Baerenzung, J., Holschneider, M., Wicht, J., Sanchez, S., Lesur, V. (2018): Modeling and Predicting the Short-Term Evolution of the Geomagnetic Field. - *Journal of Geophysical Research*, *123*, 6, pp. 4539—4560.

DOI: http://doi.org/10.1029/2017JB015115

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1029/2017JB015115

Special Section:
Magnetism in the Geosciences - Advances and Perspectives

Modeling and Predicting the Short-Term Evolution of the Geomagnetic Field

Julien Bärenzung[1,2], Matthias Holschneider[1], Johannes Wicht[3], Sabrina Sanchez[3], and Vincent Lesur[4]

[1]Institute for Mathematics, University of Potsdam, Potsdam, Germany, [2]German Research Centre for Geosciences (GFZ), Potsdam, Germany, [3]Max Planck Institute for Solar System Research, Göttingen, Germany, [4]Institut de Physique du Globe de Paris, Paris, France

Abstract We propose a reduced dynamical system describing the coupled evolution of fluid flow and magnetic field at the top of the Earth's core between the years 1900 and 2014. The flow evolution is modeled with a first-order autoregressive process, while the magnetic field obeys the classical frozen flux equation. An ensemble Kalman filter algorithm serves to constrain the dynamics with the geomagnetic field and its secular variation given by the COV-OBS.x1 model. Using a large ensemble with 40,000 members provides meaningful statistics including reliable error estimates. The model highlights two distinct flow scales. Slowly varying large-scale elements include the already documented eccentric gyre. Localized short-lived structures include distinctly ageostophic features like the high-latitude polar jet on the Northern Hemisphere. Comparisons with independent observations of the length-of-day variations not only validate the flow estimates but also suggest an acceleration of the geostrophic flows over the last century. Hindcasting tests show that our model outperforms simpler predictions bases (linear extrapolation and stationary flow). The predictability limit, of about 2,000 years for the magnetic dipole component, is mostly determined by the random fast varying dynamics of the flow and much less by the geomagnetic data quality or lack of small-scale information.

Key Points:
- The model we propose permits to estimate the reliability of certain patterns the flow at the Earth's core-mantle boundary is exhibiting
- The model we present enables to predict the evolution of the geomagnetic field and to accurately estimate prediction errors
- We show that the discrepancies between the predicted trend in length of day and the observed one can be explained by our flow model

Correspondence to:
J. Bärenzung,
baerenzung@gmx.de

Citation:
Bärenzung, J., Holschneider, M., Wicht, J., Sanchez, S., & Lesur, V. (2018). Modeling and predicting the short-term evolution of the geomagnetic field. *Journal of Geophysical Research: Solid Earth*, *123*, 4539–4560. https://doi.org/10.1029/2017JB015115

Received 17 OCT 2017
Accepted 29 APR 2018
Accepted article online 9 MAY 2018
Published online 19 JUN 2018

1. Introduction

In the Earth's outer core, turbulent motion of the electrically conducting fluid sustains the geomagnetic field through dynamo action. Part of this field, the poloidal one, crosses the mantle and can be observed at the Earth's surface and above. Because of the low conductivity of the mantle (see Jault, 2015; Velímský, 2010), once measured and modeled, the poloidal field can be estimated everywhere outside and at the outer boundary of the core (core-mantle boundary, CMB). The advection of the magnetic field at the CMB by the underlying flow produces geomagnetic field changes that can be observed at the Earth's surface. These changes are known as geomagnetic secular variation, and their close examination can thus allow for inferences on the core surface flow.

Within the outer core, the evolution of the magnetic field is prescribed by the induction equation. Under the assumption that the mantle is a perfect electrical insulator, the magnetic toroidal field, which interacts with the poloidal field inside the outer core, vanishes at the CMB. In addition, since the fluid cannot penetrate the mantle, its associated velocity field is purely two dimensional. Finally, on short timescales, diffusion can be considered as negligible in comparison to advection as shown in Roberts and Scott (1965). All in all, the induction equation expressed at the CMB can be simplified into the so-called frozen flux (FF) approximation (see Backus et al., 1996; Kahle et al., 1967). By inverting this equation, which couples the velocity field to the radial component of the magnetic field and the secular variation, fluid motions at the CMB can be recovered.

However, since the velocity field has two components for one equation, and since any flow scale can interact with the magnetic field to generate the large-scale, observable, secular variation, the inverse problem is ill posed. Additional physical assumptions help to reduce the nonuniqueness of the velocity field and thus to decrease the dimension of possible solutions. Constraints generally used in core flow inversions, such as quasi-geostrophy, columnar, tangential-geostrophy, or purely toroidal, are, for example, described in Finlay et al. (2010) and Holme (2015). Nevertheless, to obtain a unique solution, additional constraints on the velocity field need to be imposed. Typically, one enforces the energy associated with the small-scale velocity field

to rapidly decay, based on the so-called large-scale assumption (see Finlay et al., 2010; Holme, 2015). However, Baerenzung et al. (2016) have shown that although the flow is dominant at large scales, its total kinetic energy spectrum does not exhibit a strong decaying slope. Recently, other strategies have been developed to bypass the issues raised by the nonuniqueness of the velocity field. In particular, Aubert (2014) used the statistical properties of an Earth-like geodynamo simulation (the coupled Earth model of Aubert, 2013) to constrain the flow and the magnetic field a priori. A major advantage of such an approach is that the correlations between the fields at the CMB and the fields within the outer core are available allowing for the imaging of the entire outer core state.

Constraining priorily the temporal dependency of the velocity field is a more delicate operation than constraining it spatially. Optimally, one should account for the dynamics of the outer core fluid and magnetic field prescribed by the magnetohydrodynamic equations. Approaches such as variational data assimilation (see Canet et al., 2009; Li et al., 2014) allow to implement a physical dynamical model into the inversion framework. To do so, the method searches for the optimal initial conditions of the system in order for the deterministic trajectories of the different fields to explain at best the observations. The drawback of the variational method is that all data are treated simultaneously, so whenever the dimension of data or their amount are large, the algorithms become computationally expensive. An alternative to avoid such a block inversion is to operate sequentially. Kuang and Tangborn (2008) were the first to adopt a sequential assimilation algorithm in the context of geomagnetic modeling. The optimal interpolation algorithm they used proceeds recursively in two steps. In the first one, referred as the forecast, the state variables are propagated in space and time with a given physical model, in the case of Kuang and Tangborn (2008), a three-dimensional geodynamo simulation. Once observations become available, the analysis is initiated, and the state variables are corrected to serve as input for the next prediction step. Since with this approach uncertainties are not modeled, they have to be specified in an ad hoc manner. The Kalman filter (KF) is a broadly used algorithm that allows the estimation of an optimal model and its associated uncertainties (see Cohn, 1997; Evensen, 2003; Kalman, 1960; Talagrand, 1997), also proceeding through a sequence of forecast and analysis. The main advantage compared to the approach used by Kuang and Tangborn (2008) is that in the KF the evolution of the errors are also predicted, and whenever data become available, these errors are taken into account for the Bayesian update of the state variables. However, the KF can only be applied to systems exhibiting linear dynamics. When the dynamics of the system are nonlinear, as it is the case here for the geomagnetic field, the propagation of errors cannot be analytically derived. However, it can be either approximated by linearization with the extended version of the KF or represented through an ensemble of possible solutions as in the ensemble Kalman filter (EnKF). Although the forecasting potential of the EnKF has already been exploited in geomagnetic studies (see Aubert, 2015; Beggan & Whaler, 2009; Gillet, Barrois, & Finlay, 2015), it is only recently that its ability to assimilate data has been taken into account (see Barrois et al., 2017).

In the EnKF, the different fields of interest are represented through an ensemble of possible states. At the prediction step, the dynamical model of the system prescribes the spatiotemporal evolution of each individual member of the ensemble. At the analysis, covariances deriving from the forecasted fields and observations are combined in order to correct the state of the predicted ensemble. Due to limitations in available computational power, a balance between complexity of the dynamical model and size of the ensemble has to be found. Here we decided to favor accuracy in statistical representation over model complexity. The evolution of the core magnetic and velocity fields is solely modeled at the CMB level, through, respectively, the FF equation and a first-order autoregressive process (AR1). Extending the approach of Baerenzung et al. (2016) to the time domain, the parameters of the autoregressive process are assumed to derive from scale-dependent power laws and are directly estimated with the COV-OBS.x1 core magnetic field secular variation model of Gillet et al. (2013) and Gillet, Barrois, and Finlay (2015).

The article is organized as follows. Section 2 describes the physical model and the mathematical approach chosen to tackle the inverse problem. The framework is then applied to the real geophysical context, and the results shown in section 3. Finally, conclusions are drawn in section 4.

2. Modeling Strategy

2.1. Quantities of Interest and Notations

In this study three fields are of particular interest, the radial component of the magnetic field $B_r(x, t)$, the annual secular variation $\partial_t B_r(x, t)$, and the velocity field $u(x, t)$, all expressed at the Earth's CMB described by x.

The spectral counter parts of $B_r(x,t)$ and $\partial_t B_r(x,t)$ are, respectively, given by the spherical harmonics (SH) coefficients $b_{l,m}$ and $\gamma_{l,m}$ such as the following:

$$B_r(x,t) = -\sum_{l=1}^{l=+\infty}(l+1)\sum_{m=-l}^{m=+l} b_{l,m}(t)Y_{l,m}(x)\,, \tag{1}$$

$$\partial_t B_r(x,t) = -\sum_{l=1}^{l=+\infty}(l+1)\sum_{m=-l}^{m=+l} \gamma_{l,m}(t)Y_{l,m}(x)\,, \tag{2}$$

with $Y_{l,m}(x)$ being the Schmidt seminormalized SH of degree l and order m.

The two-dimensional velocity field at the CMB $u(x,t)$ is decomposed into a poloidal $\phi(x,t)$ and toroidal $\psi(x,t)$ scalar field such as the following:

$$u(x,t) = r\times\nabla_H\psi(x,t) + \nabla_H(|r|\phi(x,t))\,, \tag{3}$$

where ∇_H corresponds to the horizontal divergence operator and r is the radius of the CMB. In spectral space, the poloidal and toroidal fields are, respectively, derived from the coefficients $\phi_{l,m}$ and $\psi_{l,m}$ through the formulation:

$$\phi(x,t) = \sum_{l=1}^{l=+\infty}\sum_{m=-l}^{m=+l}\phi_{l,m}(t)Y_{l,m}(x)\,, \tag{4}$$

$$\psi(x,t) = \sum_{l=1}^{l=+\infty}\sum_{m=-l}^{m=+l}\psi_{l,m}(t)Y_{l,m}(x)\,. \tag{5}$$

The magnetic field and secular variation energy spectra are given by the following:

$$E_b(l) = (l+1)\sum_{m=-l}^{m=l} b_{l,m}^2 \tag{6}$$

$$E_\gamma(l) = (l+1)\sum_{m=-l}^{m=l} \gamma_{l,m}^2\,. \tag{7}$$

and the velocity field $u(x,t)$ poloidal and toroidal energy spectra, respectively, as

$$E_\phi(l) = \frac{l(l+1)}{2l+1}\sum_{m=-l}^{m=l}\phi_{l,m}^2 \tag{8}$$

$$E_\psi(l) = \frac{l(l+1)}{2l+1}\sum_{m=-l}^{m=l}\psi_{l,m}^2\,. \tag{9}$$

In the following, we will use normal characters and bold characters. Normal characters will refer to vectors containing the SH representation for one instant in time (epoch); for example,

$$b(t_0) = b_0 = (b_{l=1,m=0}, b_{l=1,m=1}, b_{l=1,m=-1}, \ldots)^T\,. \tag{10}$$

Bold characters refer to all single epoch vectors; for example,

$$\mathbf{b} = (b_0, b_1, \ldots b_{N-1})^T\,. \tag{11}$$

Finally, a constant formalism for some statistical quantities will be used all over the manuscript. The mean value of a distribution $p(a)$ will be written with an over bar such as the following:

$$\bar{a} = \int a p(a)\mathrm{d}a \tag{12}$$

and the covariance associated with a random variable a or between a random variable a and a random variable b will respectively be expressed as follows:

$$\Sigma_a = \overline{(a-\bar{a})(a-\bar{a})^T} \tag{13}$$

$$\Sigma_{ab} = \overline{(a-\bar{a})\left(b-\bar{b}\right)^T}. \tag{14}$$

Note that the use of bold characters for space time-dependent variables also applies to statistical quantities.

2.2. Sequential Assimilation of the Core Secular Variation and Magnetic Field

We combine a dynamic model for the magnetic and velocity fields at the CMB with the geomagnetic field model COV-OBS.x1 model by Gillet et al. (2013) and Gillet, Jault, and Finlay (2015), derived from geomagnetic data, through an EnKF approach (Evensen, 2003). The EnKF method allows the sequential assimilation of data within a two-step procedure: the forecast and the analysis. In the forecast, an ensemble of possible solutions of the magnetic and velocity fields is evolved in time until data become available. In the analysis, the predictions of the ensemble are then corrected accordingly to the data. Details on the implementation of these two steps are given in the following.

2.2.1. Forecast

2.2.1.1. Magnetic Field

As mentioned in section 1, under the assumption that the observed secular variation is solely induced by advection of the magnetic field at the CMB, the dynamical evolution of B_r is expressed by the FF equation as following: $\partial_t B_r(x,t) = -\nabla_H\left(u(x,t)B_r(x,t)\right)$. Based on the notations given in section 2.1, this equation can be written in spectral space as

$$\partial_t b(t) = \gamma(t) = -A_{b(t)}u(t) = -A_{u(t)}b(t) = -A\left(u(t)b(t)\right), \tag{15}$$

where the linear operators A_b and A_u and the third-order tensor A allow us to calculate the SH coefficients associated with the advection term $\nabla_H(u(x,t)B_r(x,t))$ when they are respectively applied to u, b and (ub).

