Geomagnetic forecasts driven by thermal wind dynamics in the Earth's core

J Aubert

HAL Id: insu-01308305
https://insu.hal.science/insu-01308305

Submitted on 27 Apr 2016

Geomagnetic forecasts driven by thermal wind dynamics in the Earth's core

J. Aubert

Institut de Physique du Globe de Paris, Sorbonne Paris Cité, Université Paris-Diderot, CNRS, 1 rue Jussieu, F-75005 Paris, France. E-mail: aubert@ipgp.fr

SUMMARY

There exists a fundamental as well as practical interest in being able to accurately forecast the future evolution of Earth's magnetic field at decadal to secular ranges. This work enables such forecasts by combining geomagnetic data with an Earth-like numerical model of a convection-driven fluid dynamo. The underlying data assimilation framework builds on recent progress in inverse geodynamo modelling, a method which estimates an internal dynamic structure for Earth's core from a snapshot of the magnetic field and its instantaneous rate of change at the surface, and takes advantage of linear relationships and long-range correlations between observed and hidden state variables. Here the method is further evolved into a single-epoch ensemble Kalman filter, in order to initialise at a given epoch an ensemble of states compatible with the observations and representative of the uncertainties in the estimation of hidden quantities. The ensemble dynamics, obtained by subsequent numerical integration of the prognostic model equations, are found to be governed by a thermal wind balance or equilibrium between buoyancy forces, the Coriolis force and the pressure gradient. The resulting core fluid flow pattern is a quasi-steady eccentric gyre organised in a column parallel to Earth's rotation axis, in equilibrium with a longitudinal hemispheric convective density anomaly pattern. The flow provides induction for the magnetic field, which also undergoes a realistic amount of diffusion. Predictions of the present magnetic field from data taken within the past century show that the ensemble has an average retaining good consistency with the true geomagnetic evolution and an acceptable spread well representative of prediction errors, up to at least a secular range. The predictability of the geodynamo thus appears to significantly exceed previous theoretical expectations based on the chaotic divergence of ensemble members. The assimilation generally outperforms the linear mathematical extrapolations from a 30-yr prediction range onwards, with a 40 per cent improvement in Earth-surface error at a secular range. The geomagnetic axial dipole decay observed over the past two centuries is predicted to continue at a similar pace in the next century, with a further loss of 1.1 ± 0.3 µT by year 2115. The focal (or minimum intensity) point of the South Atlantic geomagnetic anomaly is predicted to enter the South Pacific region in the next century, with the anomaly itself further deepening and widening. By year 2065, the minimum intensity is predicted to decrease by 1.46 ± 0.4 µT at the Earth surface and the focal point to move 12.8 ± 1.4 deg westwards with a slight northward component. This corresponds to a drift rate of 0.26 deg yr^{-1}, similar to the typical geomagnetic westward drift observed over the past four centuries. The same drift rate is also predicted until 2115 with a further (but more uncertain) intensity decrease.

Key words: Inverse theory; Dynamo: theories and simulations; Magnetic anomalies: modelling and interpretation; Rapid time variations; Satellite magnetics.

1 INTRODUCTION

A number of human technological activities are in direct interaction with Earth's magnetic field, with some requiring a precise knowledge of its evolution. One of the most critical manifestations of these

