) and slightly attenuated decadal variations (see Figure 4), we therefore conclude that our model fulfills our initial requirements, i.e. reproducing the geodynamo spatial and temporal changes on decadal and longer periods, and replicating down to short periods the spectral index indicated by geophysical records.

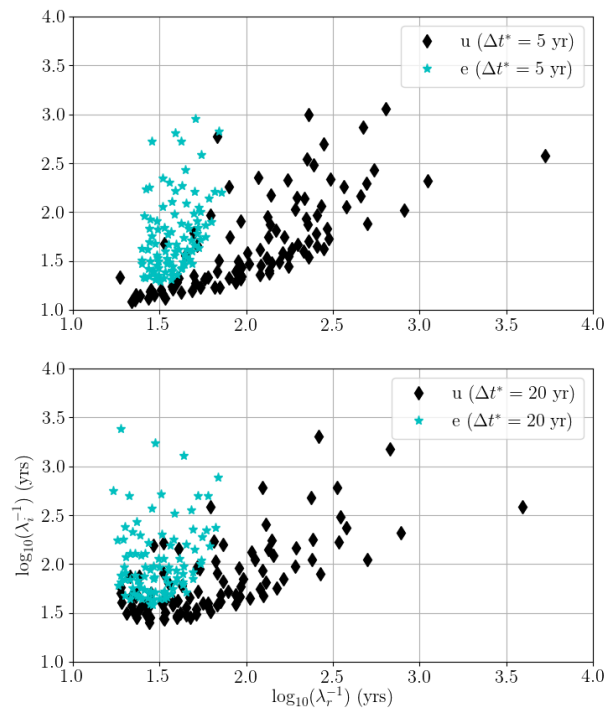


Figure 3:  $1/\lambda_{\tilde{v},e}^i$  (inverse of the angular frequency) as a function of  $1/\lambda_{\tilde{v},e}^r$  (decay time) for eigen-vectors of matrices  $D_{\tilde{v}}$  (black diamonds) and  $D_e$  (cyan stars) in the cases  $\Delta t^* = 5$  yrs (top) and  $\Delta t^* = 20$  yrs (bottom). Only eigen-values with positive imaginary parts are shown.

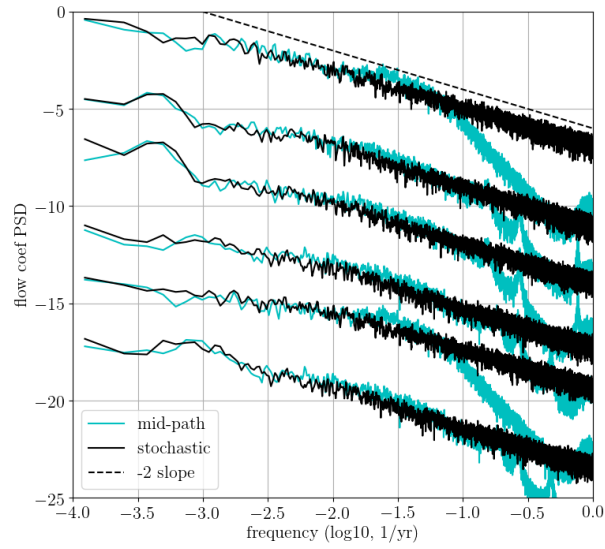


Figure 4: PSD of spherical harmonic flow coefficients for the Mid-Path dynamo (cyan) and an integration of the stochastic model over 20,000 yrs (black), with  $\Delta t = 0.2$  yr and  $\Delta t^* = 5$  yr. The dotted line corresponds to a spectral index  $p = 2$ . From top to bottom:  $t_1^0, s_2^{1,s}, t_2^{1,c}, s_2^{2,c}, t_3^0, t_3^{2,s}$ , vertically shifted by multiples of 3 decades for clarity.

### 2.3 A stochastic geomagnetic assimilation algorithm

As in BGA17, we consider an augmented state Kalman filter tool to re-analyse geomagnetic MF and SV observations. It is composed of a succession of forecast and analysis steps. Between two analysis epochs, the trajectory of an ensemble of  $N_e = 50$  realizations of the forecast state  $\{\mathbf{z}_k^f, \mathbf{b}_k^f\}_{k=1\dots N_e}$  is integrated using the discretised version of equation (3) together with the system (9), where matrices  $D_{\tilde{v},e}$  and  $L_{\tilde{v},e}$  are defined from (14).

Analyses are performed every  $\Delta t_a$ , using noisy observations  $\{\mathbf{b}_k^o(t), \dot{\mathbf{b}}_k^o(t)\}_{k=1\dots N_e}$  of the MF and its SV. MF (resp. SV) data are noised by adding a Gaussian perturbation whose statistics are described by the covariance matrix  $R_{bb}$  (resp.  $R_{\dot{b}\dot{b}}$ ) for observation errors on MF (resp. SV) Gauss coefficients. Following Aubert [2014], each analysis is performed in two steps. First we deduce the analysed MF from MF data:

$$\forall k \in [1, N_e], \mathbf{b}_k^a(t_a) = \mathbf{b}_k^f(t_a) + \mathbf{P}_{bb}^f \left[ \mathbf{P}_{bb}^f + R_{bb} \right]^{-1} \left( \mathbf{b}_k^o(t_a) - \mathbf{b}_k^f(t_a) \right). \quad (15)$$

The MF forecast covariance matrix  $\mathbf{P}_{bb}^f$  is chosen conservatively equal to  $\mathbf{P}_{bb}^*$  and kept constant over all analysis steps. As the MF Gauss coefficients are tightly constrained by observations, the result of the inversion barely changes whether  $\mathbf{P}_{bb}^f$  is updated at each analysis step or kept constant (see BGA17). At each analysis epoch  $t_a$ , the analysed magnetic field  $\mathbf{b}_k^a(t_a)$  obtained from (15) is therefore an optimal interpolation of the forecast magnetic field  $\mathbf{b}_k^f(t_a)$  and the observation  $\mathbf{b}_k^o(t_a)$ , taking into account the covariance properties of these two quantities. Next the analysis of SV data gives access to the augmented state  $\mathbf{z} = [\tilde{\mathbf{v}}^T \mathbf{e}^T]^T$ :

$$\forall k \in [1, N_e], \mathbf{z}_k^a(t_a) = \mathbf{z}_k^f(t_a) + \mathbf{P}_{zz}^f \mathbf{H}_k^T \left[ \mathbf{H}_k \mathbf{P}_{zz}^f \mathbf{H}_k^T + R_{\dot{b}\dot{b}} \right]^{-1} \left( \delta \dot{\mathbf{b}}_k^o(t_a) - \mathbf{H}_k \mathbf{z}_k^f(t_a) \right) \quad (16)$$

with, according to equations (3), (4) and (7),

$$\mathbf{H}_k = \left[ \mathbf{A}(\mathbf{b}_k^a) \mathbf{N}^{-1} \mathbf{S}_v \quad \mathbf{I}_b \right], \text{ and } \delta \dot{\mathbf{b}}_k^o(t_a) = \dot{\mathbf{b}}_k^o(t_a) - \mathbf{A}(\mathbf{b}_k^a) \mathbf{N}^{-1} \mathbf{v}_0. \quad (17)$$

At each analysis epoch  $t_a$ , the analysed augmented state  $\mathbf{z}_k^a(t_a)$ , comprising the reduced flow and error vectors, results from an update of the forecast  $\mathbf{z}_k^f(t_a)$  through a classical Bayesian linear estimation that encapsulates the resolution of the core flow inversion problem encoded in  $\mathbf{H}_k$ . Only the two diagonal blocks ( $\mathbf{P}_{\tilde{v}\tilde{v}}^f$  and  $\mathbf{P}_{ee}^f$ ) of the forecast covariance matrix  $\mathbf{P}_{zz}^f$  are non-zero – uncertainties on analysed vectors  $\mathbf{e}$  and  $\tilde{\mathbf{v}}$  are supposed independent, an assumption empirically checked.

Provided  $\Delta t_a$  is much smaller than the typical decay times  $1/\lambda_{\tilde{v}}^r$  of matrix  $D_{\tilde{v}}$  (which is the case in practice, see figure 3), we may approximate from equation (9), for any analysis epoch  $t_a = t_i$  (corresponding to the  $i^{th}$  forecast step),

$$\tilde{\mathbf{v}}^f(t_a + \Delta t_a) \simeq \tilde{\mathbf{v}}^a(t_a) + \sqrt{\Delta t} \sum_{k=1}^K \mathbf{r}_{\tilde{v}}(t_{i+k}), \quad (18)$$

with  $K = \Delta t_a / \Delta t$  the number of forecast steps between two analyses. It derives from (18) that  $\tilde{\mathbf{v}}_0^f(t_a + \Delta t_a) \simeq \tilde{\mathbf{v}}_0^a(t_a)$ . Furthermore, since  $E(\tilde{\mathbf{v}}^a(t_i) \mathbf{r}_{\tilde{v}}(t_{i+k})^T) = 0$ , we obtain for the forecast covariance matrix

$$\mathbf{P}_{\tilde{v}\tilde{v}}^f \simeq \mathbf{P}_{\tilde{v}\tilde{v}}^a + K \Delta t \mathbf{P}_{rr\tilde{v}}^* = \mathbf{P}_{\tilde{v}\tilde{v}}^a + \Delta t_a \mathbf{P}_{rr\tilde{v}}^*, \quad (19)$$

with  $P_{rr\bar{v}}^*$  from equation (14). A similar expression holds for  $P_{ee}^f$  from  $P_{rre}^*$  and  $P_{ee}^a$ .  $P_{\bar{v}\bar{v}}^a$  and  $P_{ee}^a$  are the covariance matrices of the analysed states. Off-diagonal elements of their respective cross-correlation matrices  $C_{\bar{v}\bar{v}}^a$  and  $C_{ee}^a$  are found to be small compared to unity. We thus disregard them, rendering  $P_{\bar{v}\bar{v}}^a$  and  $P_{ee}^a$  well conditioned regardless of the ensemble size. The above relation (19) generalizes equation (21) of BGA17 to the case of dense drift matrices. BGA17 either considered the second term only in (19), thus neglecting  $P_{\bar{v}\bar{v}}^a$ , or replaced  $P_{\bar{v}\bar{v}}^f$  by the whole spread of dynamo states represented by  $P_{\bar{v}\bar{v}}$ , with both approaches leading to similar results. This weak sensitivity to the definition for  $P_{\bar{v}\bar{v}}^f$  does not hold in our study because contrary to BGA17, we have here  $\sqrt{\lambda_u^r \Delta t^a} \ll 1$ . Properly accounting for both terms on the r.h.s. of (19) is thus mandatory in order to allow enough freedom for the flow to significantly change between two analyses. As such, our algorithm comes down to a proper ensemble Kalman filter [EnKF, see Evensen, 2003], as the forecast covariance matrix is updated at each analysis step based on cross-covariances within the ensemble of forecasts.