2.2.1.2. Velocity Field

The dynamical evolution of the fluid within the Earth's outer core is described by the Navier-Stokes equations. Numerically solving this equation is not only computationally expensive; it remains out of reach in Earth regime. Since on short timescales the observable magnetic field and secular variation only depend on the velocity field at the CMB, we model the outer core flow only on this surface. Following Gillet, Jault, and Finlay (2015), we choose a first-order autoregressive process (AR1) to do so. In its continuous form, this process reads

$$\partial_t u_{l,m}(t) + \frac{1}{\tau(l)}u_{l,m}(t) = \sigma(l)\dot{\omega}(t), \tag{16}$$

where $\dot{\omega}(t)$ is a Gaussian white noise process and where $u_{l,m}$ stand for either $\phi_{l,m}$ or $\psi_{l,m}$. The characteristic time $\tau(l)$ and the scaling factor $\sigma(l)$ are the scale-dependent parameters of the AR1 process. Note that these parameters will differ between the poloidal and toroidal part of the velocity field and that they will be directly estimated with the observed secular variation and core magnetic field following the procedure given in section 2.3.

2.2.2. Analysis

For the analysis step of the EnKF algorithm, the ensemble of state variables characterizing the modeled system is predicted at observation time to be corrected with the data d^o. If the latter ensemble is referred as $\{x^f\}$, and if the observations are related to the state variables through the relation $d^o = Hx + \xi^d$, where ξ^d is a normally distributed measurement noise with a 0 mean and a covariance matrix Σ_d^o, then each ensemble member x_k^f can be updated accordingly to the formulation:

$$x_k^a = x_k^f + \Sigma_{x^f}H^T\left(H\Sigma_{x^f}H^T + \Sigma_d^o\right)^{-1}\left(d_k^o - Hx_k^f\right), \tag{17}$$

where d_k^o corresponds to the data perturbed by random realizations of the measurement uncertainties (ξ_k^d) and where we recall that the covariance of a field x is referred as Σ_x and the covariance between a field x and a field y is expressed as Σ_{xy}. In equation (17), whereas the product $\Sigma_{x^f}H^T$ expresses the covariance between the state variables and the observations, the matrix $H\Sigma_{x^f}H^T + \Sigma_d^o$ describes the possible variations of the data around their mean prediction $H\bar{x}_f$.

In this study, magnetic field and secular variation data are simultaneously used to correct an ensemble of predicted magnetic field and velocity field $\{u^f, b^f\}$. The data are taken from the Gaussian model COV-OBS.x1 of Gillet, Barrois, and Finlay (2015) with a 2-year time interval Δt^A, which corresponds to the knots of the model's B-spline expansion. At a given epoch, the "observed" magnetic field b^o and secular variation γ^o are characterized by the normal distributions:

$$p(b^o) = \mathcal{N}\left(\bar{b}^o, \Sigma_b^o\right) \tag{18}$$

$$p(\gamma^o) = \mathcal{N}\left(\bar{\gamma}^o, \Sigma_\gamma^o\right), \tag{19}$$

where the covariance matrices Σ_b^o and Σ_γ^o are derived from the 100 ensemble members provided by the COV-OBS.x1 model. Because of the singular nature of these matrices, only their diagonal part is kept.

To correct each pair (u_k^f, b_k^f) of the kth forecasted ensemble member with the observations b^o and γ^o, a prediction for the observables is built according to the following relations:

$$\gamma_k^f = -A\left(u_k^f b_k^f\right) \tag{20}$$

$$b_k^{f<} = Hb_k, \tag{21}$$

where the predicted secular variation γ^f is given by the FF approximation of relation (15) and where the linear operator H simply truncates the forecasted magnetic field at the level of the observed one leading to the large-scale magnetic field $b^{f<}$.

From the ensemble $\left\{u^f, b^f, \gamma^f, Hb^f\right\}$ the covariances necessary for the analysis step of the EnKF are calculated and used to update each pair of predicted velocity field and magnetic field with the relation:

$$\begin{pmatrix} u_k^a \\ b_k^a \end{pmatrix} = \begin{pmatrix} u_k^f \\ b_k^f \end{pmatrix} + \begin{pmatrix} \Sigma_{u\gamma}^f \Sigma_{ub}^f H^T \\ \Sigma_{b\gamma}^f \Sigma_b^f H^T \end{pmatrix} \times \begin{pmatrix} \Sigma_\gamma^f + \Sigma_\gamma^o \Sigma_{\gamma b}^f H^T \\ H\Sigma_{b\gamma}^f H\Sigma_b^f H^T + \Sigma_b^o \end{pmatrix}^{-1} \begin{pmatrix} \gamma_k^o - \gamma_k^f \\ b_k^o - Hb_k^f \end{pmatrix}, \tag{22}$$

where γ_k^o and b_k^o are random realizations from the distributions of the COV-OBS.x1 model given in equations (18) and (19).

2.2.3. Numerical Implementation of the Forecast

To predict the evolution of an ensemble of velocity and magnetic fields $\{u, b\}$, equations (15) and (16) have to be numerically solved. For the magnetic field, this operation is not straightforward. Indeed, since the FF equation does not contain any diffusion mechanism, cascading magnetic energy will have a tendency to accumulate on the smallest simulated scales and to slowly contaminate the entire field through nonlinear interactions. An extra hyperdiffusion term is thus added to the FF approximation so that the evolution of the magnetic field is then prescribed by the following equation:

$$\partial_t b = -A_u b - \eta_D \Delta^4 b, \tag{23}$$

where the hypperdiffusivity term is set to $\eta_D = 9 \times 10^{13}$ km^8/year. This value is chosen so that over 100 years of pure diffusion, the magnetic energy at SH degrees 13, 26, and 39, respectively, decreases by 0.09%, 18%, and 99%. With the discretized (in time) version of equation (23), each member b_k of the ensemble of magnetic field $\{b\}$ is numerically forecasted. However, since the time step of the analysis $\Delta t^a = 2$ years is too large to ensure the stability of the algorithm, a smaller time step of $\Delta t^f = 0.5$ year has been used to predict the evolution of each b_k. Whereas the first forecast iteration is performed with an Euler scheme, the following ones are achieved with a second-order Adams-Bashforth scheme, such as the following:

$$b_k(t + \Delta t^f) = b_k(t) + \Delta t^f F_k(t) \tag{24}$$

$$b_k(t + (i+1)\Delta t^f) = b_k(t + i\Delta t^f) + \Delta t^f \left(\frac{3}{2}F_k(t + i\Delta t^f) - \frac{1}{2}F_k(t - i\Delta t^f)\right), i \geq 1, \tag{25}$$

where $F_k = -A_{u_k} b_k - \eta_D \Delta^4 b_k$. To predict the evolution of the velocity field, we use the analytical solution of equation 16:

$$u_{l,m}(t') = \exp\left(-\frac{t'-t}{\tau(l)}\right) u_{l,m}(t) + \int_t^{t'} \exp\left(-\frac{t'-s}{\tau(l)}\right) \sigma(l) d\omega(s) \quad \forall t' > t\,. \tag{26}$$

Since the velocity field is advecting the magnetic field, the former has to be known at each $t + i\Delta t^f$ epochs. Under the assumption that the flow at the CMB is in a stationary state, formulation (26) can be expressed for each member u_k of the ensemble of velocity field $\{u\}$ as the following:

$$u_k(t + (i+1)\Delta t^f) = \Gamma(\Delta t^f) u_k(t + i\Delta t^f) + \xi_k(\Delta t^f)\,, i \geq 0\,, \tag{27}$$

where the memory term Γ is a diagonal matrix with entries given by $\exp\left(-\Delta t^f / \tau(l)\right)$ and where ξ_k is a random realization of the Gaussian white noise ξ. The latter is characterized by a 0 mean and a covariance Σ_ξ deriving form the spatial covariance of the velocity field in its stationary state Σ_u^∞ such as the following:

$$\Sigma_\xi = \Sigma_u^\infty - \Gamma \Sigma_u^\infty \Gamma^T\,. \tag{28}$$

Note that at this stage, Σ_u^∞ is unknown. It will be parameterized, as shown in section 2.3.1, and its parameters will be directly estimated with the data, according to the methodology detailed in section 2.3.2.

2.3. Characterization of the Parameters of the Autoregressive Process

The numerical resolution of the EnKF algorithm requires a knowledge of Γ and ξ, the parameters of the discrete AR1 process given by equation (27). Instead of imposing these parameters, they are evaluated so that the resulting flow evolution explains at best the time series of the observed secular variation $\bar{\gamma}^o$. For this evaluation to be possible, a certain parametrization $\mathcal{M}$ of Γ and Σ_u^∞ (or Σ_ξ) is necessary. The assumptions taken to characterize $\mathcal{M}$ are detailed in the following section.

2.3.1. Parametrization of the Autoregressive Process

Following the developments of Baerenzung et al. (2016), Σ_u^∞ is derived from the poloidal and toroidal stationary spectra of the flow, the latter being assumed to behave as power laws with different spectral ranges. Γ, which contains the information on the scale-dependent temporal correlations of the velocity field, is derived from power laws with the same ranges as the flow energy spectra. Under such assumptions, the poloidal and toroidal stationary spectra of the flow and memory terms of the AR1 process are given by the following:

$$E_\phi(l) = C^i_{E_\phi} I^2_{E_\phi} l^{-P^i_{E_\phi}} \quad \text{for} \quad l \in \Delta^i_\phi \tag{29}$$

$$E_\psi(l) = C^j_{E_\psi} I^2_{E_\psi} l^{-P^j_{E_\psi}} \quad \text{for} \quad l \in \Delta^j_\psi \tag{30}$$

$$\Gamma_\phi(l) = C^i_{\Gamma_\phi} I^2_{\Gamma_\phi} l^{-P^i_{\Gamma_\phi}} \quad \text{for} \quad l \in \Delta^i_\phi \tag{31}$$

$$\Gamma_\psi(l) = C^j_{\Gamma_\psi} I^2_{\Gamma_\psi} l^{-P^j_{\Gamma_\psi}} \quad \text{for} \quad l \in \Delta^j_\psi, \tag{32}$$

where the I's are the magnitudes of the energy spectra and the memory terms, and the P^k's are their slopes within the SH ranges Δ^k's. The constants C^k's are given by the following:

$$C^k = \prod_{a=2}^{a=k} \exp\left(\log\left(l_{a\text{-}1}\right)\left(P^a - P^{a\text{-}1}\right)\right) \tag{33}$$

l_a being the SH degrees where transitions in slope occur. For a visual interpretation of relations (29) to (33), the reader can look at Figures 1 and 2.

According to this parametrization, whereas the matrix Γ is diagonal and contains both $\Gamma_\phi(l)$ and $\Gamma_\psi(l)$, the spatial covariance Σ_u^∞ derives from the poloidal and toroidal energy spectra such as the following:

$$\overline{\phi_{l,m}\phi_{l',m'}} = \frac{E_\phi(l)}{l(l+1)} \delta_{ll'} \delta_{mm'} \tag{34}$$

$$\overline{\psi_{l,m}\psi_{l',m'}} = \frac{E_\psi(l)}{l(l+1)} \delta_{ll'} \delta_{mm'} \tag{35}$$

$$\overline{\phi_{l,m}\psi_{l',m'}} = 0 \quad \forall l, l', m, m'\,. \tag{36}$$

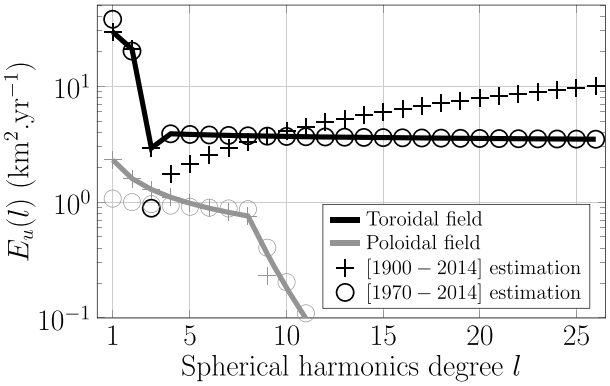

Figure 1. Prior kinetic energy spectra for the toroidal part of the velocity field (black) and for its poloidal part (gray). Estimations with the COV-OBS.x1 secular variation and magnetic field model between 1900.0 and 2014.0 (crosses) and between 1970.0 and 2014 (circles). The solid lines are the combination of the two evaluations used as a prior information to parametrize the autoregressive process for the flow.

In matrix form, Σ_u^∞ can be written as:

$$\Sigma_u^\infty = \begin{pmatrix} \frac{E_\phi(l_0)}{l_0(l_0+1)} & 0 & \cdots & 0 \\ 0 & \frac{E_\psi(l_0)}{l_0(l_0+1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{E_\psi(l_{\max})}{l_{\max}(l_{\max}+1)} \end{pmatrix}, \tag{37}$$

where l_0 and $l_{\max}$ are, respectively, the smallest and largest SH degrees of the flow spectral decomposition.

The set of parameters $\mathcal{M}$ associated with the autoregressive process can be divided into two categories, one containing the parametrization of the spatial covariances, $\mathcal{M}_\Sigma$, and the other associated with the memory terms of the process $\mathcal{M}_\Gamma$. $\mathcal{M}_\Sigma$ and $\mathcal{M}_\Gamma$ are, respectively, given by the following:

$$\mathcal{M}_\Sigma = \left\{ l_{E_\phi}, P^i_{E_\phi}, \Delta^i_\phi, l_{E_\psi}, P^j_{E_\psi}, \Delta^j_\psi \right\} \tag{38}$$

$$\mathcal{M}_\Gamma = \left\{ l_{\Gamma_\phi}, P^i_{\Gamma_\phi}, \Delta^i_\phi, l_{\Gamma_\psi}, P^j_{\Gamma_\psi}, \Delta^j_\psi \right\}, \tag{39}$$

where the index i and j are associated with the different poloidal and toroidal SH ranges.