interactions is the cosmic ray damage that can arise in low Earth orbiting satellites as they pass through regions of low geomagnetic intensity (e.g. Heirtzler 2002). Another widespread application is the determination of heading in positioning systems through embedded low-cost magnetic sensors. The geomagnetic field is generated
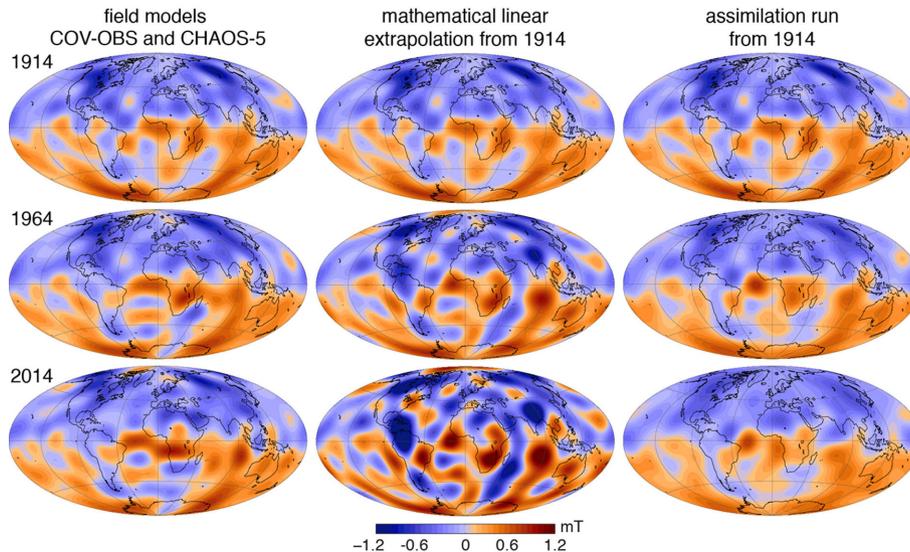
Figure 1. Evolution of the radial magnetic field at the core–mantle boundary (red is outwards, field is truncated after spherical harmonic degree $l_{obs} = 13$, see Fig. 5 for unfiltered assimilation images) between 1914 and 2014, in (left column) the geomagnetic field models COV-OBS (Gillet *et al.* 2013, epochs 1914 and 1964) and CHAOS-5 (Finlay *et al.* 2015, epoch 2014), in a mathematical linear extrapolation from the field and instantaneous rate of change in 1914 (middle column) and (right column) in the ensemble average of an assimilation run started in 1914 from a snapshot of COV-OBS surface field and secular variation data (see Section 3).

by convective motion of electrically conducting liquid iron within Earth's core, and exhibits temporal variations on timescales of years to centuries known as the secular variation (see Jackson & Finlay 2007, for a review). Notable trends of this secular variation over the past 175 yr include a 9 per cent decrease of the magnitude of the geomagnetic axial dipole (e.g. Gillet *et al.* 2013) as well as a broadening and deepening (at an average rate of 3.2 μT century^{-1}) of the region of low geomagnetic intensity at the Earth surface known as the South Atlantic anomaly (SAA; Finlay *et al.* 2010b). Should they continue in the future, such trends could probably be regarded as sizably changing our geomagnetic environment and its technological effects within only decades. This establishes the practical interest of being able to forecast the future evolution of the geomagnetic field at the decadal to secular range.

Current operational forecasts of the geomagnetic field are carried out in the International Reference Geomagnetic Field (IGRF) model issued every five years by the scientific community (most recently Thébault *et al.* 2015). This model comprises a global description of the present field at the Earth surface, together with a prediction for the five years to come. In order to achieve this prediction, most contributions to the IGRF rely on the assumption that the presently observed rate of secular variation does not change in the future, thus enabling a mathematical linear extrapolation of the field using its first time derivative. In an attempt to refine this description and model the second-time derivative or geomagnetic acceleration, the past two IGRF iterations have also seen contributions injecting insight from physical laws governing the geodynamo into the forecast. Estimating core-surface flows responsible for the present secular variation can, for instance, provide a slightly nonlinear prediction (Hamilton *et al.* 2015; Whaler & Beggan 2015). Data assimilation strategies combining the geomagnetic data with numerical models of the geodynamo are also coming to the fore (Fournier *et al.* 2010; Kuang *et al.* 2010; Li *et al.* 2014; Fournier *et al.* 2015, this last study using an earlier version of the methodological framework presented here). When critically assessing the quality of these advanced predictions (Fournier *et al.* 2015; Whaler & Beggan 2015), one generally finds it difficult to improve on the