### 3 Results: core flow re-analysis over 1880–2015

We apply the algorithm defined in sections §2.2–2.3 to the COV-OBS.x1 field model [Gillet et al., 2015b] over  $[t_s, t_e] = [1880, 2015]$ , which provides MF and SV Gauss coefficient observations. Also provided with COV-OBS.x1 are time-dependent posterior error covariances on the geomagnetic field model, which are used to define matrices  $R_{bb}$  and  $R_{\bar{b}\bar{b}}$  (only diagonal elements are considered here). We consider  $\Delta t_a = 1$  yr,  $\Delta t = 4$  months, and compare two solutions, obtained with  $\Delta t^* = 5$  yrs and  $\Delta t^* = 20$  yrs. In order to attest to the coherence of our solutions with respect to the geodynamo model covariances on the one hand, and to the data error covariances on the other hand, we give in Appendix B the time evolution of several normalised misfits.

#### 3.1 SV signature of errors of representativeness

Gauss coefficient SV predictions from the analysed augmented state confirm the good fit to SV data (see the two examples shown in Figure 5). The separation between  $\mathbf{e}$  and the SV  $\mathbf{f}$  generated by large length-scale fields appears rather insensitive to the choice of the free parameter  $\Delta t^*$ . A significant fraction of interannual to decadal changes is absorbed by the contribution from the errors of representativeness, in agreement with Figure 2 and the dynamo series used to construct the stochastic model. This is in contrast with the results by GJF15, where most of the rapid SV changes were associated with  $\mathbf{f}$ . It is possibly due to the weak formalism [Sasaki, 1970] they used to jointly invert all epochs, or to the fact they integrate the impact of  $e_r$  into  $R_{\bar{b}\bar{b}}$ . In the present study, the contributions to such rapid changes are more balanced between  $\mathbf{e}$  and  $\mathbf{f}$ . This behaviour is found for all Gauss coefficients.

Interestingly,  $\mathbf{f}$  contributes positively, over the whole studied era, to the dipole decay, with a value  $\approx 6 \pm 2$  nT/yr. This observation confirms results already obtained by Barrois et al. [2018a] over the period 1955–2015. It is to be related to a long-lived eccentric westward gyre (see §3.2), validating back to 1880 the scenario put forward by Finlay et al. [2016a], where a positive dipole decay results on average from the interaction of the eccentric gyre with asymmetric main field patches distribution. This view must nevertheless be nuanced as we find a larger contribution from  $e_r$  to  $\dot{j}_1^0$  ( $\approx 10 \pm 5$  nT/yr over 1880–2015). This is coherent with the dynamo series shown in Figure 2, where  $e_r$  can be a significant source of SV on decadal changes in particular for the dipole (see Figure 2a). The analysed

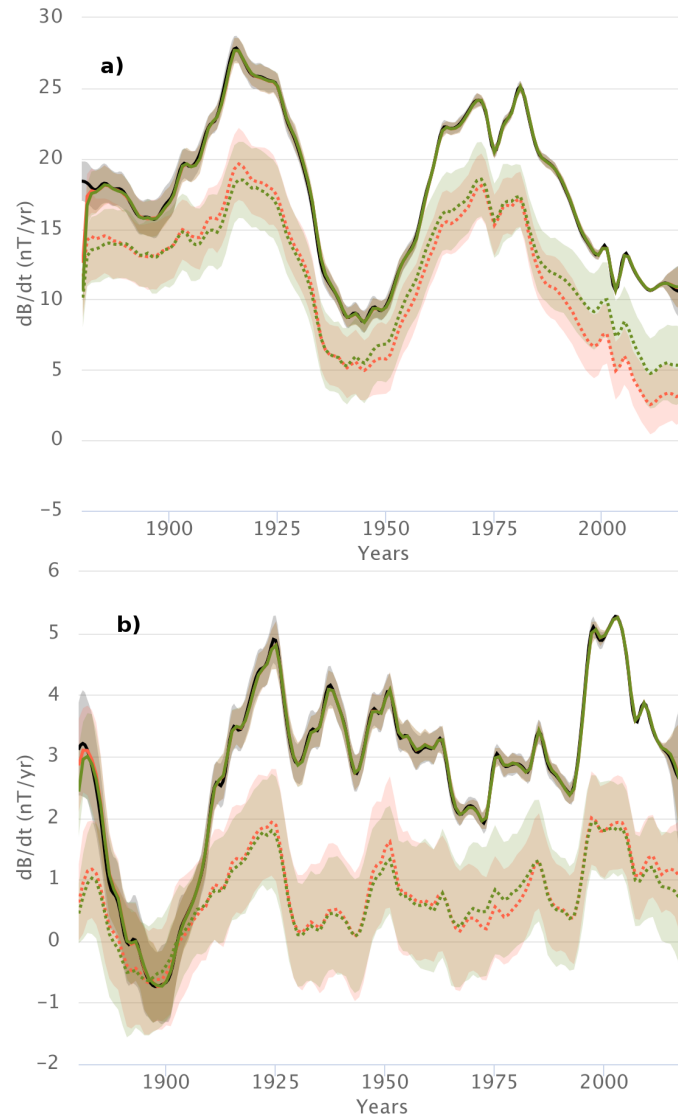


Figure 5: SV predictions (in nT/yr) for Gauss coefficients  $dg_1^0/dt$  (a) and  $dh_4^3/dt$  (b), obtained for the analysed states using  $\Delta t^* = 5$  yrs (green) and 20 yrs (red), compared to the COV-OBS.x1 SV data (the grey shaded area represents  $\pm 1\sigma$  data uncertainties). Dotted lines correspond to the isolated contribution from  $e_r$ . Shaded areas represent the  $\pm 1\sigma$  dispersion within the ensemble of solutions. Total SV model predictions almost superimpose for both values of  $\Delta t^*$ .



series of  $e_r$  are almost always positively correlated to the observed SV. A similar trend is found for the geodynamo series, although for the result of our reanalysis it may also partly result from the least-squares formalism used here (see GJF15).

### 3.2 Motions around the gyre

A planetary scale, axially columnar, eccentric westward gyre was first put forward by Pais and Jault [2008] with kinematic QG flow inversions from satellite SV models, and later confirmed by several authors using various methods to account for errors of representativeness [e.g. Gillet et al., 2009, Aubert, 2013, Baerenzung et al., 2016, Kloss and Finlay, 2019]. As Aubert [2014] or Pais et al. [2015], we recover its existence back to the 19th century (see Figure 6). The obtained flow structure is largely dominated by equatorially symmetric features ( $> 90\%$  in kinetic energy over the whole studied era, for the ensemble average solution). This estimate is coherent with earlier results obtained using the ‘Coupled-Earth’ dynamo as a prior, first by Aubert [2014] inverting the Gauss coefficient data over the observatory era (see their figure 6), and later by Barrois et al. [2018b] considering virtual observatory data over the satellite era.

We observe a (non systematic) tendency of the flow trajectories at the CMB to follow isocontours of  $B_r$  (see the polar and Aitoff projections in Figure 6). It is particularly true for instance around the patch of intense  $B_r$  below Siberia from 1920 to 1980 [but no more in the most recent epochs, in link with the intense SV patches observed in this region from satellites, see Finlay et al., 2016b], or around the reversed flux patch below Patagonia from 1950 onward (but not for the one below Africa). This was already observed with numerical simulations [e.g. Aubert, 2005, Finlay and Amit, 2011, Schaeffer et al., 2017, Aubert et al., 2017], as well as in core flow inversions from magnetic data by Amit and Olson [2004]. This dynamic alignment of the magnetic and velocity fields is commonly found in MHD turbulence in the presence of a strong magnetic field [Tobias et al., 2011], where both fields organize so as to minimize their interaction. It suggests maps of the radial field at the core surface alone, even without SV, carry information on the flow on the CMB as well as deep in the core, because of the axial invariance in this rapidly rotating system. The information from the dynamics (z-invariance and alignment) must be partly embedded in cross-correlations obtained from geodynamo simulations, as these latter together with knowledge of  $B_r$  at the CMB have lead to reliable inferences of the magnetic field deep in the core in twin experiments from numerical dynamos [e.g. Aubert, 2014].