2.3.2. Estimation of the Parameters Associated With the Autoregressive Process

To estimate the parameters $\mathcal{M}$, which enable the flow to optimally explain $\bar{\boldsymbol{\gamma}}^o$, the secular variation time series of the COV-OBS.x1 model, one can maximize the distribution $p(\mathcal{M}|\bar{\boldsymbol{\gamma}}^o)$. In order to obtain the latter distribution, the joint posterior distribution $p(\mathbf{u}, \mathbf{b}, \mathcal{M}|\bar{\boldsymbol{\gamma}}^o)$ can be marginalized in the following manner:

$$p(\mathcal{M}|\bar{\boldsymbol{\gamma}}^o) = \int p(\mathbf{u}, \mathbf{b}, \mathcal{M}|\bar{\boldsymbol{\gamma}}^o) \mathrm{d}\mathbf{u}\mathrm{d}\mathbf{b} = \frac{1}{p(\bar{\boldsymbol{\gamma}}^o)} \iint p(\boldsymbol{\gamma}^o|\mathbf{u}, \mathbf{b}, \mathcal{M}) p(\mathbf{b}) p(\mathbf{u}|\mathcal{M}) p(\mathcal{M}) \mathrm{d}\mathbf{u}\mathrm{d}\mathbf{b}\,. \tag{40}$$

On the second line of equation (40), $p(\bar{\boldsymbol{\gamma}}^o|\mathbf{u}, \mathbf{b}, \mathcal{M})$ is the likelihood distribution, $p(\mathbf{b})$, $p(\mathcal{M})$ and $p(\mathbf{u}|\mathcal{M})$ are the prior distributions for, respectively, the magnetic field, the parameters of the AR1 process, and the velocity field knowing $\mathcal{M}$, and finally, $p(\bar{\boldsymbol{\gamma}}^o)$ is the prior distribution of the secular variation data, which is a constant. The derivation of each distribution entering relation (40) is detailed in the following.

2.3.2.1. Likelihood Distribution $p(\bar{\boldsymbol{\gamma}}^o|\mathbf{u}, \mathbf{b}, \mathcal{M})$

The likelihood distribution is a measure of the statistical properties of the data once reality is known. In our case we assume that the real secular variation is at any time given by the FF equation such as $\boldsymbol{\gamma} = -A(\mathbf{bu})$. The data being the COV-OBS.×1 Gaussian model for the secular variation, the likelihood distribution is therefore given by the following:

$$p(\bar{\boldsymbol{\gamma}}^o|\mathbf{u}, \mathbf{b}, \mathcal{M}) = \mathcal{N}\left(-A(\mathbf{bu}), \Sigma^o_\gamma\right). \tag{41}$$

The possible correlations in time of the errors associated with the COV-OBS.x1 secular variation are neglected; therefore, Σ^o_γ is a block diagonal matrix with Σ^o_γ blocks (the covariance matrices of the COV-OBS.x1 secular variation at different epochs).

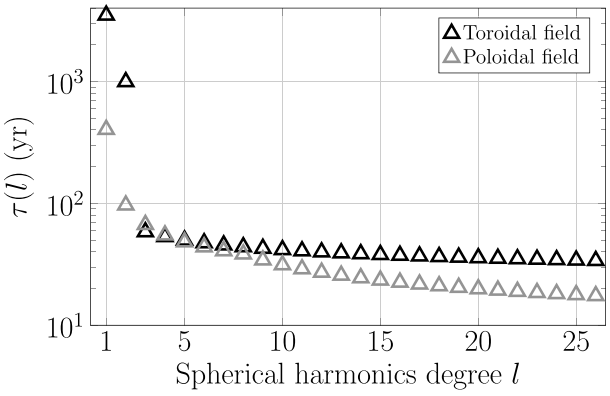

Figure 2. Prior characteristic timescale $\tau(l)$ for the autoregressive process of the flow, associated with the toroidal part of the velocity field (in black) and its poloidal part (in gray).

2.3.2.2. Prior Distribution of the Magnetic Field $p(\mathbf{b})$

This distribution is decomposed into two parts. The first one describes the statistical properties of the large-scale, observable field $\mathbf{b}^<$, whereas the second part expresses our prior knowledge on the small-scale field $\mathbf{b}^>$. The large-scale magnetic field is characterized by the prior distribution:

$$p(\mathbf{b}^<) = \mathcal{N}(\bar{b}^o, \Sigma^o_b)\,, \tag{42}$$

where $\bar{b}^o$ and Σ^o_b are, respectively, composed of the COV-OBS.x1 magnetic fields b^o and covariance matrices Σ^o at different epochs. As for the covariance of the COV-OBS.x1 secular variation, the correlations in time of the magnetic field model errors are neglected and Σ^o_b is a block diagonal matrix composed of Σ^o_b blocks.

The small-scale magnetic field is chosen to be at any time isotropically distributed, with a 0 mean and a covariance $\Sigma_{b^>}$ deriving from the extrapolation of the large-scale field spectrum $E_{b^<}(l)$. Here we use the formulation proposed by Buffett and Christensen (2007) to characterize the magnetic field spectrum at the CMB. It reads

$$E_{b^>}(l) = C_1 \chi^l \tag{43}$$

where $\chi = 0.99$. To determine the constant C_1, we used the COV-OBS.x1 magnetic field sampled every 2 years between 1900.0 and 2014.0 and performed a weighted least squares fit of the associated energy spectra between SH degree $l = 2$ and $l = 13$. We obtained that $C_1 = 7.15 \times 10^9$ nT^2. Another type of extrapolation has also been tried, assuming an exponential decay of the magnetic field spectrum. Although we do not show the results associated with this modeling, we observed that such an assumption would provide insufficient levels of energy at small scales, leading to suboptimal predictions of the magnetic field evolution. From the extrapolation given in equation (43) we construct the covariance of the small-scale magnetic field $\Sigma_{b^>}$ at a given time through the relation:

$$\overline{b^>_{l,m} b^>_{l',m'}} = \frac{E_{b^>}(l)}{(l+1)(2l+1)} \delta_{ll'} \delta_{mm'} . \tag{44}$$

Neglecting a priori the temporal correlations between the small scales of the magnetic field, the full covariance $\Sigma_{b^>}$ is simply a block diagonal matrix where every block is identical and given by $\Sigma_{b^>}$. The prior distribution associated with the small-scale magnetic field is thus given by the following:

$$p(\mathbf{b}^>) = \mathcal{N}(\mathbf{0}, \Sigma_{b^>}) , \tag{45}$$

and we have $p(\mathbf{b}) = p(\mathbf{b}^<) p(\mathbf{b}^>)$.

2.3.2.3. Prior Distribution of the AR1 Parameters $p(\mathcal{M})$

The parameters $\mathcal{M}$ depend on the magnitudes I's, the slopes P's, and the SH ranges Δ's of the flow stationary spectra and the AR1 memory terms. Whereas the ranges Δ's will be a priori imposed and therefore considered as known, the magnitudes and slopes are completely undetermined. To reflect this lack of knowledge, we characterize them by uniform distributions such as the following:

$$p(I) = \mathcal{U}(0, \infty) \tag{46}$$

$$p(P) = \mathcal{U}(-\infty, \infty) . \tag{47}$$

The full prior distribution of $\mathcal{M}$ is simply the product of the prior distributions of each individual AR1 parameter.

2.3.2.4. Prior Distribution of the Velocity Field Conditioned by the AR1 Parameters $p(\mathbf{u}|\mathcal{M})$

The dynamical evolution of the velocity field is prescribed the first-order autoregressive process described in section 2.2. In its discrete form, we recall that the process can be written as follows:

$$u(t + \Delta t) = \Gamma(\Delta t, \mathcal{M}) u(t) + \xi(\Delta t, \mathcal{M}) . \tag{48}$$

If the parameters $\mathcal{M}$ describing Γ and ξ are known, and if we assume that the velocity field at the CMB is in its stationary regime, the prior distribution of the velocity field on a given time window can be expressed as follows:

$$p(\mathbf{u}|\mathcal{M}) = \mathcal{N}(0, \Sigma_{\mathbf{u}|_{\mathcal{M}}}) \tag{49}$$

where the covariance matrix $\Sigma_{\mathbf{u}|_{\mathcal{M}}}$ is given by:

$$\Sigma_{u|\mathcal{M}} = \begin{pmatrix} \Sigma_{u|\mathcal{M}}^{\infty} & \Gamma_{|\mathcal{M}} \Sigma_{u|\mathcal{M}}^{\infty} & \cdots & \Gamma_{|\mathcal{M}}^{(N-1)} \Sigma_{u|\mathcal{M}}^{\infty} \\ \Sigma_{u|\mathcal{M}}^{\infty} \Gamma_{|\mathcal{M}}^{T} & \Sigma_{u|\mathcal{M}}^{\infty} & \cdots & \Gamma_{|\mathcal{M}}^{(N-2)} \Sigma_{u|\mathcal{M}}^{\infty} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{u|\mathcal{M}}^{\infty} \Gamma_{|\mathcal{M}}^{(N-1)^T} & \Sigma_{u|\mathcal{M}}^{\infty} \Gamma_{|\mathcal{M}}^{(N-2)^T} & \cdots & \Sigma_{u|\mathcal{M}^{\infty}} \end{pmatrix} . \tag{50}$$

2.3.2.5. Mariginalization

Because of the nonlinear term entering the likelihood distribution in equation (40), the marginalization of $p(\mathbf{u}, \mathbf{b}, \mathcal{M}|\bar{\gamma}^o)$ with respect to **B** and **u** results in a complex distribution, which is hard to handle numerically.

This is why, following the development of Baerenzung et al. (2016), the posterior distribution of the AR1 parameters given the secular variation $p(\mathcal{M}|\bar{\boldsymbol{\gamma}}^o)$ is approximated by the following distribution:

$$p(\mathcal{M}|\bar{\boldsymbol{\gamma}}^o) = \frac{\exp\left[-\frac{1}{2}\bar{\boldsymbol{\gamma}}^{\mathbf{o}T}\Sigma^{-1}_{\mathcal{M}_{|\bar{\gamma}^o}}\bar{\boldsymbol{\gamma}}^{\mathbf{o}}\right]}{(2\pi)^{\frac{d_\gamma}{2}}|\boldsymbol{\Sigma}_{\mathcal{M}_{|\bar{\gamma}^{\mathbf{o}}}}|^{\frac{1}{2}}} \times \frac{p(\mathcal{M})}{p(\bar{\boldsymbol{\gamma}}^o)}, \tag{51}$$

where d_γ is the dimension of the secular variation vector. To construct the matrix $\boldsymbol{\Sigma}_{\mathcal{M}_{|\bar{\gamma}^o}}$, the covariance between $\bar{\gamma}^o$ at a time t_α and $\bar{\gamma}^o$ at a time t_β, with respect to the distribution $p(\bar{\boldsymbol{\gamma}}^o|\mathbf{u},\mathbf{b},\mathcal{M})p(\mathbf{b})p(\mathbf{u}|\mathcal{M})$ is calculated for every combination of epochs considered. The component at a row index i and a column index j of the resulting covariance matrix $\Sigma^{t_\alpha t_\beta}_{\mathcal{M}_{|\bar{\gamma}^o}}$ reads

$$\begin{aligned}\left(\Sigma^{t_\alpha t_\beta}_{\mathcal{M}_{|\bar{\gamma}^o}}\right)_{ij} = &\left(\Sigma^{o^{t_\alpha t_\beta}}_{\gamma}\right)_{ij} + \left(A_{\bar{b}_o}(t_\alpha)\Sigma^{t_\alpha t_\beta}_{u_{|\mathcal{M}}}A^T_{\bar{b}_o}(t_\beta)\right)_{ij} \\ &+ A_{imn}\left(\Sigma^{t_\alpha t_\beta}_{u_{|\mathcal{M}}}\right)_{mr}\left(\Sigma^{o^{t_\alpha t_\beta}}_{b}\right)_{ns}A_{jrs}\,,\end{aligned} \tag{52}$$

where we recall that the third-order tensor A is defined such as $A_{ijk}(u)_j(b)_k = \left(A_b u\right)_i = \left(A_u b\right)_i$, and where the Einstein summation convention applies to the tensor indexes m, n, r, and s.

3. Geophysical Application

3.1. Numerical Setup

The data entering our EnKF assimilation scheme consist of magnetic field and secular variation SH coefficients from the COV-OBS.x1 model of Gillet, Barrois, and Finlay (2015) spanning the time period between 1900 and 2014. The coefficients are taken within a 2-year sampling, corresponding to the knots of COV-OBS.x1 B-spline temporal expansion. The magnetic and velocity fields simulations are performed through a pseudospectral approach on the Gaussian-Legendre grid provided by Schaeffer (2013). Both the poloidal and toroidal parts of the velocity field are expanded up to SH degree $l = 26$, and the radial component of the magnetic field is expressed up to SH degree $l = 39$ in order for the field to possess a large enough diffusion range. Whereas the COV-OBS.×1 magnetic field is always taken up to SH degree $l = 13$, the expansion of the COV-OBS.x1 secular variation depends on the variance level associated with each degree. If globally at a certain scale the standard deviation of the secular variation is larger than the absolute value of the mean secular variation, the total field is truncated at this scale. Under such a condition, the COV-OBS.x1 secular variation is taken up to SH degrees $l = 10$, $l = 11$, $l = 12$, and $l = 13$ for the respective time windows [1900–1923], [1924–1943], [1944–1963], and [1964–2014]. Finally, the state of the system is characterized by 40,000 pairs of magnetic field and velocity field at the CMB.