already excellent field forecast provided by simple linear extrapolation, which typically achieves an Earth-surface root mean square magnetic field error of only about 100 nT or 0.2 per cent of the present field intensity after 5 yr. This result is not surprising when one considers how little the geomagnetic field accelerates over the course of five years. The geomagnetic secular acceleration timescale is indeed on the order of $\tau_{SA} \approx 10$ yr for large-scale structures of the field (Christensen *et al.* 2012), meaning that the rate of secular variation of these structures will only significantly change on a timescale of several τ_{SA}. Such description obviously ignores sudden core flow changes and their possible connection with geomagnetic acceleration jumps (e.g. Olsen & Mandea 2008). However these jumps are not necessarily large (e.g. Christensen *et al.* 2012) and a kinematic description such as the linear extrapolation is sufficient most of the times for the practical purpose of field prediction.

The situation changes when one considers forecasts at a range of several decades. The quality of the linear extrapolation significantly degrades, as illustrated, for instance, in Fig. 1, where neither the amplitude nor the pattern of the radial magnetic field at the core–mantle boundary are correctly predicted after 50 yr in a forecast started in 1914. Here we have much better odds of obtaining more consistent results by making use of physical laws, as also illustrated in Fig. 1. For this promise to materialise, one needs of course a sufficiently realistic numerical model of the geodynamo, as well as data assimilation techniques able to make the most of the injected geomagnetic data. Recently (Aubert *et al.* 2013) we have introduced the coupled Earth (CE) dynamo, a numerical model of a convection-driven fluid dynamo that appears appropriate for this task. In several key areas, the output of this model presents an interesting similarity with the behaviour of the geodynamo over the past four centuries. First, the static morphology of the generated field is Earth-like on the basis of objective quantitative criteria (Christensen *et al.* 2010). Second, the core-surface field kinematics have a pattern similar to that obtained in models of the historical field evolution (Jackson *et al.* 2000). These kinematics are underlain by core surface and deep core flow patterns similar to those retrieved by several inverse methods (Pais & Jault 2008; Gillet *et al.* 2009, 2015; Aubert 2013), and include a

significant westward drift at equatorial position beneath the Atlantic region and a deep, eccentric, axially columnar, gyre-like circulation. This organisation of the CE dynamo model owes to a particular configuration of buoyancy sources and core–mantle–inner core couplings supported by reasonable geodynamic arguments. Lastly, its large-scale geomagnetic acceleration time constant is $\tau_{SA} = 13$ yr, favourably comparing with geomagnetic estimates (Christensen *et al.* 2012, recall that time is rescaled in the CE model by matching the secular variation timescale to its Earth value, so that the value of the acceleration timescale is non-trivial).