### 3.3 Areas of large flow acceleration

We define the flow acceleration at the CMB for the analysed flow, with finite differences, as

$$a_{\theta,\phi}(t_a + \Delta t_a/2) = (u_{\theta,\phi}(t_a + \Delta t_a)) - u_{\theta,\phi}(t_a)) / \Delta t_a, \quad (20)$$

and the corresponding r.m.s. acceleration profiles as a function of latitude,

$$\hat{a}_{\theta,\phi}(\theta) = \left[ \frac{1}{t_e - t_s} \int_{t_s}^{t_e} \frac{1}{2\pi} \int_0^{2\pi} a_{\theta,\phi}^2(\theta, \phi, t) d\phi dt \right]^{1/2}. \quad (21)$$

These are shown in Figure 7a: the largest accelerations are by far associated with azimuthal motions in an equatorial belt below  $10^\circ$  latitude, as already put forward by GJF15 under the QG assumption, and by Kloss and Finlay [2019] who use both equatorial symmetries of an

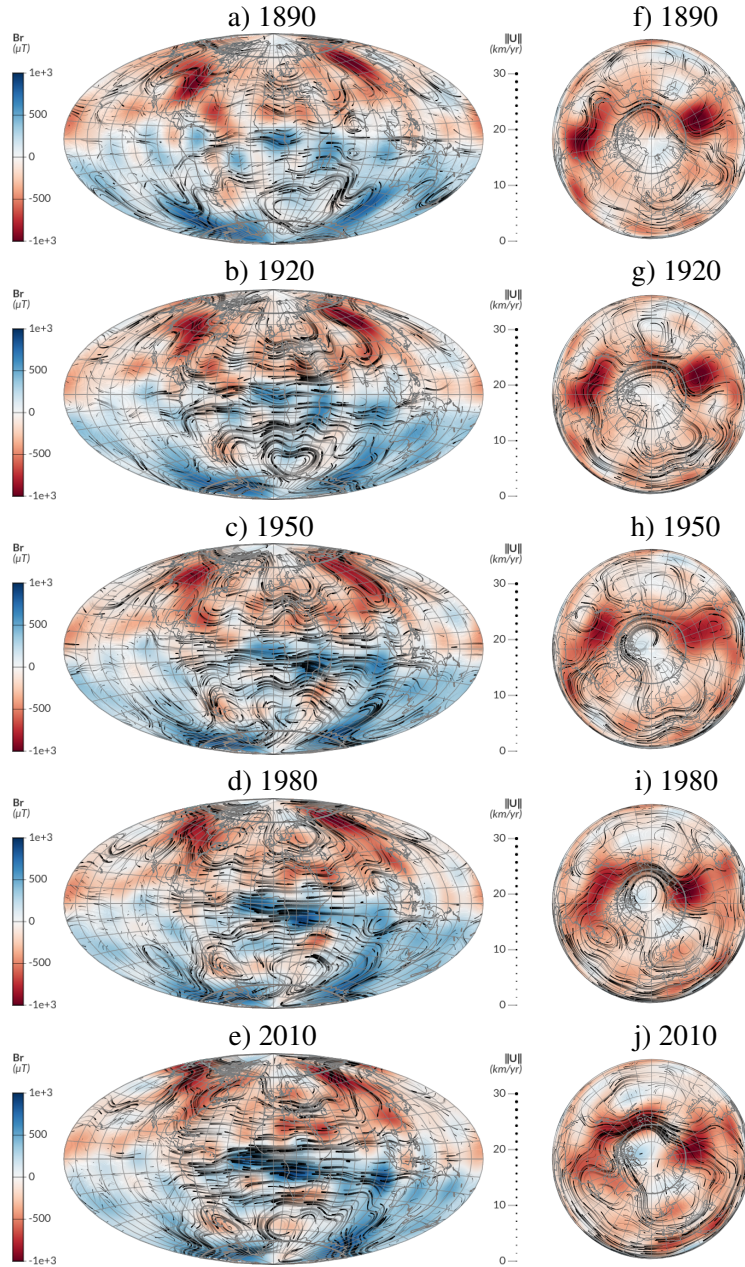


Figure 6: Maps of the radial magnetic field at the CMB (colormap, in  $\mu T$ ) superimposed with core surface flow streamlines (black lines, which thickness give the velocity norm in  $km/yr$ ), every 30 yrs from 1890 (top) to 2010 (bottom), using  $\Delta t^* = 20$  yrs, for the ensemble average flow solution. a–e) Aitoff projection centered on Greenwich. f–j) view from North pole.

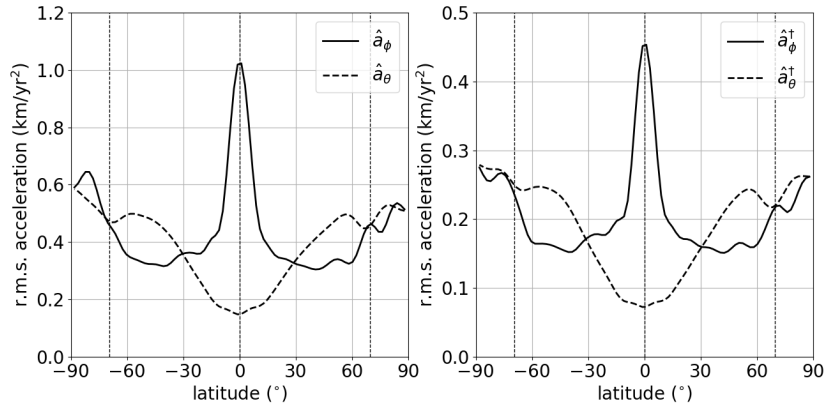


Figure 7: a) r.m.s. acceleration latitudinal profiles  $\hat{a}_{\theta,\phi}$ , for the model obtained using  $\Delta t^* = 20$  yrs. b)  $\hat{a}_{\theta,\phi}^\dagger$  for the band-passed filtered flow (Note different scales). Thin dotted lines mark the position of the tangent cylinder and of the equator.

inertial wave basis to decompose transient motions. To a lesser extent, we also find large accelerations in the azimuthal direction around and within the tangent cylinder (cylinder tangent to the inner core whose axis is aligned with the Earth’s rotation vector), and in the meridional direction at mid to high latitudes.

At Earth-like parameters, transient motions tend to be axially invariant [Jault, 2008, Gillet et al., 2011]. As a consequence, the rapid Earth’s rotation should hinder time-dependent meridional flow velocities (or accelerations) at the positions of the tangent cylinder and of the equator at the CMB. This is the rationale for the QG approximation used in several core flow reconstructions [e.g. Pais and Jault, 2008, Gillet et al., 2009, Livermore et al., 2017]. If we recover, at the precise location of the tangent cylinder and at the equator, a local minimum of  $\hat{a}_\theta$  (see Figure 7a), our results however suggest that these two singular locations cannot be considered as strictly forbidden for meridional flows. This is in agreement with the results from numerical dynamos [Aubert et al., 2017, Schaeffer et al., 2017]. The correlation between local maxima in  $\hat{a}_\phi$  (found at the equator and on the tangent cylinder in the Southern hemisphere) and minima of  $\hat{a}_\theta$  may result partly from the mass conservation weakly imposed with the geodynamo spatial norm – although SV data also likely participate to such features, as GJF15 make similar observation while they impose no specific constraint on the tangent cylinder.

We present in Figure 8 time-longitude diagrams for the acceleration. Large acceleration patterns arise at preferential longitudes, and are primarily non-axisymmetric, confirming similar previous estimates by GJF15 [and more recently by Kloss and Finlay, 2019]. The largest recurrent patterns for  $a_\phi$  are found below the Western equatorial Pacific (see Figure 8a), East of the area of large magnetic secular acceleration found in advanced dynamo simulations with heterogeneous forcing at both the inner and outer boundaries [see Figure 3 of Aubert, 2018]. As Barrois et al. [2018b] we find there, over the satellite era, an Eastward acceleration. This feature does not appear outstanding with regard to previous epochs, which display alternating East-West decadal fluctuations of the flow acceleration as large as  $5 \text{ km/yr}^2$ . The next larger acceleration patterns are seen in the meridional direc-

tion with mid-latitudes eddies. Accelerations are generally the largest between 50°E and 150°E, where the gyre detaches from the tangent cylinder (Figure 8*b*). Such features are primarily symmetrical with respect to the equator.

In order to focalize on interannual flow changes, we also consider the flow  $u_{\theta,\phi}^\dagger$  band-pass filtered in the period range [4, 9.5] yrs where GJF15 found enhanced power in zonal compared to non-zonal motions (see their Figure 9). The corresponding r.m.s. acceleration profiles  $\hat{a}_{\theta,\phi}^\dagger(\theta)$ , shown in 7*b*, are very similar to those obtained for the non-filtered flow, but with a much weaker amplitude. Continuous records from the satellite era give access to sharper acceleration features, with enhanced interannual changes (see Figure 8*c*). While subdecadal interannual acceleration are more obvious since 2000, they are ubiquitous at earlier epochs. As for non-filtered flow solutions, these are dominated by non-zonal patterns. As far as the azimuthal symmetry or the latitudinal distribution of kinetic energy are concerned, we thus see no evidence for significantly different behaviours over the various period ranges. Interestingly, recent satellite data suggest intense interannual azimuthal changes over all longitudes, whereas they were weaker between the Greenwich meridian and 90°W prior to 2000. Similarly to the results of Kloss and Finlay [2019], our time-longitude maps are not supportive of a systematic longitudinal wave-like propagation of acceleration patterns. The propagation of non-axisymmetric patterns at interannual periods would constitute an argument in favor of a stratified layer at the top of the core, because there is no large length-scale magnetostrophic waves at such periods [see Labbé et al., 2015]. Our results therefore indicate that geomagnetic records do not require a stratified layer at the top of the core, as advocated for by Chulliat et al. [e.g. 2015], Buffett et al. [e.g. 2016].