3.2. Estimation of the Flow Optimal Autoregressive Parameters

To simulate the spatiotemporal evolution of the flow at the CMB, the parameters of the autoregressive process have to be estimated. We recall that their posterior distribution, $p(\mathcal{M}|\bar{\boldsymbol{\gamma}}^o)$, is expressed in equation (51). By maximizing this distribution, it is possible to get the optimal parameters for the AR1 process. However, instead of estimating both temporal and spatial parameters simultaneously, we proceed in two steps. First, only the spatial covariance of the velocity field is evaluated following the method proposed and tested by Baerenzung et al. (2016). This approach consists in maximizing the distribution $p(\mathcal{M}|\bar{\boldsymbol{\gamma}}^o)$ in which only the block diagonal part of the covariance matrix $\Sigma_{\mathcal{M}_{|\bar{\gamma}^o}}$ is kept. Once the spatial covariance is determined, it is assumed to be known, and the maximum of $p(\mathcal{M}_\Gamma|\bar{\boldsymbol{\gamma}}^o,\mathcal{M}_\Sigma)$ is calculated.

Following the developments of section 2.3.1, the AR parameters are decomposed into scale-dependent power laws exhibiting different spectral ranges. Baerenzung et al. (2016) showed that if the stationary spectra of the flow are decomposed into two spectral ranges, the optimal scales where transition in slope occurs are $l = 3$ and $l = 8$ for, respectively, the toroidal and poloidal energy spectra. Here whereas we keep the same decomposition for the spectrum associated with the poloidal field, more degrees of freedom are allowed for the toroidal field spectrum. Since we wish to accurately determine the spatiotemporal evolution of the eccentric gyre, toroidal field component at SH degree $l = 1$ and $l = 2$, the main components of the gyre, are free to exhibit any variance level and characteristic time. Similarly, toroidal field components at SH degree $l = 3$ are also assumed to be unconstrained by surrounding velocity field scales. This choice is motivated

Table 1
Combined Optimal Covariance Parameters $\mathcal{M}_\Sigma$ of the Flow Autoregressive Process, Within the 1970.0–2014.0 and the 1900.0–2014.0 Periods

Flow field	Range index i	Δ_i	l_i	P_i
Toroidal	1	1	5.41	0
	2	2	4.56	0
	3	3	1.71	0
	4	[4 , 26]	2.05	5.8×10^{-2}
Poloidal	1	[1 , 8]	1.52	0.54
	2	[8 , 26]	663	6.4

Note. l_i corresponds to the magnitudes and P_i to the slopes of the prior stationary spectra, within the spectral ranges Δ_i (see equations (29) and (30), and (34) and (35)).

by the particular low level of energy that these scales are exhibiting over recent epochs (see Baerenzung et al., 2016; Whaler et al., 2016). Finally, one spectral range is used to characterize the toroidal field spatial variance and memory effects between SH degrees $l = 4$ and $l = 26$. All in all, the AR1 parameters associated with the toroidal field exhibit the four respective spectral ranges, $\Delta_0 = [1]$, $\Delta_1 = [2]$, $\Delta_2 = [3]$, and $\Delta_3 = [4, 26]$.

As mentioned in the beginning of the section, the estimation of the stationary energy spectra parameters is performed between 1900.0 and 2014.0, taking the COV-OBS.x1 magnetic field and secular variation every $\Delta t^A = 2$ years. In Figure 1 the resulting power law spectra are displayed with crosses. As already observed in Baerenzung et al. (2016), the toroidal field (in black), and in particular its large scales (SH degree $l = 1$ and $l = 2$), exhibits a much larger energetic level than the poloidal field (in gray). The toroidal energy spectrum also presents a strong increase of energy toward its smallest scales. This contradiction with the results of Baerenzung et al. (2016) is likely attributed to a slight underestimation of the COV-OBS.x1 secular variation uncertainties. Indeed, secular variation components that are not consistent with the FF approximation can only be explained by an artificial injection of small-scale velocity field. Therefore, in order to better estimate the small-scale energy spectra of the velocity field, we performed different estimations of the stationary spectra parameters by varying the time window in which the evaluation is computed. We found that the largest period where the spectra did not exhibit an anomalous behavior was 1970.0–2014.0. The resulting prior spectra are shown in Figure 1 with circles. Combining the small-scale spectra of the 1970.0–2014.0 evaluation to the large-scale ones of the 1900.0–2014.0 estimation, we get the final prior spectra for both the toroidal and poloidal field displayed in Figure 1 with solid lines. The values associated with the spectra parameters are given in Table 1.

The spatial covariances of the AR1 process being characterized, the evaluation of the memory terms is now performed by maximizing the distribution $p(\mathcal{M}_\Gamma|\bar{\gamma}^o, \mathcal{M}_\Sigma)$ within the 1900.0–2014.0 time window. The results, expressed through the scale-dependent characteristic time $\tau(l) = -\frac{\Delta_t}{\log(\Gamma(l))}$, are shown in Figure 2, and the parameters of the memory terms are given in Table 2. The most striking feature we observe is the very long memory time of the order of thousand years, of degrees $l = 1$ and $l = 2$, associated with the main components of the eccentric gyre. This indicates that this structure has to be very persistent over time in order to optimally explain the observed secular variation. These values should nevertheless be taken with care since they are evaluated on a comparatively short time window of 114 years. In contrast with the large-scale field, the toroidal field components at SH degree varying from $l = 3$ to $l = 26$ exhibit much lower characteristic memory, with decaying times of $\tau(l = 3) \sim 50$ years and $\tau(l = 26) \sim 30$ years. Note that this limiting time is similar to the *e*-folding time of the geodynamo as calculated by Hulot et al. (2010) and Lhuillier et al. (2011). The characteristic times associated with the poloidal field indicate a similar behavior but with $\tau(l)$ varying from $\tau(l = 1) \sim 400$ years to $\tau(l = 8) \sim 40$ years.

Table 2
Optimal Parameters for the Memory Term of the Autoregressive Process $\mathcal{M}_\Gamma$, Evaluated Within the 1900–2014 Time Window

Flow field	Range index i	Δ_i	$1 - l_i$	P_i	$\tau(\Delta_i)$
Toroidal	1	1	1.43×10^{-4}	0	3495
	2	2	5.03×10^{-4}	0	994
	3	3	8.53×10^{-3}	0	58
	4	[4 , 26]	5.42×10^{-3}	5.72×10^{-3}	[53 , 34]
Poloidal	1	[1 , 8]	1.24×10^{-3}	1.13×10^{-3}	[403 , 38]
	2	[8 , 26]	-1.47×10^{-2}	2.65×10^{-2}	[38 , 17]

Note. l_i and P_i are, respectively, the magnitudes and the slopes of the assumed power laws within the spectral ranges Δ_i. Also given are the characteristic times of the AR1 process $\tau = -\frac{1}{\log(\Gamma)}$ expressed in years.

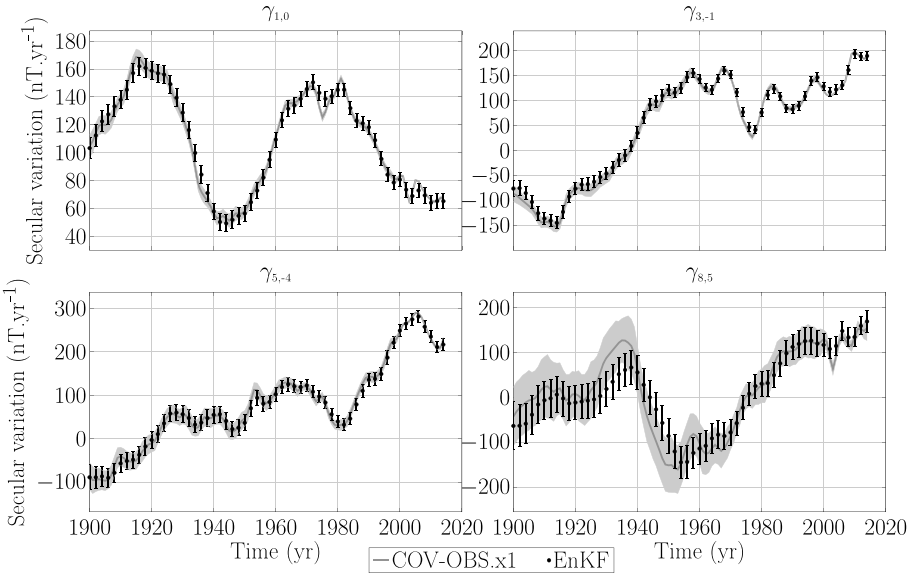

Figure 3. Time series of some selected secular variation coefficients $\gamma_{l,m}$ at the CMB. Black dots and gray lines respectively correspond to $\bar{\gamma}^A_{l,m}$ (EnKF) and $\bar{\gamma}^o_{l,m}$ (COV-OBS.x1). Error bars and gray shaded areas provide the associated standard deviations. EnKF = ensemble Kalman filter.

3.3. Implementation of the EnKF and Misfit to the Data

With the AR1 process parametrization in place, the estimation of the velocity and magnetic field at the CMB are ready for the EnKF algorithm. To initialize the fields in 1900.0, we applied the Gibbs sampling algorithm proposed by Baerenzung et al. (2016) and which is detailed in Appendix A. However, instead of sampling the joint posterior distribution of the flow and the magnetic field at the 1900.0 epoch only, we sampled the distribution characterizing simultaneously the fields in 1900.0, 1950.0, and 2000.0, in order to constrain the initial state with recent observations.

Once the initial fields were obtained, the EnKF algorithm was numerically solved, alternating analysis every $\Delta t^A = 2$ years and forecasts consisting of four successive predictions with a time step of $\Delta t^f = 0.5$ year. The consistency of the model with the data is evaluated through the prediction misfit proposed by Evensen (2003). For respectively the magnetic field and the secular variation, this quantity reads

$$\chi^2_b = \left\langle \frac{1}{d_b} (\bar{b}^o - H\bar{b}^f)^T \left(\Sigma^o_b + H\Sigma^f_b H^T\right)^{-1} (\bar{b}^o - H\bar{b}^f) \right\rangle_{1900.0}^{2014.0} \tag{53}$$

$$\chi^2_\gamma = \left\langle \frac{1}{d_\gamma} (\bar{\gamma}^o - \bar{\gamma}^f)^T \left(\Sigma^o_\gamma + \Sigma^f_\gamma\right)^{-1} (\bar{\gamma}^o - \bar{\gamma}^f) \right\rangle_{1900.0}^{2014.0}, \tag{54}$$

where the notation $< \ldots >_{T_0}^{T_F}$ is associated with the time averaging between T_0 and T_F of the quantity lying within the brackets and where d_b and d_γ are the dimensions of, respectively, the observed magnetic field and secular variation. We recall that γ^f and b^f correspond to the predictions of the secular variation and the magnetic field at every analysis epochs. The value of $\chi^2_\gamma = 1.08$ indicates that globally, the predictions for the secular variation are consistent with its observations. For the magnetic field, the mean prediction misfit of $\chi^2_b = 0.58$ suggests a slight overfit of the data. This effect is certainly due to the fact that γ^o directly derives from the time series of b^o and that therefore their respective uncertainties are correlated. Yet these correlations are neglected in our model.

The ability of our model to reproduce the time evolution of the observed secular variation is illustrated in Figure 3. Indeed, the time series of some selected secular variation coefficients $\bar{\gamma}^A_{l,m}$ and associated standard deviation (black dots with erros bars) compare well with their observed counter parts $\bar{\gamma}^o_{l,m}$, which are shown with gray lines and shaded areas for the corresponding standard deviations.

3.4. General Properties of the Flow at the CMB

Because of the sequential nature of the EnKF algorithm, the accuracy of the estimated fields is not constant over time but improves whenever new data are assimilated. This effect is well illustrated in Figure 4 displaying

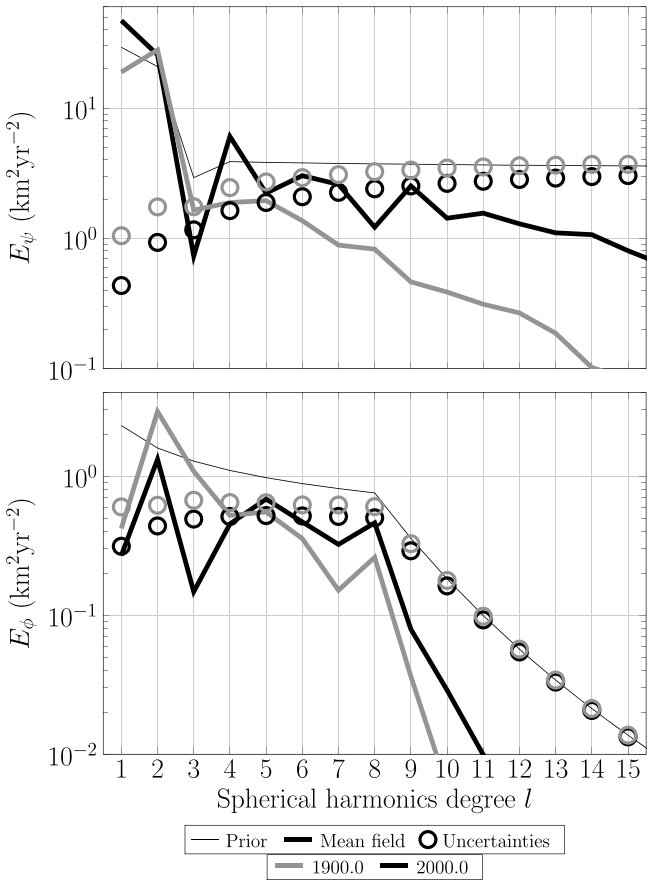

Figure 4. Toroidal (top) and poloidal (bottom) energy spectra associated with the ensemble mean fields (thick lines) and standard deviation (circles), for the 1900.0 (gray) and 2014.0 (black) epochs. Thin lines correspond to the prior spectra.

at two different epochs, 1900.0 (gray) and 2000.0 (black), the energy spectra of the toroidal (top) and poloidal (bottom) mean velocity fields (thick solid lines) and uncertainty fields (circles). Whereas in 1900.0 the mean toroidal field exhibits a level of energy larger or of the same order than the variance of the field up to SH degree $l = 3$, in 2000.0 reliable information becomes available up to SH degree $l = 9$, although some components of the field at SH degree $l = 3$ and $l = 8$ remain uncertain. Above these scales the posterior variance of the flow rapidly reaches its prior level. One can also notice that the larger the scale, the stronger the variance reduction over time. The latter observation, which is also valid for the poloidal field, is linked to the characteristic times $\tau(l)$ associated with the different flow scales (see Figure 2). Indeed, the fact that $\tau(l)$ is a strictly decaying function of l, implies that small-scale velocity field will exhibit a higher randomization rate than large scales, and so between two analysis step, the prior variance will increase faster at small scales.