Performing an efficient data assimilation first requires the statistical distribution of the observable part of the model to overlap that of the data. The CE dynamo model satisfies this condition (see in particular the corrected misfit results in Aubert 2014b), and it is rather encouraging that the best compatibility is obtained in the last decade, despite the tight data distributions characteristic of highly accurate satellite measurements. The next condition is to be able to infer a physically and statistically consistent internal structure of the hidden part of the model from the data, a difficult task given the respective sizes of the data and model state vectors ($O(10^2)$ versus $O(10^7)$ elements; Aubert & Fournier 2011). The inverse geodynamo modelling framework (Aubert *et al.* 2013; Aubert 2014a) performs this task by taking advantage of linear relationships and long-range correlations existing in the CE dynamo model between observed and hidden state variables (see Section 2). In addition to fitting the core-surface magnetic field, inverse geodynamo modelling solutions present the advantage of accounting for the core-surface secular variation with a realistic balance between magnetic advection by an Earth-like core flow morphology and magnetic diffusion. One weakness of earlier implementations was the handling and propagation of errors related to the estimation of hidden quantities, most notably the unobservable, small-scale magnetic field at the core–mantle boundary. This problem is alleviated in the updated implementation presented here by using an ensemble approach akin to a single step of an ensemble Kalman filter (EnKF; Evensen 1994, 2003). At a given epoch, the method determines a number of core states compatible with the observed core-surface magnetic field and secular variation, and accounting for the spread of the unobservable part of this field. The subsequent numerical integration of the prognostic model equations initialised by these ensemble members provides an average forecast (the best estimator in a least-squares EnKF formulation), as well as a typical spread of the ensemble. This enables a more precise quantification of the so-called butterfly effect, which in the geodynamo has an *e*-folding time (the time taken for an *e*-fold chaotic amplification of the distance between two initially close trajectories) of a few decades (Hulot *et al.* 2010). Furthermore, the ensemble average also reveals the key thermal wind dynamic balance (e.g. Aubert 2005), which is realistically accounted for in the CE dynamo model and governs the forecast evolution. The aim of this paper is thus to present, validate and exploit the first geophysically consistent forecasts of the secular geomagnetic field evolution. In Section 2 we detail the geomagnetic data sources, numerical model and data assimilation framework. Section 3 presents the results, which are then discussed in Section 4.

2 MODELS AND METHODS

2.1 Geomagnetic field models

One prospect of geomagnetic assimilation is to directly handle data with as little pre-processing as possible, as provided, for instance,

by satellite missions such as the recently launched Swarm constellation. In practice, it is however far more convenient to work with higher level products such as mathematical models of the geomagnetic field which provide a set of spherical harmonic coefficients (the Gauss coefficients, see e.g. Finlay *et al.* 2010b) developed on a spline temporal basis for a given epoch range. The first obvious reason is that numerical dynamo simulations such as the one used here also follow a spherical harmonic decomposition in the lateral spatial directions. Another more fundamental reason is that much work is involved into the elaboration of field models, for instance in the data selection and quality check and in the separation of the geomagnetic sources of internal and external origins. This high level of expertise presumably enables more reliable estimates of the field emanating from the Earth interior than one would obtain by starting from scratch. Finally, the instantaneous rate of change of the geomagnetic field can be accurately computed on the spline basis and represents a valuable source of information that can be assimilated in the dynamo model in order to constrain the core flow. As in a previous study (Aubert 2014a), the main geomagnetic field model used is COV-OBS (Gillet *et al.* 2013), spanning 1840–2010. This model further provides arguably reliable data error statistics projected into the spherical harmonic space, thus ready for use in a data assimilation framework. For the epoch range 2010–2015, the field model CHAOS-5 (Finlay *et al.* 2015) is also used.

At a given epoch, geomagnetic data for the magnetic field **B** and its instantaneous rate of change **Ḃ** is supplied as a set of Gauss coefficients $g_l^m, h_l^m, \dot{g}_l^m, \dot{h}_l^m$ in the spectral space. Since we are dealing with the internal field, these data are retained only up to spherical harmonic degree and order $l_{obs} = m_{obs} = 13$. The numerical dynamo model presented in Section 2.2 uses a poloidal–toroidal potential decomposition of the fields, together with a complex-valued spherical harmonic basis Y_l^m, such that the magnetic field **B** at the core surface (radius $r_{CMB} = 3485$ km) is expressed as:

$$\mathbf{B}(r_{CMB}, \theta, \varphi) = \nabla \times \nabla \times [W(r_{CMB}, \theta, \varphi) r_{CMB} \mathbf{e}_r] \quad (1)$$

$$+ \nabla \times [Z(r_{CMB}, \theta, \varphi) r_{CMB} \mathbf{e}_r], \quad (2)$$

$$W(r_{CMB}, \theta, \varphi) = \sum_{m=-m_{obs}}^{m_{obs}} \sum_{l=|m|}^{l_{obs}} W_l^m(r_{CMB}) Y_l^m(\theta, \varphi), \quad (3)$$