### 3.4 Geostrophic motions and length-of-day predictions

Though axisymmetric (zonal) core surface flows are usually subdominant in the kinetic energy budget, they are of importance for several reasons. First, geostrophic (i.e. zonal and invariant along the rotation axis) motions are singular as the projection of the Coriolis force vanishes on geostrophic cylinders. As a consequence, large length-scales Alfvén waves in the core only project onto the zonal component (the so-called torsional waves). Second, geostrophic motions carry angular momentum, and thus generate a signature in length-of-day (LOD) observations, an independent constraint on the core dynamics [for a thorough discussion of these issues, we refer for instance to Jault and Finlay, 2015].

In Figure 9 we compare LOD predictions from our flow models with geodetic observations (and with predictions from QG models by GJF15). We recover the amplitude and phase for decadal LOD changes from 1930 onwards, while towards earlier epochs we witness a lag between geophysical records and model predictions, as usually found since Jackson et al. [1993]. To our knowledge, the only exception to such reports is the stochastic EnKF re-analysis by Baerenzung et al. [2018], who impose the predicted and geodetic LOD linear trends to match. The good fit in amplitude suggests the magnitude of decadal fluctuations in the Mid-Path dynamo is compatible with the geophysical configuration. From 1940 onwards, it is comparable to that found by GJF15. However, in this latter study the fit to LOD changes was subject to the somewhat ad hoc adjustment of the AR-1 decay time (see their Table 1), while in our case it entirely derives from the choice of dynamo model. In comparison to the single epoch prediction with the Coupled-Earth dynamo norm by Aubert [2014], we improve in particular the fit around 1950, with an overall amplitude of LOD variations slightly closer to geodetic data. In terms of uncertainties, the dispersion within our ensemble of LOD predictions is similar to that by Baerenzung et al. [2018],

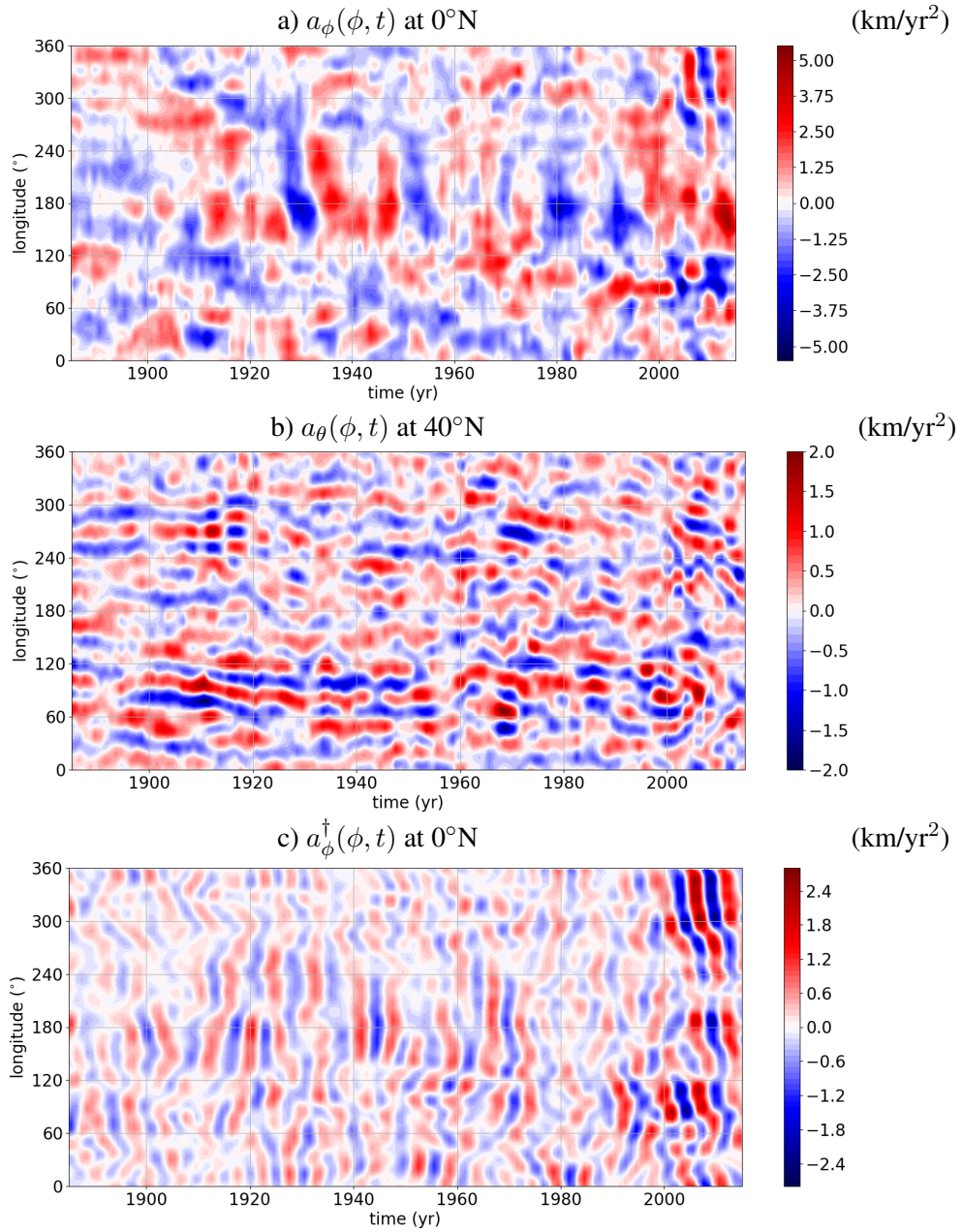


Figure 8: Time-longitude maps of the flow acceleration (in km/yr<sup>2</sup>) for the ensemble average model obtained with  $\Delta t^* = 20$  yrs. a)  $a_\phi$  at the equator. b)  $a_\theta$  at  $40^\circ\text{N}$ . c)  $a_\phi^\dagger$  at the equator, band-pass filtered between 4 and 9.5 yrs.

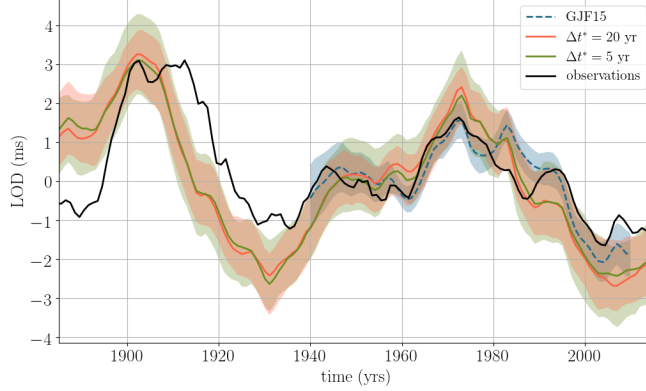


Figure 9: Comparison between geodetic LOD observations (black) and its predictions obtained using  $\Delta t^* = 5$  yrs (green) and 20 yrs (red). In dotted blue the inversion from GJF15. Shaded areas represent the  $\pm 1\sigma$  dispersion within the ensemble of solutions. LOD observations concatenate the Lunar97 series [Gross, 2001] with the C04 series [Bizouard and Gambis, 2009] provided by the International Earth Rotation and Reference System Service (IERS). These latter have been corrected for contributions from atmospheric angular momentum and solid tides, and for the observed 1.4 ms/cy centennial trend [Stephenson et al., 1984] – see BGA17 for details.

about twice larger than the one proposed by GJF15.

We now consider LOD changes filtered for periods within  $[4, 9.5]$  yrs, in order to focus on the 6 yr signal interpreted as the signature of torsional waves by Gillet et al. [2010, 2017]. If zonal motions inside the core are invariant along the Earth’s rotation axis, LOD changes essentially depend on coefficients  $t_1^0$  and  $t_3^0$  [see Jault and Finlay, 2015]. We show in Figure 10 the filtered series of these coefficients,  $(t_1^{0\dagger}, t_3^{0\dagger})$ , which are very similar to those by GJF15 from 1940 onwards. The phase between their ensemble of models and ours is almost identical, while the amplitude in  $t_1^{0\dagger}$  and  $t_3^{0\dagger}$  is slightly smaller in the present study, in particular during the recent satellite era. The similarity between the two studies is striking, given the major differences between their algorithm and ours:

- (i) we follow here a sequential approach with an EnKF, while GJF15 employ a weak formalism [see Sasaki, 1970];
- (ii) we account for errors of representativeness via an augmented state approach, while the impact of  $e_r$  is accounted for via the data error covariance matrix in GJF15.