In physical space, the gain of flow accuracy over time is particularly striking for the toroidal part of the velocity field as shown on Figure 5. In this figure, the toroidal (left) and poloidal (right) mean velocity fields are displayed with black arrows for three different epochs, 1900.0 (top), 1950.0 (middle), and 2000.0 (bottom). Color maps, representing the 90% confidence interval on the velocity field orientation, provide information on locations where the mean flow direction can be reliably estimated (violet and blue) or not (red). In 1900.0, very little of the eccentric gyre can be confidently estimated. Only the westward flow below Africa and the Atlantic Ocean, the southern branches of the gyre, and the northern circulation around and partially inside the tangent cylinder (the cylinder tangent to the inner core and aligned with the axis of rotation of the Earth), appear as reliable patterns. At later times the gyre is better defined, and many of its small-scale structures become visible. Globally, uncertainties on the toroidal part of the flow are decreasing with time. This does not seem to be the case for the poloidal field, for which in 1950.0 reliable patterns are covering a larger surface of the CMB than in 2000.0. Nevertheless, the root-mean-square (r.m.s.) velocity of the poloidal field and associated standard deviation of 3.49 $\pm$ 0.23 in 1950.0 and 2.49 $\pm$ 0.21 in 2000.0 indicate that the global uncertainty level of the poloidal field and its magnitude have been decreasing between 1950.0 and 2000.0.

Contrary to flow models constrained to be geostrophic (see Amit & Paris, 2013; Bloxham & Jackson, 1991), upwelling and downwelling fluid motions, associated with the poloidal field close to the CMB, are not particularly located around the equator. Instead, a strong and persistent poloidal structure evolves below the Indian Ocean and South Africa. According to Bloxham (1986), such a poloidal field could be at the source, through the expulsion of magnetic flux from the outer core, of the intense reversed flux patch located there. Although the FF equation cannot model the transport of magnetic structures between the core and its outer boundary, the poloidal flow estimation hints at upwelling and downwelling connected to subsurface flow, which are usually connected to spread or concentration of magnetic flux, respectively.

Other specific features of the velocity field are in apparent contradiction with a possible geostrophic state of the outer core flow. This includes the reliable part of the toroidal field penetrating the tangent cylinder, or its component crossing the equator below India and South America, as already reported in other studies (see Barrois et al., 2017). Clearly visible is the violation by the eccentric gyre of the equatorial symmetry condition imposed by quasi-geostrophy (see Amit & Olson, 2004). Indeed, the flow responsible for the westward drift, together with the circulations around the tangent cylinder, exhibits different levels of intensity in the Northern Hemisphere than in the Southern Hemisphere.

These visual observations are confirmed in Figure 6, where the r.m.s. velocity and associated standard deviation of the toroidal part of the flow, are measured in different locations of the CMB. One can observe on the top right of Figure 6 that the flow evolving below Africa and the Atlantic Ocean within the areas shown

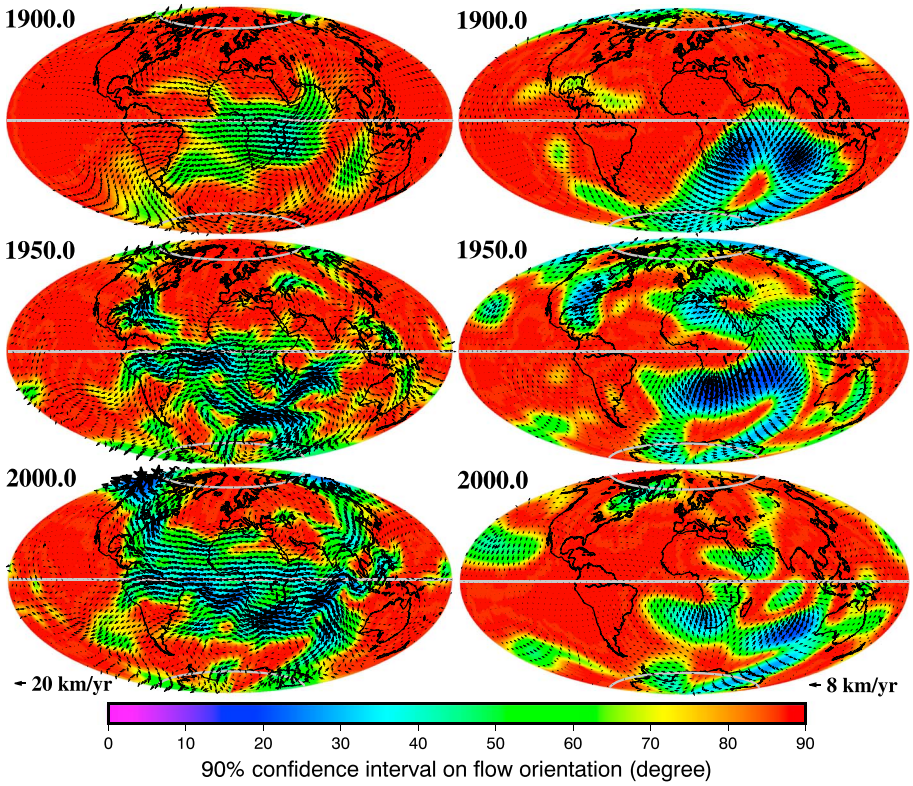

Figure 5. Toroidal (left column) and poloidal (right column) velocity fields in 1900.0 (top), 1950.0 (middle), and 2000.0 (bottom). Arrows correspond the mean fields (over the ensemble), and color maps are displaying the 90% confidence interval on their orientation. Note that the scaling for the velocity field is different between the toroidal and poloidal part of the flow.

on bottom right of the figure is at any time more intense in the south than in the north. For the circulations around the tangent cylinder, both southern and northern part exhibit similar levels of energy between 1900.0 and 1960.0 as shown on the bottom left of Figure 6. However, in 1960.0 the flow below Alaska and the eastern part of Siberia starts to accelerate and intensifies almost continuously over the last decades to reach a r.m.s. velocity around 23 km/year in 2014.0. Although this acceleration has already been observed during the satellite era by Livermore et al. (2017), here we notice that it has been persistent for many decades. We can further note that although the toroidal part of the flow exhibits some clear deviation from geostrophy, comparisons of its r.m.s. velocity over the entire CMB (black symbols on the top left of Figure 6) with the r.m.s. velocity of its equatorial symmetric part (gray symbols) show that the latter component remains at any time dominant.

3.5. Predictions

The ability of a model to successfully predict the system evolution not only suggests that the model correctly captures the dynamics on the considered timescales but also points toward useful applications. We test our model with so-called hindcast simulations, which means that the analysis steps in the assimilation are only carried out until a time T_0 and then integrated as a free model run up to a time T_F. The free run corresponds to the forecast, or prediction, which can be compared with the data. Here we use six different T_0 values, 1940.0, 1960.0, 1980.0, 1990.0, 2000.0, and 2010.0, and compare predictions for $T_F = 2015$ with the respective epoch in the CHAOS-6 magnetic field model by Finlay et al. (2016). This means that we attempt predictions over periods ranging from 5 to 55 years. As a measure to quantify the quality of our mean predictions, we calculated the r.m.s. difference $\sqrt{dP}$, between the mean forecasted magnetic field $\bar{b}^f$ and the CHAOS-6 field b^c in 2015.0. The latter reads

$$\sqrt{dP} = \left(\sum_{l=1}^{l=13} (l+1) \sum_{m=-l}^{m=l} \left(\bar{b}^f_{l,m} - b^c_{l,m} \right)^2 \right)^{0.5} . \tag{55}$$

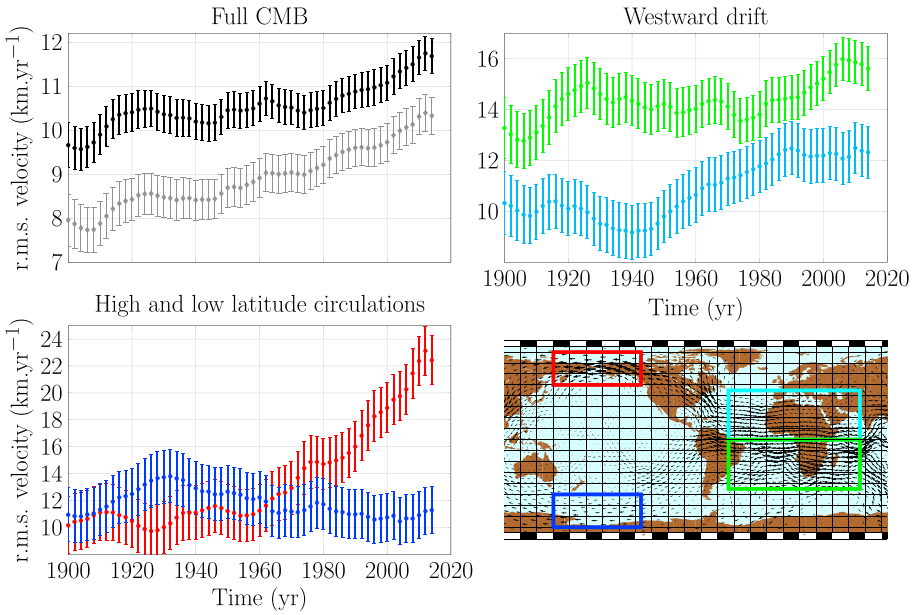

Figure 6. Root-mean-square (r.m.s.) velocity of the toroidal part of the flow and associated standard deviation in different locations of the core-mantle boundary (CMB). Colored curves correspond to the r.m.s. velocity within the areas surrounded by the same colored contour on the bottom right of the figure. The black and gray symbols on the top left corresponds to the r.m.s. velocity of, respectively, the toroidal flow and its equatorial symmetric part over the entire surface of the CMB. The arrows on the bottom right are associated with the 1980.0–2014.0 time-averaged mean toroidal field.

The values of $\sqrt{dP}$, which are evaluated at the Earth's surface, are summarized in Table 3. As it can be expected, the longer the forecast, the larger the discrepancies between predicted and observed magnetic field. For short-term predictions our results are consistent with previous studies. In particular the 5-year hindcast between 2010.0 and 2015.0, which leads to a r.m.s. difference of $\sqrt{dP} = 66$ nT, is equivalent to the 63-nT optimal value obtained by Whaler and Beggan (2015) for a hindcast between 2010.0 and 2014.5 (for the same hindcast period we get $\sqrt{dP} = 55$ nT). Looking at longer term forecasts, one can observe that the accuracy of the mean predicted field degrades rapidly with time. The AR1 process characterizing the evolution of the flow is certainly responsible for this. Indeed, with such a process, the different components of the mean velocity field are simply decaying over time. Therefore, over periods that do not exceed the characteristic times of the dominant flow scales, our mean predicted magnetic field would not differ much from a field, which would be advected by a static flow (SF). Nevertheless, the AR1 offers an important advantage compared to the SF assumption: the possibility to simulate the randomization over time of the different velocity field components. This additional feature indeed allows us to better estimate the growth rate of errors associated with mean magnetic field forecasts and thus provide accurate measures of the reliability of our predictions.