This resemblance gives some confidence in the obtained interannual zonal core flow changes from 1940 onwards. Furthermore, our ensemble of models adequately fits sub-decadal LOD changes over 1940–2005. However, the lower amplitude of  $t_{1,3}^{0\dagger}$  over the satellite era leads to a much less convincing fit to geodetic records, compared to the model by GJF15. Tests performed with larger SV data uncertainties over the satellite era show that this is not related to the COV-OBS.x1 SV data error model (the matrix  $R_{bb}$ , characterised by much weaker variances since 2000). A possible origin to this misfit is the partition of interan-

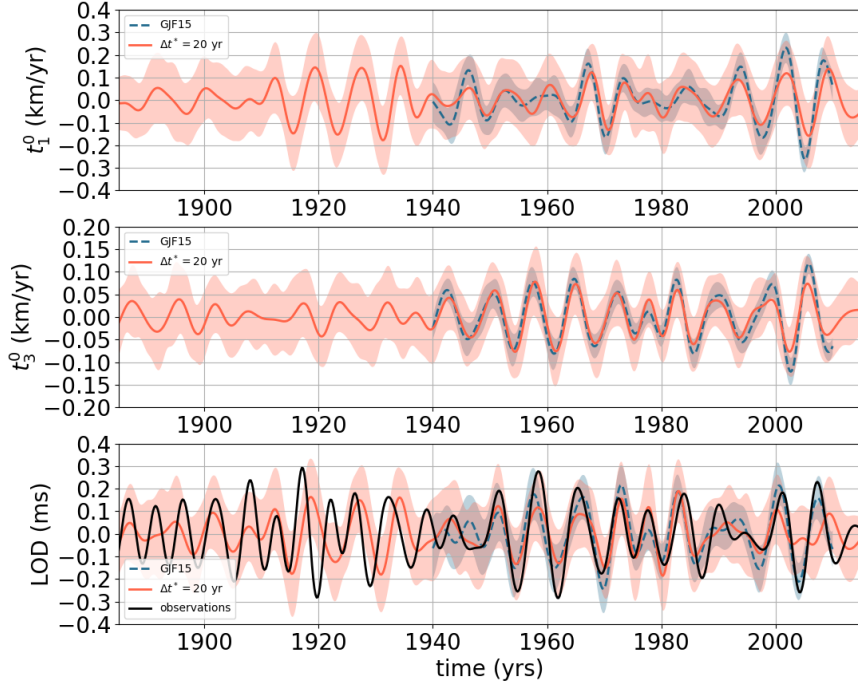


Figure 10: Band-pass filtered series (between 4 and 9.5 yrs) for  $t_1^0$  (top),  $t_3^0$  (middle) and the LOD (bottom) for our ensemble of solutions with  $\Delta t^* = 20$  yr (red), the ensemble of solutions from GJF15 (dotted green). For LOD series, the black curve represents the geophysical data. Shaded areas represent the  $\pm 1\sigma$  dispersion within the ensemble of solutions.

nual SV changes between **e** and **f**. The prior variances, implicitly stated by the stochastic model, are derived from the Mid-Path geodynamo fluctuations at decadal periods, and are continued with a spectral index  $p = 2$  towards interannual time-scales (see Figure 2). The large amount of interannual SV accounted for by  $e_r$  in the present study contrasts with Figure 4 in GJF15, possibly linked with the weak formalism they employ (as opposed to the sequential method used here). We cannot exclude relative over-estimation of  $e_r$  in the present study (in comparison with GJF15), as (i) within the augmented state framework  $e_r$  is related to SV data through the identity observation operation (see equation 17), and (ii) the temporal correlation functions for  $e_r$  in GJF15 are less sharp than the Laplacian implicitly stated by the AR-1 process used for  $e_r$ .

We recover large oscillations in  $t_1^{0\dagger}$  from 1910 onward, while the amplitude of  $t_3^{0\dagger}$  is much reduced prior to 1940. Prior to this epoch precisely, our ensemble of models do not fit well sub-decadal LOD changes, suggesting we are not able to recover interannual variations for zonal coefficients of degrees  $n \geq 3$  before 1940, and thus to obtain any hint on the propagation of torsional waves at such early epochs. This transpires in the

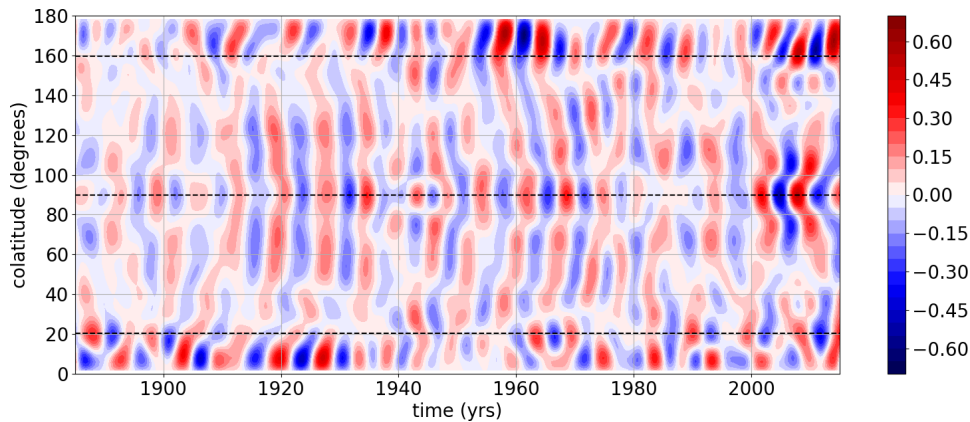


Figure 11: Zonal motions band-pass filtered for periods within 4 and 9.5 yrs (in km/yr), as a function of colatitude, for the ensemble average model obtained using  $\Delta t^* = 20$  yrs. Dotted lines represent the position of the tangent cylinder and of the equator.

time-latitude map of the filtered zonal motion,

$$u_z^\dagger(\theta, t) = (2\pi)^{-1} \oint u_\phi^\dagger(\theta, \phi, t) d\phi, \quad (22)$$

shown in Figure 11 – concerning the sensitivity of the torsional waveforms to the choice of period band-width, we refer to Gillet [2019]. The studied era can be separated in several periods:

- (i) prior to 1910, the ensemble average  $u_z^\dagger$  is very weak, and no interannual motions emerge. The larger errorbars in COV-OBS.x1 at this early era most probably do not allow to recover accurately enough interannual changes.
- (ii) over 1910–1940,  $u_z^\dagger(\theta, t)$  shows no propagation and a geostrophic velocity maximum at the equator, a logical consequence of the weak  $t_3^{0\dagger}$ , with a dominant signature of  $t_1^{0\dagger}$ .
- (iii) after 1940,  $u_z^\dagger$  is organized as equatorward propagating patterns, with an apparent velocity similar to that first highlighted by Gillet et al. [2010]. This is coherent with out-of-phase  $t_1^{0\dagger}$  and  $t_3^{0\dagger}$  series (see Figure 10).
- (iv) the signal seems less organised in the 1990's, while from 2000, more complex patterns in  $u_z^\dagger$  emerge, with features common to Figure 10 in GJF15 (though less intense).

Interestingly, we recover as GJF15 large intensities of the filtered zonal motions within the equatorial belt at latitudes below  $10^\circ$ . We also observe within the tangent cylinder particularly large values of  $u_z^\dagger$ , with a poleward propagation below the Northern polar cap. These patterns, which are not symmetric with respect to the equator, may indicate the propagation of torsional waves in the cavity encapsulated by the tangent cylinder and the inner core, as seen in numerical simulations [see Jault and L egaut, 2005, Teed et al., 2015].



## 4 Discussion

### 4.1 A primordial role played by errors of representativeness

We propose a stochastic model of the large length-scale core surface dynamics that replicates the statistics of a geodynamo numerical simulation in the spatial domain, recovers its temporal spectrum at decadal and longer periods, and proposes a compensation for its too low variations on shorter time-scales. The origin of the observed SV is partitioned between induction associated with large length-scale fields, and errors of representativeness. These latter encompass all nonlinear induction processes involving unresolved velocity and magnetic fields, plus diffusion.

By integrating our model into an EnKF algorithm, we perform a re-analysis of geomagnetic observations over the period 1880–2015. A significant fraction of interannual to decadal SV changes is accounted for by errors of representativeness. Nevertheless, the large scale flow still accounts for significant SV changes, and its resulting variations are sufficient to predict accurately interannual to decadal LOD changes over 1940–2005. The weight of the two contributions ( $\mathbf{f}$  and  $e_r$ ) to the SV directly derives from the Mid-Path geodynamo simulation.

The question of how accurately the large length-scale flows should fit the data is crucial. It appears that subgrid processes become a larger contribution to SV when increasing the rotation rate along the dynamo path [from  $E = 3 \times 10^{-5}$  to  $E = 10^{-8}$ , see Aubert et al., 2017]. Should we expect even less SV from large length-scale fields with closer to Earth-like parameters? Even if the large length-scale MF morphology is unaltered along the parameter path followed to produce Mid-Path [see Aubert et al., 2017], this issue remains open. Indeed, rapid fluctuations in large-scale motions are under-estimated in the Mid-Path dynamo: currently available models lack the needed temporal variability to capture interannual changes, the reason why we complement them with stochastic equations. However, one key ingredient is the temporal resolution at which one aims at fitting SV changes: since correlation times for  $e_r$  are shorter than for the contribution from large length-scale fields, the long period SV should be relatively less imprinted by errors of representativeness. In any case, our study warns against too closely interpreting SV changes (as recorded with unprecedented accuracy with the *Swarm* mission of ESA) as the signature of large length-scale dynamics.