Table 3
Root-Mean-Square Difference at the Earth's Surface (in nT), Between the CHAOS-6 Magnetic Field in 2015.0 and the Mean Magnetic Field Predicted From T_0 to 2015.0

T_0	$\sqrt{dP}$
1940	2,875
1960	2,007
1980	1,220
1990	712
2000	328
2010	66

The last statement is well illustrated in Figure 7, which presents the hindcast tests in terms of energy spectra at the Earth's CMB. These results are compared with hindcasts where the flows are assumed to be static from T_0 (SF model). In the figure, thick black lines and thin gray lines are the spectra of, respectively, the CHAOS-6 reference field b^c and the predicted ensemble means $\bar{b}^f$ in 2015.0. The spectra of the difference between predicted and observed fields, which are referred to prediction errors, are displayed with gray and black triangles for, respectively, the EnKF and SF forecasts. The latter can be compared with the spectra of the predicted errors (the spectra of the standard deviation of the magnetic predicted field) displayed with thick gray lines (EnKF) and thin black lines (SF). One can first observe that prediction errors are of the same order of magnitude for the EnKF

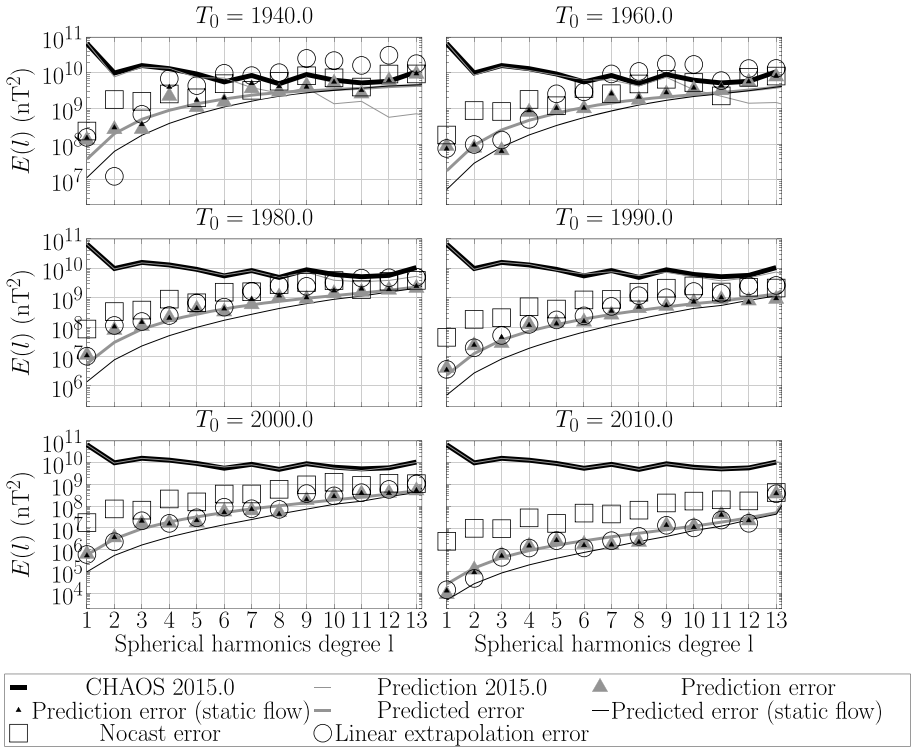

Figure 7. Results of the hindcast tests from T_0 to $T_F = 2015.0$, expressed in terms of energy spectra at the core-mantle boundary. Spectra of the observed magnetic field in 2015.0 (thick black lines), its mean prediction (thin gray lines), the prediction error (gray triangles), and the predicted error (thick gray lines). The thin black lines and the black triangles are associated with, respectively, the predicted and the prediction errors when the velocity field is assumed to be static from T_0. The circles and squares, respectively, correspond to the energy spectra of the linear extrapolation error and the no cast error.

and for the SF approaches. However, whereas predicted errors are consistent with prediction errors for the EnKF model, uncertainties derived from the SF model clearly underestimate the prediction errors. This shows that modeling flow dispersion with time is mandatory to explain the loss of information in magnetic field predictions.

The comparison with two other more trivial prediction methods further allows to judge the advantage of our more sophisticated approach. "No cast" refers to the assumption that the field remains identical to the field at T_0. "Linear extrapolation" uses the secular variation at T_0 to linearly extrapolate the field from T_0 to 2015: $b_l(2015) = b^o(T_0) + (2015 - T_0)\gamma^o$. The prediction errors of these two trivial methods are shown as circles and squares in Figure 7. Linear extrapolation and assimilation prediction errors remain similar for the two shortest prediction periods. However, for predictions beyond the 10-year horizon, our assimilation formalism starts to pay off. For the longest prediction period of 55 years, the errors in the two trivial methods already exceed the spectral energy at degree 6, while the assimilation predictions remains appropriate until degree 9.

Working with an ensemble also permits the evaluation of statistical properties of quantities like inclination or declination that depend nonlinearly on the state variables. Figure 8 compares the inclination (left) and declination (right) for 1990 (black lines) and 2015 (red lines) with the ensemble mean predictions using $T_0 = 1990$ (yellow lines). The prediction errors are quantified by the absolute local difference in degrees and are shown as color maps in the top two panels. Color maps in the bottom two panels show the 90% confidence interval of the ensemble prediction.

As expected, inclination errors are large for the small or vanishing values around the equator. This is illustrated in the top panel of the left column and also well captured by the larger variance in the ensemble used to predict the error shown in the lower left panel. We calculated the area where the prediction error remains within the 90% confidence interval defined by the prediction ensemble, which amounts to 89.9% of the total surface. This indicates that the inclination uncertainties provide a good estimate of the prediction error.

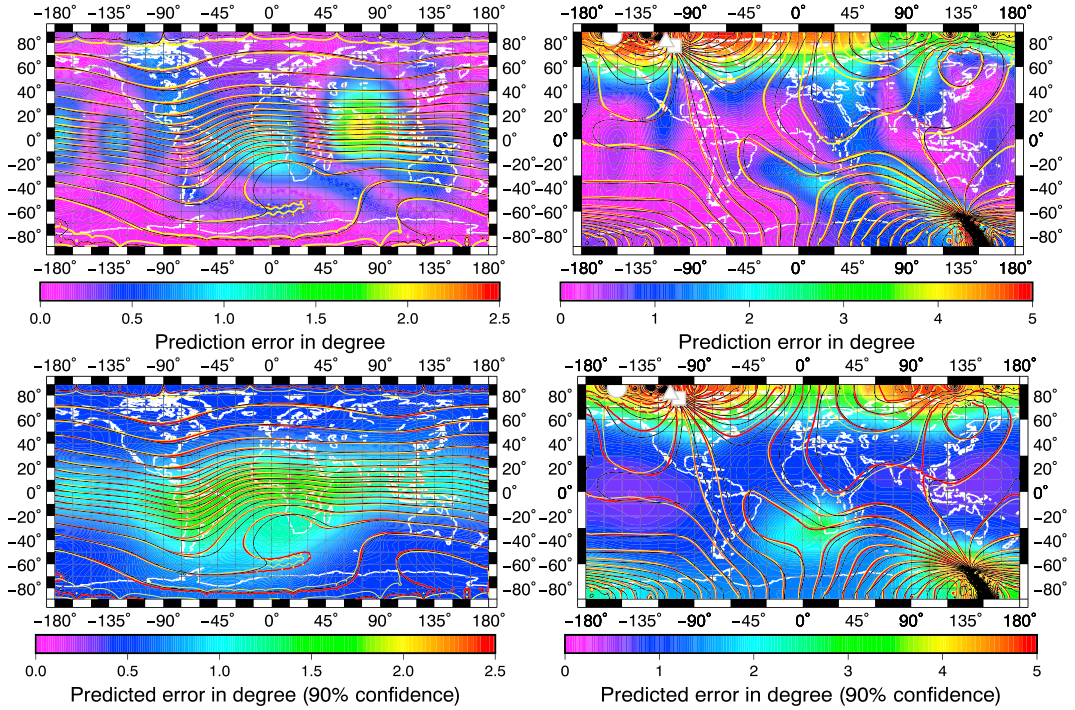

Figure 8. The 1990.0–2015.0 hindcast. Isocontours of the observed inclination (left column) and declination (right column), in 1990.0 (black) and 2015.0 (yellow), and their predictions in 2015.0 (red). Color maps correspond to the absolute value of prediction error (top) and to the 90% confidence interval on the prediction (bottom). The white symbols on the right are associated with the north magnetic dip pole, observed in 1990.0 (square) and in 2015.0 (circle) and predicted in 2015.0 (triangle).

Since the declination is undefined at the north and south dip poles, the surrounding regions show larger prediction errors and ensemble variances, as is demonstrated in the right column of Figure 8. The area where prediction errors remain within the 90% confidence interval amounts to only 81.1% of the surface, which suggests a somewhat inferior error prediction likely related to the dip poles. The error prediction seems more reliable in the Southern Hemisphere alone where the relative area increases to 92.2%.

Although not clearly visible in Figure 8 because the color has been saturated at 5°, the prediction error for the location of the north magnetic dip pole (NMDP) is also particularly large. Its real and predicted position for 2015 has been marked by a gray circle and triangle, respectively. According to Chulliat et al. (2010), the rapid acceleration of the NMDP drift during the 1990s can be explained by the expulsion of magnetic flux below the New Siberian Islands. This may be the reason for the larger prediction error, since such expulsions are not modeled in the FF approximation used here. The NMDP drift is much better captured, however, when the hindcast test starts at $T_0 \geq 2000.0$.

Finally, we note that our model accurately predicts the evolution of the inclination and declination associated with South Atlantic Anomaly despite the significant changes between 1990.0 and 2015.0. This highlights the potential usefulness of the method for forecasting core's magnetic field features.

3.6. Predictability

In order to quantify the different sources of forecast errors, we analyze the secular variation, which is responsible for advancing the field from T_0 to T_F: $b(T_F) = b(T_0) + \int_{T_0}^{T_F} \gamma(s)\mathrm{d}s$. Under the FF approximation, γ depends on the flow and the magnetic field, through the relation $\gamma = -\nabla_H(ub)$. When separating b into the observable part $b^<$ and the nonobservable small-scale part $b^>$, we can distinguish three error sources:

$$\gamma' = -\nabla_H(u'b) - \nabla_H(ub^{<\prime}) - \nabla_H(ub^{>\prime}). \tag{56}$$

Here the primed quantities on the right-hand side indicate deviations from the ensemble expectation value, for example, $u' = u - \bar{u}$.

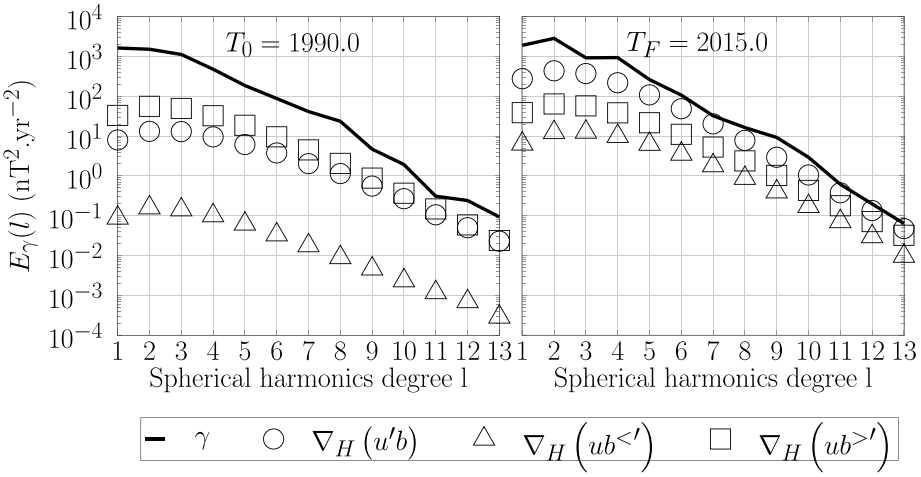

Figure 9. Sources of secular variation (secular variation) error in magnetic field predictions, at the initial time $T_0 = 1990.0$ (left) and at the forecast time $T_F = 2015.0$ (right). Energy spectra at the Earth's surface of the secular variation (black lines) and of different types of error (symbols) generated when estimating the secular variation. Triangles and circles correspond to the errors induced by, respectively, the large-scale and small-scale variable parts of the magnetic field. Squares are associated with the uncertainties arising from the variable part of the velocity field.

Figure 9 compares the spectra of the different error contributions to the secular variation spectrum at the Earth's surface for $T_0 = 1990$ (left) and $T_F = 2015$ (right). The last analysis step performed at T_0 directly constrains $b^<$ and leads to a small related variance and thus a small error contribution that lies about 2 orders of magnitude below the u and $b^>$ related errors. At the end of the 2015 forecast, however, the $b^<$ errors have grown significantly but remain the smallest contribution. Since $b^>$ is mostly constrained by the apriorily assumed statistics at all times, its related errors change only mildly. While the u related error is smaller than the $b^>$ error at 1990, it becomes the dominant contribution at 2015 due to the increase in flow dispersion.

Thus, neither the observational error in the large-scale magnetic field nor the lack of knowledge on the small-scale contributions is the limiting factor for the predictions but rather the randomization of the different velocity field scales over time.

We also tested the predictability range for the magnetic field. Starting a forecast in 2014, we let the system evolve until the scale per scale variance of the magnetic field exceeds the mean predicted energy. Like in the forecast from 1940 to 2015 illustrated in the upper left panel of Figure 7, after 75 years only, the variance in contributions beyond degree $l = 7$ exceeds the mean energy. The predictability limit further decreases to $l = 5$, $l = 3$, and $l = 2$ after 160, 400, and 640 years, respectively. After 1,950 years, even the dipole energy is exceeded by the respective variance level.

3.7. Variations in the Length of Day

The rotation rate of the Earth, and therefore the length of day (LOD), varies with time. Many geophysical phenomena are responsible for these variations, and over timescales larger than a decade, four of them are expected to explain the departures of the LOD (Λ) from its typical time of $T_0 = 86{,}400$ s . These mechanisms are the following:

The tidal friction (TF). By deforming the Earth's surface, tidal forces induce a dissipation of energy in the Earth-Moon system. As a consequence, the rotation rate of the Earth decreases, and thus, the LOD increases with a rate of $\dot{\Lambda}_{\mathrm{TF}} \sim 2.4$ ms/cy, as estimated by Williams and Boggs (2016).

The glacial isostatic adjustment (GIA). During the last glaciation period, a huge amount of ice accumulated over the polar caps and compressed the mantle. When the ice melted, the mantle tended to regain its initial shape at a rate, which depends on its viscosity profile. These mantle displacements are still generating a continuous decay of the Earth's oblateness J_2. Since the total angular momentum of the Earth has to be conserved, a decrease of J_2 is necessarily accompanied by an increase of its rotation rate. Based on his GIA model, Peltier (2015) has shown that the associated rate at which the LOD is shortened is of $\dot{\Lambda}_{\mathrm{GIA}} \sim -0.6$ ms/cy.