### 4.2 The high latitude jet revisited

Livermore et al. [2017] highlight a monotonous increase of a westward jet velocity on the vicinity of the tangent cylinder in the Northern Pacific hemisphere – from 16 to 42 km/yr over 2000–2016, i.e. an acceleration of  $\approx 1.6$  km/yr<sup>2</sup> over 16 yrs, see their figure 5. The jet studied by Livermore et al. corresponds to the Northern Pacific portion of the eccentric gyre put forward by Pais and Jault [2008]. In our core flow reconstruction, the evolution of the jet differs in several ways from the above description (see Figure 12):

- (i) similarly to the re-analysis by Barrois et al. [2018b], and in contradiction with the symmetric model by Livermore et al. [2017], the jet recovered here is not symmetric with respect to the equator for the recent era;
- (ii) it is not monotonous over the satellite era, instead reaching a maximum intensity (of  $\approx 40$  km/yr) around 2011 in the Northern hemisphere where it is the largest (around 160°W);

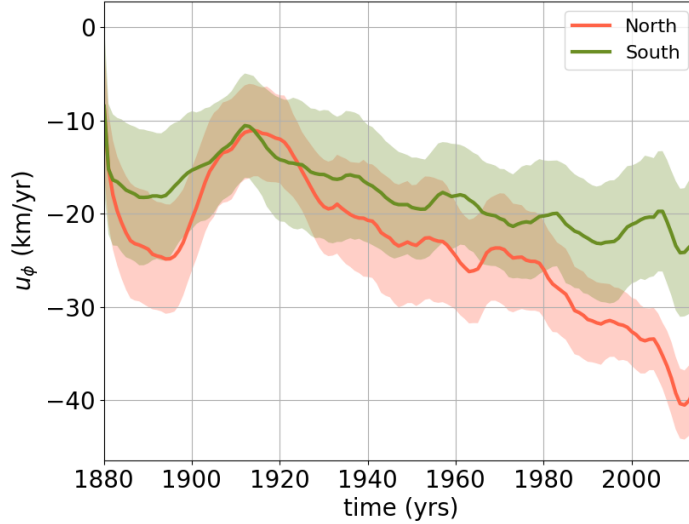


Figure 12: Azimuthal velocity  $u_\phi$  at the CMB in the vicinity of the tangent cylinder (at latitude  $68.2^\circ$ ) in both hemispheres, averaged over longitudes  $\phi \in [140, 180]^\circ\text{W}$ , for the model obtained using  $\Delta t^* = 20$  yrs. The choice of longitudes cover is such that it covers the position where the jet is the strongest over the studied era. Thick lines correspond to the ensemble average, and the shaded areas to  $\pm\sigma$  within the ensemble.

- (iii) averaged over the satellite era, the acceleration is  $\approx 0.34 \text{ km/yr}^2$ , much lower than previous estimates, essentially because the jet's speed in our reconstruction is already as large as about  $34 \text{ km/yr}$  in 2000. The simple design of the model by Livermore et al. (no errors of representativeness, reduced flow geometry) probably lead their reconstructed polar jet to overfit the SV observations.

We nevertheless witness locally a peak acceleration at about  $1.1 \text{ km/yr}^2$  around 2008, larger than any previous event in the past century, but it is not maintained over the whole satellite era (with a deceleration from 2011 onwards). Interestingly, we see in the Northern hemisphere a gradual increase of the jet from 1910 onward. Such an evolution is associated with a reinforcement of the gyre over the past century (see Figure 6), already noted by Pais et al. [2015]. Our flow solutions require in the tangent cylinder latitudes some departure from equatorial symmetry. Given the estimated uncertainties (see the errorbars in Figure 12), the existence of an ageostrophic component in this area seems constrained by observations. The equatorial asymmetry of polar vortices is expected given the impenetrability of the inner core, and has been put forward in earlier core flow inversion studies [Olson and Aurnou, 1999, Pais and Hulot, 2000, Amit and Olson, 2006].

### 4.3 A re-analysis tool for confronting geodynamo simulations to modern data

We recover with our algorithm a planetary scale, axially columnar, eccentric westward gyre over the whole studied era, back to the late 19th century. The flow is predominantly equatorially symmetric over all epochs. The main time-dependent features are non-axisymmetric and located within the equatorial belt. Nevertheless the high latitude jet in the Pacific hemisphere, which grows over the past century, shows some evidence of equatorial asymmetry especially towards the recent epochs. Our flow solutions tend to minimize induction, as the flow is often aligned with isolines of  $B_r$ . Maps of the magnetic field at the CMB, combined with the axial invariance of the flow due to the fast rotation, and with the hypothesis of field alignment in the whole core, may then provide partial indications on the field orientation deeper in the core. This has several potential consequences. First, differences in iso-contours of  $B_r$  in the Northern and Southern hemispheres would be interpreted either as a source of SV or as a sign of ageostrophy. This information could be used as a constraint for recovering the flow from maps of  $B_r$  and  $\partial_t B_r$ . Second, QG dynamical models [Jault and Finlay, 2015], which currently disconnect the field at the core surface from that in the bulk, may for this reason miss crucial information.

By confronting the result of the re-analysis (e.g. the average predictions plus the dispersion within the ensemble of forecasts) to magnetic and LOD observations, the algorithm defined in §2.2–2.3 may be seen as a tool to validate the Earth-likeness of geodynamo simulations. This could be performed by estimating the a posteriori probability of the analysed state, given the a priori spatial and temporal distributions provided by the dynamo model [for an example where only spatial constraints are used, see Aubert, 2013]. Such an approach may be used to assess whether the eccentric gyre structure results from a non axisymmetric coupling between the inner core and the mantle (pre-supposed in the Mid-Path model), or if the symmetry breaking naturally emerges from isotropic primitive equations as in the study by Schaeffer et al. [2017] – see also the discussion in Gillet [2019].

Our algorithm may also be tailored to directly accept magnetic records (instead of Gauss coefficients), as performed by Barrois et al. [2018b] with ground and virtual observatory series. As such it would become a geomagnetic assimilation tool suitable for the production of models such as the International Geomagnetic Reference Field model [Thébaud et al., 2015], with built-in 5 yr SV predictions. This is an opportunity offered by the open-source Python code proposed by Huder et al. [2019], and available at <https://gricad-gitlab.univ-grenoble-alpes.fr/Geodynamo/Pygeodyn>.

## 5 Acknowledgements

We thank Hagay Amit and an anonymous referee for their thorough reviews, which helped improve the quality of our manuscript. NG and LH were partially supported by the French Centre National d’Etudes Spatiales (CNES) for the study of Earth’s core dynamics in the context of the *Swarm* mission of ESA. JA acknowledges support from the Fondation Simone et Cino Del Duca of Institut de France (2017 Research Grant).

## References

Hagay Amit and Ulrich R Christensen. Accounting for magnetic diffusion in core flow inversions from geomagnetic secular variation. *Geophys. J. Int.*, 175(3):913–924, 2008.

- Hagay Amit, Maélie Coutelier, and Ulrich R Christensen. On equatorially symmetric and antisymmetric geomagnetic secular variation timescales. *Physics of the Earth and Planetary Interiors*, 276:190–201, 2018.
- Hagay Amit and Peter Olson. Helical core flow from geomagnetic secular variation. *Phys. Earth Planet. Int.*, 147(1):1–25, 2004.
- Hagay Amit and Peter Olson. Time-average and time-dependent parts of core flow. *Phys. Earth Planet. Int.*, 155:120–139, 2006.
- J. Aubert. Steady zonal flows in spherical shell fluid dynamos. *J. Fluid Mech.*, 542:53–67, 2005.
- J Aubert. Geomagnetic forecasts driven by thermal wind dynamics in the Earth’s core. *Geophys. J. Int.*, 203(3):1738–1751, 2015.
- Julien Aubert. Flow throughout the Earth’s core inverted from geomagnetic observations and numerical dynamo models. *Geophys. J. Int.*, 192(2):537–556, 2013.
- Julien Aubert. Earth’s core internal dynamics 1840–2010 imaged by inverse geodynamo modelling. *Geophys. J. Int.*, page ggu064, 2014.
- Julien Aubert. Geomagnetic acceleration and rapid hydromagnetic wave dynamics in advanced numerical simulations of the geodynamo. *Geophys. J. Int.*, 214(1):531–547, 2018.
- Julien Aubert, Christopher C Finlay, and Alexandre Fournier. Bottom-up control of geomagnetic secular variation by the Earth’s inner core. *Nature*, 502(7470):219–223, 2013.
- Julien Aubert, Thomas Gastine, and Alexandre Fournier. Spherical convective dynamos in the rapidly rotating asymptotic regime. *J. Fluid Mech.*, 813:558–593, 2017.
- Julien Baerenzung, Matthias Holschneider, and Vincent Lesur. The flow at the Earth’s core-mantle boundary under weak prior constraints. *J. Geophys. Res.: Solid Earth*, 121(3):1343–1364, 2016.
- Julien Baerenzung, Matthias Holschneider, Johannes Wicht, Sabrina Sanchez, and Vincent Lesur. Modeling and predicting the short-term evolution of the geomagnetic field. *J. Geophys. Res.: Solid Earth*, 123(6):4539–4560, 2018.
- O Barrois, N Gillet, and J Aubert. Contributions to the geomagnetic secular variation from a reanalysis of core surface dynamics. *Geophys. J. Int.*, 211(1):50–68, 2017.
- O Barrois, N Gillet, J Aubert, O Barrois, MD Hammer, CC Finlay, Y Martin, and N Gillet. Erratum:contributions to the geomagnetic secular variation from a reanalysis of core surface dynamics and assimilation of ground and satellite magnetic measurements: inference of core surface magnetic and velocity field changes. *Geophysical Journal International*, 216(3):2106–2113, 2018a.
- O Barrois, MD Hammer, CC Finlay, Y Martin, and N Gillet. Assimilation of ground and satellite magnetic measurements: inference of core surface magnetic and velocity field changes. *Geophys. J. Int.*, 215(1):695–712, 2018b.