The global sea level rise and ice melting (RSL-IM). Over the last century, the melting of glaciers and ice caps has led to an augmentation of the sea level, as reported in Church and Whit (2011)and Hay et al. (2015). On a global scale, this redistribution of mass at the Earth's surface increases the oblateness of the Earth and therefore lengthens the LOD. Although the rate of this lengthening $\dot{\Lambda}_{\text{RSL-IM}}$ remains uncertain, the analysis of the variations of the Earth's oblateness performed by Cheng et al. (2013) provides some constraints on it. Indeed, Cheng et al. (2013) showed that between 1976 and the 1990s, J_2 was decreasing but not sufficiently fast to be only explained by the GIA. This means that during this period of time $\dot{\Lambda}_{\text{GIA}} + \dot{\Lambda}_{\text{RSL-IM}} > -0.6$ ms/cy. After the 1990s, the effects of the RSL-IM on the LOD are becoming more intense since they completely compensate the variations of J_2 induced by the GIA.

The core-mantle coupling. Jault et al. (1988) have shown that geostrophic motions of the liquid outer core can modify the core angular momentum (CAM). Yet for the Earth to conserve its total angular momentum, fluctuations of the CAM are compensated by variations of the Earth's rotation rate. As shown by Jault and Finlay (2015), the excess in LOD induced by the outer core flow Λ_{FLOW} can be evaluated from the velocity field at the Earth's CMB through the following relation:

$$\Lambda_{\text{FLOW}} = 1.232(\psi_{1,0} + 1.776\psi_{3,0} + 0.08\psi_{5,0} + 0.002\psi_{7,0})\,. \tag{57}$$

Although it is well known that the outer core flow is responsible for decadal variations in the LOD (see Holme, 2015), possible long-term effects are not clear.

3.7.1. Observations and Contradictions

Using compilations of ancient and medieval eclipses and lunar occultations of stars, Stephenson et al. (2016) reconstructed a time series of the excess in LOD between 720 BC and 2015. They highlight, in particular, two main features of the variation in the LOD over the last millenia. First, the LOD exhibits very low frequency fluctuations with a typical period of about 1,500 years. According to Dumberry and Bloxham (2006), these variations are likely to be induced by the outer core flow. Second, Stephenson et al. (2016) could also observe that the LOD was globally increasing at a rate of approximately 1.8 ms/cy. This latter value can be very well explained by the joint effects of TF and GIA since $\dot{\Lambda}_{\text{TF}} + \dot{\Lambda}_{\text{GIA}} \sim 1.8$ ms/cy.

However, over the last 200 years, the time series derived by Gross (2001) with a combination of accurate astronomical techniques indicates that the LOD was globally increasing at a rate of ~1.4 ms/cy. This value is lower than the 1.8 ms/cy predicted by the GIA and TF, yet it should, at least over the last century, be larger than 1.8 ms/cy because of the RSL-IM. If one does not consider a possible influence of the outer core flow, this means that a mechanism increasing the rotation rate of the Earth is missing. This discrepancy between expected and observed LOD constitutes the enigma formulated by Munk (2002).

3.7.2. Estimation of Possible Long-Term Effects of the Outer Core Flow on the LOD

Since the cumulative long-term effects of the GIA, the TF, and the RSL-IM on the LOD are uncertain, we evaluated the influence they should have according to our velocity field solution and to the observed LOD time series (Λ_{OBS}). To do so, the real excess in the LOD is assumed to be given by

$$\Lambda(t) = at + b + \Lambda_{\text{FLOW}}, \tag{58}$$

where the trend $a = \dot{\Lambda}_{\text{TF}} + \dot{\Lambda}_{\text{GIA}} + \dot{\Lambda}_{\text{RSL-IM}}$, b is a constant and Λ_{FLOW} is given by equation (57). Assuming that a and b are a priori unknown (uniform prior over infinite ranges), the posterior distribution of a can be expressed as follows:

$$p(a|\Lambda_{\text{OBS}}) \sim \int p(\Lambda_{\text{OBS}}|a, b, \Lambda_{\text{FLOW}})p(\Lambda_{\text{FLOW}})p(b)\mathrm{d}b\mathrm{d}\Lambda_{\text{FLOW}}\,. \tag{59}$$

Because of the abrupt change in the Earth's oblateness during the 1990s, we restrict the analysis to pre-1990 epochs. Observed LOD variations Λ_{OBS} are taken from Gross (2001), who also provides uncertainty estimates $\Sigma_{\Lambda_{\text{OBS}}}$. The likelihood distribution is approximated by a Gaussian distribution such as the following:

$$p(\Lambda_{\text{OBS}}|a, b, \Lambda_{\text{FLOW}}) = \mathcal{N}\left(\Lambda_{\text{OBS}} - a\mathbf{t} - b - \bar{\Lambda}_{\text{FLOW}}, \Sigma_{\Lambda_{\text{OBS}}}\right). \tag{60}$$

The prior distribution of Λ_{FLOW} is also assumed to be Gaussian with a mean and a covariance deriving from the ensemble of Λ_{FLOW} time series calculated with equation (57). The distribution $p(a|\Lambda_{\text{OBS}})$ is thus also a Gaussian distribution. Our computation suggests a mean of $\bar{a} = 2.2$ ms/cy and a large standard deviation of $\sigma_a = 1.8$ ms/cy which embraces the values discussed above.

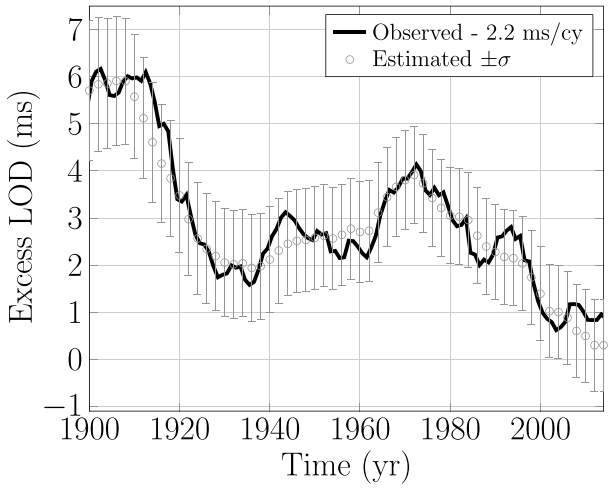

Figure 10. Observed variations in length of day (LOD; black curve) taken from Gross (2001) and extended after 1997.0 with a time series provided by the Earth orientation center. A trend of 2.2 ms/cy estimated in section 3.7 has been removed from the observed time series. The variations in LOD induced by the outer core flow and associated standard deviation are shown with gray circles and error bars.

The mean estimated trend $\bar{a}$ lies well between the 1.8 and 2.4 ms/cy interval prescribed by tidal forces and by the measurements of the Earth's oblateness. Furthermore, once the optimal trend is removed from Λ_{OBS}, comparisons with the core flow contribution exhibit a good agreement on decadal timescales as illustrated in Figure 10. Although the large uncertainty levels associated with $\bar{a}$ and $\bar{\Lambda}_{FLOW}$ forbids any definitive conclusions on the impact of the last century RSL-IM on the variations in the LOD, our results suggest that the outer core flow could be compensating it by globally inducing an increase of the Earth's rotation rate. The equivalent shortening of the LOD due to the global augmentation of the CAM over the last century would be consistent with the oscillatory trend of the LOD estimated by Stephenson et al. (2016), which is decaying between approximately the years 1500 and 2200.

4. Conclusion

We have employed a sequential data assimilation framework to model the dynamics of the geomagnetic field and the flow at the top of Earth's core in the twentieth century using the COV-OBS.x1 model of Gillet, Jault, and Finlay (2015) as observations. The method extends the approach in Baerenzung et al. (2016) to the time domain, as a sequential propagation in time of flow and field inversions under weak prior constraints. The prior is a dynamical model that combines the induction equation in the FF approximation with a simple AR1 process describing the flow evolution. The latter comprises a memory term and stochastic forcing, which are both constrained by the secular variation observations following the ideas presented in Baerenzung et al. (2016).

We use an ensemble approach, the EnKF (Evensen, 2003), where the dynamical model uncertainties are characterized by statistically sound covariances. Using the AR1 process falls short of integrating a proper Navier-Stokes equation but allows us to forward a large ensemble of 40,000 members in time in order to characterize the errors and prior covariances.

The optimal parameters characterizing the flow prior spatial and temporal properties points to particularly long timescales of several centuries to millennia for the toroidal field at SH degrees $l = 1$ and 2. This flow contribution can be attributed to a large-scale slowly evolving gyre, which has also been identified in core flow inversions (Gillet, Jault, & Finlay, 2015; Pais & Hulot, 2000) and numerical simulations (Aubert, 2013; Schaeffer et al., 2017). The most prominent feature of the gyre is the well-documented pronounced westward drift at low and middle latitudes of the Atlantic hemisphere. Smaller-scale flows have characteristic timescales in the decadal range that are consistent with previous estimates (Christensen & Tilgner, 2004; Hulot et al., 2010). Typical related features are local modifications of the gyre in the Southern Hemisphere or the acceleration of the westward flow at and around the tangent cylinder underneath Alaska and the eastern part of Siberia, which has already been reported by Livermore et al. (2017). This points to an important contribution of ageostrophic motions to the dominantly geostrophic overall core flow.

We further tested the capability of our model to forecast the evolution of magnetic and flow fields through hindcast experiments. Comparisons of the magnetic field evolution with linear extrapolations and no casts (in which the field is assumed static) show that our more sophisticated model significantly improves predictions beyond 10 or 15 years.

For predictions of a mean model, we could observe that the advection of the magnetic field by a SF or by a velocity field controlled with a AR1 process does not differ much. Accounting for the flow acceleration, through a second order autoregressive process, for example, would certainly improve the quality of the averaged predicted magnetic field as shown by Whaler and Beggan (2015). Nevertheless, SFs do not allow the proper propagation of uncertainties associated with the magnetic field, since the underlying mechanism for flow dispersion over time is not considered. Within their incorporation in our AR1 model, predicted and prediction errors become consistent. Such match also reveals that the characteristic times estimated for the AR1 process ensure a realistic randomization rate of the fields.

However, it is the dispersion of the velocity field itself that seems to dominate the uncertainties in secular variation estimations, what limits therefore the predictability of the geomagnetic field. Within this limitation, the scale-dependent predictability corresponds to 1950, 640, 400, 160, and 75 years for degrees $l = 1, 2, 3, 5$, and 7, respectively. A more realistic dynamical model such as a geodynamo simulation would possibly extend the predictability horizon. Nevertheless, the enormous numerical power required to perform dynamo simulations at more extreme parameters would preclude the type of EnKF approach followed here. Even with compromises in dynamo model parameters and the ensemble size, the computational costs would still increase by orders of magnitude. Moreover, since dynamo simulations are strongly nonlinear, the system bears an intrinsic sensitivity to initial perturbations. This amounts to an important *e*-folding time, which is estimated to be about 30 years with a temporal rescaling within the secular variation timescale (Hulot et al., 2010; Lhuillier et al., 2011). Since this characteristic time is not so different from the one associated with the small length scales of our flow model, we expect that the rate at which information is lost in the system will be somewhat equivalent in both modeling strategies.

Predictions of decadal LOD variations from changes in angular momentum of our core flow model yield good resemblance to corresponding independent observations. Furthermore, we observe a global increase of the CAM over the last century that we attribute to an acceleration of the geostrophic contribution of the gyre. The resulting decrease in LOD could explain the difference between the recently observed trend in LOD changes and the physically expected one as highlighted by Munk (2002). It would in addition be consistent with the low-frequency oscillatory behavior of the LOD reported by Stephenson et al. (2016).

Appendix A: Gibbs Sampling

The Gibbs sampling algorithm permits to randomly draw an ensemble characterizing statistically a given joint distribution, by recursively sampling conditional probability distributions deriving from it. In our case, the joint distribution of interest is $p(\mathbf{u}, \mathbf{b}|\gamma^o)$. Therefore, at a step n, the algorithm samples alternatively the two following distributions:

$$p(\mathbf{u}^n|\mathbf{b}^{n-1}, \gamma_o, \mathcal{M}) = \mathcal{N}\left(\bar{\mathbf{u}}^n, \Sigma_{u^n}\right) \tag{A1}$$

$$p(\mathbf{b}^n|\mathbf{u}^n, \gamma_o, \mathcal{M}) = \mathcal{N}\left(\bar{\mathbf{b}}^n, \Sigma_{b^n}\right) \tag{A2}$$

with

$$\bar{\mathbf{u}}^n = \bar{\mathbf{u}} + \Sigma_{u_{|\mathcal{M}}} \mathbf{A}_{\mathbf{b}^{n-1}}^{\mathbf{T}} \mathbf{R}_{\mathbf{u}^n}^{-1} \left(\bar{\gamma}^o + \mathbf{A}_{\mathbf{b}^{n-1}} \bar{\mathbf{u}}\right) \tag{A3}$$

$$\Sigma_{\mathbf{u}^n} = \Sigma_{\mathbf{u}_{|\mathcal{M}}} - \Sigma_{\mathbf{u}_{|\mathcal{M}}} \mathbf{A}_{\mathbf{b}^{n-1}}^{\mathbf{T}} \mathbf{R}_{\mathbf{u}^n}^{-1} \mathbf{A}_{b^{n-1}} \Sigma_{\mathbf{u}_{|\mathcal{M}}} \tag{A4}$$

$$\bar{\mathbf{b}}^n = \bar{\mathbf{b}} + = \Sigma_{\mathbf{b}} \mathbf{A}_{\mathbf{u}^n}^{\mathbf{T}} \mathbf{R}_{\mathbf{b}^n}^{-1} \left(\bar{\gamma}^o + \mathbf{A}_{\mathbf{u}^n} \bar{\mathbf{b}}\right) \tag{A5}$$

$$\Sigma_{\mathbf{b}^n} = \Sigma_{\mathbf{b}} - \Sigma_{\mathbf{b}} \mathbf{A}_{\mathbf{u}^n}^{\mathbf{T}} \mathbf{R}_{\mathbf{b}^n}^{-1} \mathbf{A}_{\mathbf{u}^n} \Sigma_{\mathbf{b}} \,. \tag{A6}$$

and where

$$\mathbf{R}_{\mathbf{u}^n} = \left(\mathbf{A}_{\mathbf{b}^{n-1}} \Sigma_{\mathbf{u}_{|\mathcal{M}}} \mathbf{A}_{\mathbf{b}^{n-1}}{}^{\mathbf{T}} + \Sigma_{\gamma}^{o}\right) \tag{A7}$$

$$\mathbf{R}_{\mathbf{b}^n} = \left(\mathbf{A}_{\mathbf{u}^n} \Sigma_{\mathbf{b}} \mathbf{A}_{\mathbf{u}^n}^{\mathbf{T}} + \Sigma_{\gamma}^{\mathbf{o}}\right) . \tag{A8}$$

We recall that the matrices $\mathbf{A_b}$ and $\mathbf{A_u}$ allow us to evaluate the secular variation $\boldsymbol{\gamma}$ when they are, respectively, applied to $\mathbf{u}$ and $\mathbf{b}$.