- Christian Bizouard and Daniel Gambis. The combined solution C04 for Earth orientation parameters consistent with international terrestrial reference frame 2005. In *Geodetic reference frames*, pages 265–270. Springer, 2009.
- Claire Bouligand, Nicolas Gillet, Dominique Jault, Nathanaël Schaeffer, Alexandre Fournier, and Julien Aubert. Frequency spectrum of the geomagnetic field harmonic coefficients from dynamo simulations. *Geophys. J. Int.*, 207(2):1142–1157, 2016.
- Bruce Buffett, Nicholas Knezek, and Richard Holme. Evidence for MAC waves at the top of earth’s core and implications for variations in length of day. *Geophys. J. Int.*, 204(3): 1789–1800, 2016.
- A Chulliat and S Maus. Geomagnetic secular acceleration, jerks, and a localized standing wave at the core surface from 2000 to 2010. *J. Geophys. Res.: Solid Earth*, 119(3): 1531–1543, 2014.
- Arnaud Chulliat, Patrick Alken, and Stefan Maus. Fast equatorial waves propagating at the top of the Earth’s core. *Geophys. Res. Lett.*, 42(9):3321–3329, 2015.
- Geir Evensen. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.
- Céline Eymin and Gauthier Hulot. On core surface flows inferred from satellite magnetic data. *Phys. Earth Planet. Int.*, 152(3):200–220, 2005.
- Christopher C Finlay and Hagay Amit. On flow magnitude and field-flow alignment at earth’s core surface. *Geophysical Journal International*, 186(1):175–192, 2011.
- Christopher C Finlay, Julien Aubert, and Nicolas Gillet. Gyre-driven decay of the Earth’s magnetic dipole. *Nature communications*, 7:10422, 2016a.
- Christopher C Finlay, Nils Olsen, Stavros Kotsiaros, Nicolas Gillet, and Lars Tøffner-Clausen. Recent geomagnetic secular variation from Swarm. *Earth, Planets and Space*, 68(1):1–18, 2016b.
- A. Fournier, J. Aubert, and E. Thébault. Inference on core surface flow from observations and 3-D dynamo modelling. *Geophys. J. Int.*, 186:118–136, 2011.
- Alexandre Fournier, Julien Aubert, and Erwan Thébault. A candidate secular variation model for IGRF-12 based on Swarm data and inverse geodynamo modelling. *Earth, Planets and Space*, 67(1):1–17, 2015.
- Alexandre Fournier, Lars Nerger, and Julien Aubert. An ensemble Kalman filter for the time-dependent analysis of the geomagnetic field. *Geochem. geophys. geosyst.*, 14(10): 4035–4043, 2013.
- N Gillet, D Jault, and CC Finlay. Planetary gyre, time-dependent eddies, torsional waves, and equatorial jets at the Earth’s core surface. *J. Geophys. Res.: Solid Earth*, 120(6): 3991–4013, 2015a.
- Nicolas Gillet. 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*. 2019. URL <https://hal.archives-ouvertes.fr/hal-02042703>.

- Nicolas Gillet, Olivier Barrois, and Christopher C Finlay. 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):1–14, 2015b.
- Nicolas Gillet, Dominique Jault, and Elisabeth Canet. Excitation of travelling torsional normal modes in an earths core model. *Geophys. J. Int.*, 210(3):1503–1516, 2017.
- Nicolas Gillet, Dominique Jault, Elisabeth Canet, and Alexandre Fournier. Fast torsional waves and strong magnetic field within the Earths core. *Nature*, 465(7294):74–77, 2010.
- Nicolas Gillet, Dominique Jault, CC Finlay, and Nils Olsen. Stochastic modeling of the Earth’s magnetic field: inversion for covariances over the observatory era. *Geochem. geophys. geosyst.*, 14(4):766–786, 2013.
- Nicolas Gillet, MA Pais, and Dominique Jault. Ensemble inversion of time-dependent core flow models. *Geochem. geophys. geosyst.*, 10(6), 2009.
- Nicolas Gillet, Nathanaël Schaeffer, and Dominique Jault. Rationale and geophysical evidence for quasi-geostrophic rapid dynamics within the Earth’s outer core. *Phys. Earth Planet. Int.*, 187(3):380–390, 2011.
- Richard S Gross. A combined length-of-day series spanning 1832–1997: Lunar97. *Phys. Earth Planet. Int.*, 123(1):65–76, 2001.
- R. Holme. Large scale flow in the core. In P. Olson and G. Schubert, editors, *Treatise in Geophysics, Core Dynamics*, volume 8, chapter 4, pages 91–113. Elsevier, 2015.
- Loic Huder, Nicolas Gillet, and Franck Thollard. Pygeost: a python package for geomagnetic data assimilation using stochastic processes, *submitted to Geoscientific Model Development*. 2019.
- Andrew Jackson, Jeremy Bloxham, and David Gubbins. Time-dependent flow at the core surface and conservation of angular momentum in the coupled core-mantle system. *Dynamics of Earth’s Deep Interior and Earth Rotation*, pages 97–107, 1993.
- D. Jault and C. C. Finlay. Waves in the core and mechanical core-mantle interactions. In Gerald Schubert and P. Olson, editors, *Treatise on Geophysics, Core Dynamics, 2nd edition*, volume 8, chapter 8.09, pages 225–244. Elsevier, Oxford, 2015.
- D Jault and G Légaut. Alfvén waves within the earths core. *Fluid dynamics and dynamos in astrophysics and geophysics*, pages 277–293, 2005.
- Dominique Jault. Axial invariance of rapidly varying diffusionless motions in the earths core interior. *Physics of the Earth and Planetary Interiors*, 166(1-2):67–76, 2008.
- Clemens Kloss and Christopher C Finlay. Time-dependent low-latitude core flow and geomagnetic field acceleration pulses. *Geophysical Journal International*, 217(1):140–168, 2019.
- F Labbé, D Jault, and N Gillet. On magnetostrophic inertia-less waves in quasi-geostrophic models of planetary cores. *Geophysical & Astrophysical Fluid Dynamics*, 109(6):587–610, 2015.

- Vincent Lesur, Ingo Wardinski, Julien Baerenzung, and Matthias Holschneider. On the frequency spectra of the core magnetic field gauss coefficients. *Physics of the Earth and Planetary Interiors*, 276:145–158, 2018.
- Philip W Livermore, Rainer Hollerbach, and Christopher C Finlay. An accelerating high-latitude jet in Earth’s core. *Nature Geoscience*, 10(1):62–68, 2017.
- D. Lloyd and D Gubbins. Toroidal fluid motions at the top of Earth’s core, 1990.
- Stefano Maffei and Andrew Jackson. Kinematic validation of a quasi-geostrophic model for the fast dynamics in the earths outer core. *Geophys. J. Int.*, 210(3):1772–1786, 2017.
- M Mandeau, R. Holme, A. Pais, K. Pinheiro, A. Jackson, and G. Verbanac. Geomagnetic jerks: rapid core field variations and core dynamics. *Space Sci. Rev.*, 155:147–155, 2010.
- Maurits C Metman, Philip W Livermore, Jonathan E Mound, and Ciarán D Beggan. Modelling decadal secular variation with only magnetic diffusion. *Geophys. J. Int.*, 2019.
- H.-C. Nataf and N. Schaeffer. Turbulence in the core. In Gerald Schubert and P. Olson, editors, *Treatise on Geophysics, Core Dynamics, 2nd edition*, volume 8, pages 161–181. Elsevier, Oxford, 2015.
- Arnold Neumaier and Tapio Schneider. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software (TOMS)*, 27(1):27–57, 2001.
- Peter Olson and Jonathan Aurnou. A polar vortex in the Earth’s core. *Nature*, 402(6758):170–173, 1999.
- A Pais and G Hulot. Length of day decade variations, torsional oscillations and inner core superrotation: evidence from recovered core surface zonal flows. *Phys. Earth Planet. Int.*, 118(3-4):291–316, 2000.
- MA Pais and Dominique Jault. Quasi-geostrophic flows responsible for the secular variation of the Earth’s magnetic field. *Geophys. J. Int.*, 173(2):421–443, 2008.
- Maria A Pais, Anna L Morozova, and Nathanaël Schaeffer. Variability modes in core flows inverted from geomagnetic field models. *Geophys. J. Int.*, 200(1):402–420, 2015.
- Steffen Rau, Ulrich Christensen, Andrew Jackson, and Johannes Wicht. Core flow inversion tested with numerical dynamo models. *Geophys. J. Int.*, 141(2):485–497, 2000.
- Sabrina Sanchez. *Assimilation of geomagnetic data into dynamo models, an archeomagnetic study*. PhD thesis, Université Sorbonne Paris Cité, 2016.
- Yoshikazu Sasaki. Some basic formalisms in numerical variational analysis. *Monthly Weather Review*, 98(12):875–883, 1970.
- Nathanaël Schaeffer, Dominique Jault, H-C Nataf, and Alexandre Fournier. Turbulent geodynamo simulations: a leap towards Earth’s core. *Geophys. J. Int.*, 211(1):1–29, 2017.

- Francis Richard Stephenson, Leslie V Morrison, and Gerald James Whitrow. Long-term changes in the rotation of the Earth: 700 bc to ad 1980 [and discussion]. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 313(1524):47–70, 1984.
- Robert J Teed, Chris A Jones, and Steven M Tobias. The dynamics and excitation of torsional waves in geodynamo simulations. *Geophys. J. Int.*, 196(2):724–735, 2014.
- Robert J Teed, Chris A Jones, and Steven M Tobias. The transition to earth-like torsional oscillations in magnetoconvection simulations. *Earth Planet. Sci. Lett.*, 419:22–31, 2015.
- Erwan Thébault, Christopher C Finlay, Ciarán D Beggan, Patrick Alken, Julien Aubert, Olivier Barrois, Francois Bertrand, Tatiana Bondar, Axel Boness, Laura Brocco, et al. International geomagnetic reference field: the 12th generation. *Earth, Planets and Space*, 67(1):1–19, 2015.
- Steven M Tobias, Fausto Cattaneo, and Stanislav Boldyrev. MHD dynamos and turbulence. In *Ten chapters on turbulence*, pages 305–404. Cambridge University Press, 2011.
- K. Whaler. Geomagnetic evidence for fluid upwelling at the core-mantle boundary. *Geophys. J. R. Astr. Soc.*, 86:563–588, 1986.
- Johannes Wicht and U. R. Christensen. Torsional oscillations in dynamo simulations. *Geophys. J. Int.*, 181(3):1367–1380, 2010.

## A Geodynamo versus stochastic model

We show in Figure 13 the target (geodynamo) correlation matrices  $C_{uu}^*$  and  $C_{ee}^*$ , together with the difference between these latter and the correlation matrices  $C_{uu}$  and  $C_{ee}$  of the fields obtained after numerical integration of the stochastic model (9), using matrices as defined in equation (14), over the same time-span (20,200 yr) and  $\Delta t = 0.2$  yr. Except for some noise level, our stochastic model reproduces for both  $e$  and  $u$  the spatial cross-correlations present in the geodynamo simulation.

We also compare in Figure 14 the magnitude of the stochastic model series to those of the geodynamo model. The r.m.s. of the stochastic model coefficients is slightly less than that of the target simulation. With longer sampling rates  $\Delta t^*$ , the loss of energy is more important, and is relatively larger towards small length-scales. Figure 14 shows that this effect is not predominately the consequence of the projection onto the PC. For this reason we neglect in our study the loss of information due to the projection on the PC, which is in any case secondary in comparison with the effect of errors of representativeness.

## B Coherence of the solutions

In order to measure the coherence of the solution to the inverse problem described in §2.3 with respect to both the geodynamo priors and the geomagnetic data, we define normalised



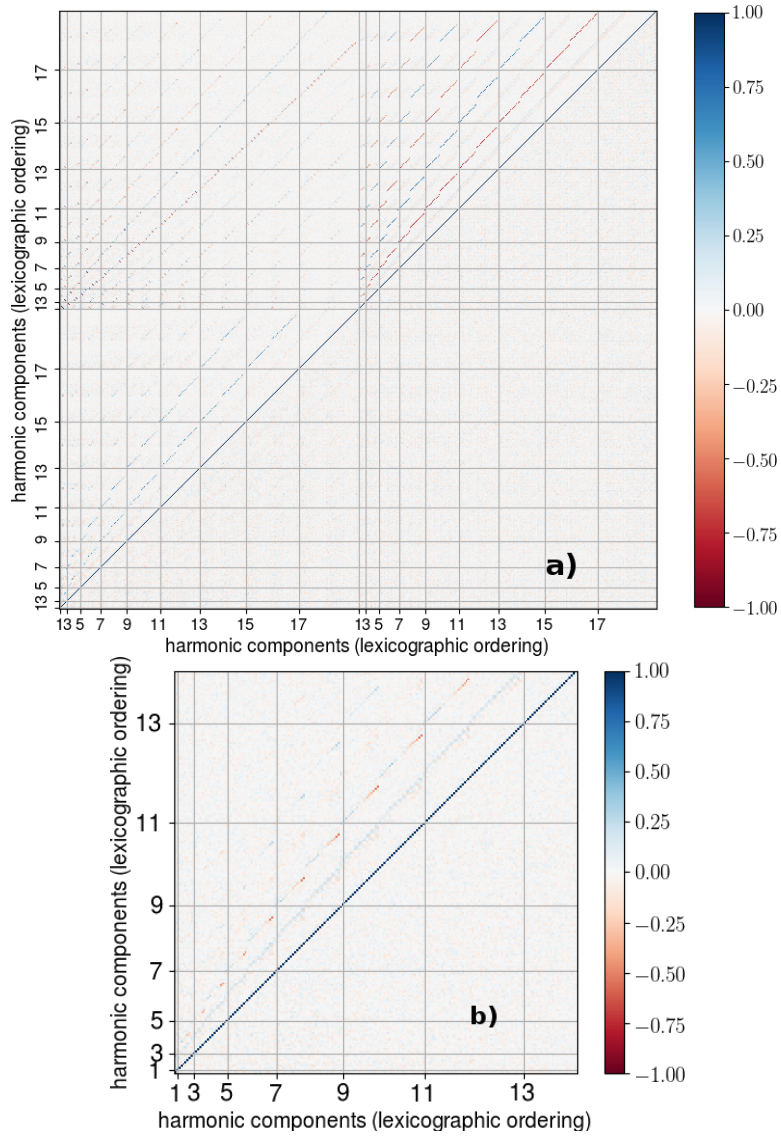


Figure 13: Upper diagonal blocks: correlation matrices for coefficient series of the vector  $\mathbf{u}$  (a) and  $\mathbf{e}$  (b) from the Mid-Path dynamo sampled every  $\Delta t^* = 5$  yrs. Lower diagonal blocks: difference between these matrices and the correlation matrices obtained after integration of the stochastic model. Spherical harmonic components on the axes are lexicographically ordered as follows:  $(n, m) = (1, 0), (1, -1), (2, 0), (2, 1), (2, -1), (2, 2), (2, -2), (3, 0) \dots$  – for  $\mathbf{u}$  toroidal coefficients are stored before the poloidal ones.

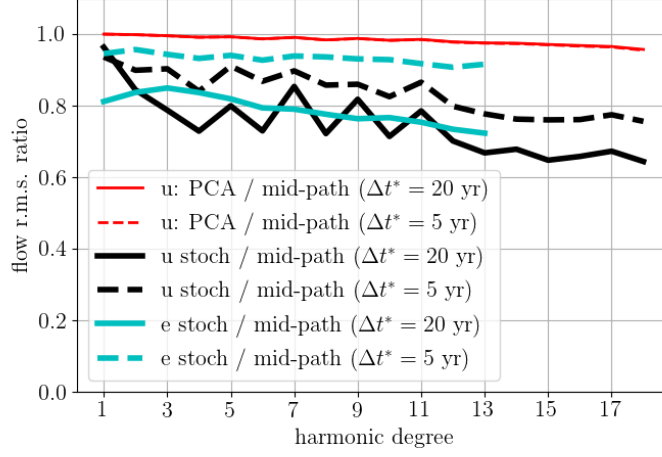


Figure 14: R.m.s. of  $\mathbf{u}$  (black) and  $\mathbf{e}$  (cyan) coefficients obtained after integration of the stochastic model (average over harmonic orders) as a function of spherical harmonic degree, normalised by the r.m.s. of the sampled Mid-Path dynamo series, for  $\Delta t^* = 5$  yrs (dotted lines) and  $\Delta t^* = 20$  yrs (full lines). In red the reduction effect of the PCA on the geodynamo series.

measures of SV prediction errors, core surface flow and errors of representativeness terms:

$$\left\{ \begin{array}{l} \chi_b(t) = \left[ \frac{1}{N_e} \sum_k \left( \dot{\mathbf{b}}_k^o(t) - \mathbf{H}_k \mathbf{z}_k^a(t) \right)^T \mathbf{R}_{bb}^{-1} \left( \dot{\mathbf{b}}_k^o(t) - \mathbf{H}_k \mathbf{z}_k^a(t) \right) \right]^{1/2}, \\ \chi_u(t) = \left[ \frac{1}{N_e} \sum_k \left( \dot{\mathbf{u}}_k^a(t) - \mathbf{u}_0 \right)^T \mathbf{P}_{uu}^{-1} \left( \dot{\mathbf{u}}_k^a(t) - \mathbf{u}_0 \right) \right]^{1/2}, \\ \chi_e(t) = \left[ \frac{1}{N_e} \sum_k \left( \dot{\mathbf{e}}_k^a(t) - \mathbf{e}_0 \right)^T \mathbf{P}_{ee}^{-1} \left( \dot{\mathbf{e}}_k^a(t) - \mathbf{e}_0 \right) \right]^{1/2}. \end{array} \right. \quad (23)$$

The above normalised measures, shown in Figure 15, are rather independent of our choice for the sampling period  $\Delta t^*$  for the geodynamo. They illustrate that, using the geodynamo prior cross-covariance matrix as a measure, the ensemble of flow models lie within  $\pm 1\sigma$  over the whole studied era, and the ensemble of  $e_r$  models is around  $\pm 1\sigma$  over the whole studied era, with values slightly above unity prior to 1980. It appears that the ensemble of augmented state solutions fits SV observations well within  $\pm 1\sigma$ . After a spin-up time of 5 to 10 yrs,  $\chi_b$  reaches  $\approx 0.25$  on average (using the COV-OBS.x1 prior data uncertainties as a measure). Such low values question on a possible overfitting of SV data. This may be related either to the employed algorithm, or to a considered geodynamo model rendering too much errors of representativeness.

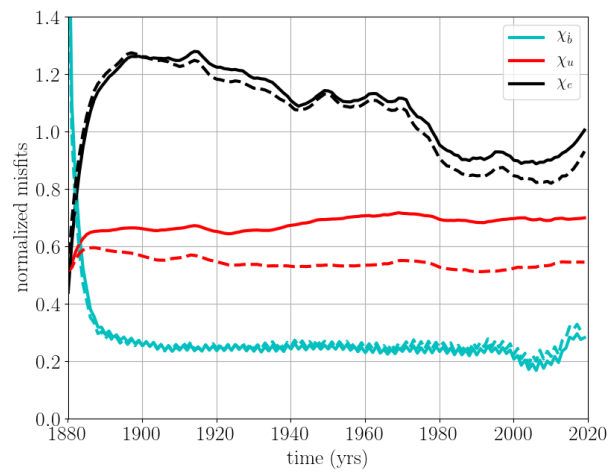


Figure 15: SV data misfit  $\chi_b$  (cyan), core surface flow misfit  $\chi_u$  (red) and representativeness error misfit  $\chi_e$  (black) as a function of time, for  $\Delta t^* = 5$  yrs (full lines) and 20 yrs (dotted lines).