In this study, three epochs were considered simultaneously, 1900.0, 1950.0, and 2000.0. Since over long periods of time the real temporal correlations of the slow varying components of the flow probably differ from the one induced by the autoregressive process, the characteristic times associated with the degree $l = 1$ and $l = 2$ of the toroidal field and $l = 1$ of the poloidal field were reestimated with a 50-year time step.

Acknowledgments
The authors wish to thank the reviewers of this manuscript, Ciaran Beggan and Richard Holme, for their comments and suggestions to improve the overall quality of the manuscript. This work has been supported by the German Research Foundation (DFG) within the Priority Program SPP1788 "Dynamic Earth." The mean flow coefficients and associated standard deviation used for this study can be downloaded at https://zenodo.org/record/1216857#.Ws36N605aV6.

References

Amit, H., & Olson, P. (2004). Helical core flow from geomagnetic secular variation. *Physics of the Earth and Planetary Interiors*, *147*, 1–25. https://doi.org/10.1016/j.pepi.2004.02.006

Amit, H., & Pais, M. A. (2013). Differences between tangential geostrophy and columnar flow. *Geophysical Journal International*, *194*, 145–157. https://doi.org/10.1093/gji/ggt077

Aubert, J. (2013). Flow throughout the Earth's core inverted from geomagnetic observations and numerical dynamo models. *Geophysical Journal International*, *192*, 537–556. https://doi.org/10.1093/gji/ggs051

Aubert, J. (2014). Earth's core internal dynamics 1840–2010 imaged by inverse geodynamo modelling. *Geophysical Journal International*, *197*, 1321–1334. https://doi.org/10.1093/gji/ggu064

Aubert, J. (2015). Geomagnetic forecasts driven by thermal wind dynamics in the Earth's core. *Geophysical Journal International*, *203*, 1738–1751. https://doi.org/10.1093/gji/ggv394

Aubert, J., Christopher, C., & Fournier, A. (2013). Bottom-up control of geomagnetic secular variation by the Earth/'s inner core. *Nature*, *502*, 219–223. https://doi.org/10.1038/nature12574

Backus, G., Parker, R., & Constable, C. (1996). *Foundations of geomagnetism*. Cambridge, UK: Cambridge University Press.

Baerenzung, J., Holschneider, M., & Lesur, V. (2016). The flow at the Earth's core-mantle boundary under weak prior constraints. *Journal of Geophysical Research: Solid Earth*, *121*, 1343–1364. https://doi.org/10.1002/2015JB012464

Barrois, O., Gillet, N., & Aubert, J. (2017). Contributions to the geomagnetic secular variation from a reanalysis of core surface dynamics. *Geophysical Journal International*, *211*(1), 50–68. https://doi.org/10.1093/gji/ggx280

Beggan, C. D., & Whaler, K. A. (2009). Forecasting change of the magnetic field using core surface flows and ensemble Kalman filtering. *Geophysical Research Letters*, *36*, L18303. https://doi.org/10.1029/2009GL039927

Bloxham, J. (1986). The expulsion of magnetic flux from the Earth's core. *Geophysical Journal*, *87*, 669–678. https://doi.org/10.1111/j.1365-246X.1986.tb06643.x

Bloxham, J., & Jackson, A. (1991). Fluid flow near the surface of Earth's outer core. *Reviews of Geophysics*, *29*, 97–120. https://doi.org/10.1029/90RG02470

Buffett, B. A., & Christensen, U. R. (2007). Magnetic and viscous coupling at the core-mantle boundary: Inferences from observations of the Earth's nutations. *Geophysical Journal International*, *171*, 145–152. https://doi.org/10.1111/j. 1365-246X.2007.03543.x

Canet, E., Fournier, A., & Jault, D. (2009). Forward and adjoint quasi-geostrophic models of the geomagnetic secular variation. *Journal of Geophysical Research*, *114*, B11101. https://doi.org/10.1029/2008JB006189

Cheng, M., Tapley, B. D., & Ries, J. C. (2013). Deceleration in the Earth's oblateness. *Journal of Geophysical Research: Solid Earth*, *118*, 740–747. https://doi.org/10.1002/jgrb.50058

Christensen, U. R., & Tilgner, A. (2004). Power requirement of the geodynamo from ohmic losses in numerical and laboratory dynamos. *nature*, *429*, 169–171. https://doi.org/10.1038/nature02508

Chulliat, A., Hulot, G., & Newitt, L. R. (2010). Magnetic flux expulsion from the core as a possible cause of the unusually large acceleration of the north magnetic pole during the 1990s. *Journal of Geophysical Research*, *115*, B07101. https://doi.org/10.1029/2009JB007143

Church, J. A., & White, N. J. (2011). Sea-Level rise from the late 19th to the early 21st century. *Surveys in Geophysics*, *32*, 585–602. https://doi.org/10.1007/s10712-011-9119-1

Cohn, S. E. (1997). An introduction to estimation theory. *Journal of the Meteorological Society of Japan. Ser. II*, *75*, 257–288. https://doi.org/10.2151/jmsj1965.75.1B257

Dumberry, M., & Bloxham, J. (2006). Azimuthal flows in the Earth's core and changes in length of day at millennial timescales. *Geophysical Journal International*, *165*, 32–46. https://doi.org/10.1111/j.1365-246X.2006.02903.x

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*, 343–367. https://doi.org/10.1007/s10236-003-0036-9

Finlay, C. C., Dumberry, M., Chulliat, A., & Pais, M. A. (2010). Short timescale core dynamics: Theory and observations. *Space Science Reviews*, *155*, 177–218. https://doi.org/10.1007/s11214-010-9691-6

Finlay, C. C., Olsen, N., Kotsiaros, S., Gillet, N., & Tøffner-Clausen, L. (2016). Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model. *Earth, Planets, and Space*, *68*, 112. https://doi.org/10.1186/s40623-016-0486-1

Gillet, N., Barrois, O., & Finlay, C. C. (2015). Stochastic forecasting of the geomagnetic field from the COV-OBS.x1 geomagnetic field model, and candidate models for IGRF-12. *Earth, Planets, and Space*, *67*, 71. https://doi.org/10.1186/s40623-015-0225-z

Gillet, N., Jault, D., & Finlay, C. C. (2015). Planetary gyre, time-dependent eddies, torsional waves, and equatorial jets at the Earth's core surface. *Journal of Geophysical Research: Solid Earth*, *120*, 3991–4013. https://doi.org/10.1002/2014JB011786

Gillet, N., Jault, D., Finlay, C. C., & Olsen, N. (2013). Stochastic modeling of the Earth's magnetic field: Inversion for covariances over the observatory era. *Geosystems, Geochemistry, Geophysics*, *14*, 766–786. https://doi.org/10.1002/ggge.20041

Gross, R. S. (2001). A combined length-of-day series spanning 18321997: Lunar97. *Physics of the Earth and Planetary Interiors*, *123*(1), 65–76. https://doi.org/10.1016/S0031-9201(00)00217-X

Hay, C. C., Morrow, E., Kopp, R. E., & Mitrovica, J. X. (2015). Probabilistic reanalysis of twentieth-century sea-level rise. *Nature*, *517*, 481–484. https://doi.org/10.1038/nature14093

Holme, R. (2015). Large-scale flow in the core. In G. Schubert (Ed.), *Treatise on geophysics second edition* (pp. 91–113). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-444-53802-4.00138-X

Hulot, G., Lhuillier, F., & Aubert, J. (2010). Earth's dynamo limit of predictability. *Geophysical Research Letters*, *37*, L06305. https://doi.org/10.1029/2009GL041869

Jault, D. (2015). Illuminating the electrical conductivity of the lowermost mantle from below. *Geophysical Journal International*, *202*, 482–496. https://doi.org/10.1093/gji/ggv152

Jault, D., & Finlay, C. (2015). Waves in the core and mechanical core-mantle interactions. *Treatise on Geophysics*, *8*, 225–245. https://doi.org/10.1016/B978-0-444-53802-4.00150-0

Jault, D., Gire, C., & Le Mouel, J. L. (1988). Westward drift, core motions and exchanges of angular momentum between core and mantle. *Nature*, *333*, 353–356. https://doi.org/10.1038/333353a0

Kahle, A. B., Vestine, E. H., & Ball, R. H. (1967). Estimated surface motions of the Earth's core. *Journal of Geophysical Research*, *72*(3), 1095–1108. https://doi.org/10.1029/JZ072i003p01095

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*, 35–45. https://doi.org/10.1115/1.3662552

Kuang, W., & Tangborn, A. (2008). MoSST_DAS: The first generation geomagnetic data assimilation framework. *Communications in Computational Physics*, *3*, 85–108.

Lhuillier, F., Aubert, J., Hulot, G., & Earth's dynamo limit of predictability controlled by magnetic dissipation (2011). *Geophysical Journal International*, *186*, 492–508. https://doi.org/10.1111/j.1365-246X.2011.05081.x

Li, K., Jackson, A., & Livermore, P. W. (2014). Variational data assimilation for a forced, inertia-free magnetohydrodynamic dynamo model. *Geophysical Journal International*, *199*, 1662–1676. https://doi.org/10.1093/gji/ggu260

Livermore, P. W., Hollerbach, R., & Finlay, C. C. (2017). An accelerating high-latitude jet in Earth's core. *Nature Geoscience*, *10*, 62–68. https://doi.org/10.1038/ngeo2859

Munk, W. (2002). Twentieth century sea level: An enigma. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 6550–6555. https://doi.org/10.1073/pnas.092704599

Pais, A., & Hulot, G. (2000). Length of day decade variations, torsional oscillations and inner core superrotation: Evidence from recovered core surface zonal flows. *Physics of the Earth and Planetary Interiors*, *118*, 291–316. https://doi.org/10.1016/S0031-9201(99)00161-2

Peltier, W. (2015). The history of the Earths rotation: Impacts of deep Earth physics and surface climate variability. *Treatise on Geophysics Second Edition*, *9*, 221–279.

Roberts, P. H., & Scott, S. (1965). On analysis of the secular variation. *Journal of geomagnetism and geoelectricity*, *17*(2), 137–151. https://doi.org/10.5636/jgg.17.137

Schaeffer, N. (2013). Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations. *Geochemistry, Geophysics, Geosystems*, *14*, 751–758. https://doi.org/10.1002/ggge.20071

Schaeffer, N., Jault, D., Nataf, H.-C., & Fournier, A. (2017). Turbulent geodynamo simulations: A leap towards Earth's core. *Geophysical Journal International*, *211*, 1–29. https://doi.org/10.1093/gji/ggx265

Stephenson, F. R., Morrison, L. V., & Hohenkerk, C. Y. (2016). Measurement of the Earth's rotation: 720 BC to AD 2015. *Proceedings of the Royal Society of London Series A*, *472*(20160404). https://doi.org/10.1098/rspa.2016.0404

Talagrand, O. (1997). Assimilation of observations, An introduction. *Journal of the Meteorological Society of Japan Series II*, *75*, 191–209. https://doi.org/10.2151/jmsj1965.75.1B191

Velímský, J. (2010). Electrical conductivity in the lower mantle: Constraints from CHAMP satellite data by time-domain EM induction modelling. *Physics of the Earth and Planetary Interiors*, *180*, 111–117. https://doi.org/10. 1016/j.pepi.2010.02.007

Whaler, K. A., & Beggan, C. D. (2015). Derivation and use of core surface flows for forecasting secular variation. *Journal of Geophysical Research: Solid Earth*, *120*, 1400–1414. https://doi.org/10.1002/2014JB011697

Whaler, K. A., Olsen, N., & Finlay, C. C. (2016). Decadal variability in core surface flows deduced from geomagnetic observatory monthly means. *Geophysical Journal International*, *207*, 228–243. https://doi.org/10.1093/gji/ggw268

Williams, J. G., & Boggs, D. H. (2016). Secular tidal changes in lunar orbit and Earth rotation. *Celestial Mechanics and Dynamical Astronomy*, *126*, 89–129. https://doi.org/10.1007/s10569-016-9702-3