$$Z(r_{CMB}, \theta, \varphi) = 0. \quad (4)$$

Here the Earth mantle is assumed to be insulating, such that the toroidal potential Z vanishes at the core surface and the poloidal potential W can be downward continued from its value at Earth's surface. A geocentric set of spherical coordinates (r, θ, φ) is used, $\mathbf{e}_r$ being the unit vector in the radial direction. The Gauss coefficients are converted into this framework to form the data vectors **d** and **ḋ** enclosing the core-surface values of W and its time derivative:

$$\mathbf{d} = \left[\frac{r_{CMB}}{l} \left(\frac{a}{r_{CMB}} \right)^{l+2} (g_l^m + i h_l^m) \right]_{0 < l \leq l_{obs}}^{\mathrm{T}}, \quad (5)$$

$$\dot{\mathbf{d}} = \left[\frac{r_{CMB}}{l} \left(\frac{a}{r_{CMB}} \right)^{l+2} (\dot{g}_l^m + i \dot{h}_l^m) \right]_{0 < l \leq l_{obs}}^{\mathrm{T}}. \quad (6)$$

Here $a = 6371.2$ km is the Earth's radius, i is the imaginary unit and the superscript T denotes the transpose. The data vectors **d**, **ḋ** are supplemented with their respective data error covariance matrices **R** and **Ṙ**, such that

$$\mathbf{R} = E(\boldsymbol{\epsilon} \boldsymbol{\epsilon}'), \ \dot{\mathbf{R}} = E(\dot{\boldsymbol{\epsilon}} \dot{\boldsymbol{\epsilon}}') \quad (7)$$

where E stands for the expectation value, ϵ is the difference between **d** in the geomagnetic field model and its true value in Earth's core and the prime denotes the transpose complex conjugate. Estimates of such data error covariance matrices are supplied by model COV-OBS, as a result of the ensemble approach underlying this model. For model CHAOS-5, we adopt diagonal data error covariance matrices with the following formulation:

$$R_{l,m,l',m'} = \delta_{l,l'}\delta_{m,m'} \frac{r_{\mathrm{CMB}}^2}{l^2(l+1)^2} \left(\frac{a}{r_{\mathrm{CMB}}}\right)^{2l+4} \frac{\epsilon_B^2}{l_{\mathrm{obs}}}, \tag{8}$$

$$\dot{R}_{l,m,l',m'} = \delta_{l,l'}\delta_{m,m'} \frac{r_{\mathrm{CMB}}^2}{l^2(l+1)^2} \left(\frac{a}{r_{\mathrm{CMB}}}\right)^{2l+4} \frac{\epsilon_{\dot{B}}^2}{l_{\mathrm{obs}}}. \tag{9}$$

These correspond to flat energy spectra at the Earth's surface, with integral error levels $\epsilon_B = 5$ nT for **B** and $\epsilon_{\dot{B}} = 2$ nT yr^{-1} for **Ḃ**. Such levels ensure a rather smooth connection of errors between the two geomagnetic field models, left and right of epoch 2010.

2.2 Numerical dynamo model

The CE numerical dynamo model solves for Boussinesq convection for a velocity field **u** and a density anomaly field C, magnetic induction in the magnetohydrodynamic approximation for a magnetic field **B**, in an electrically conducting spherical fluid shell between radii r_{ICB} and r_{CMB}, with the present Earth's core aspect ratio $r_{\mathrm{ICB}}/r_{\mathrm{CMB}} = 0.35$. This shell is coupled to an electrically conducting, axially rotating solid inner core of radius r_{ICB} and to an insulating mantle shell between radii r_{CMB} and $1.83 r_{\mathrm{CMB}}$, also axially rotating. The whole system has a constant solid-body rotation rate Ω defining the planetary rotating frame. The physical formulation of the fluid spherical shell problem can be found in Aubert & Fournier (2011), and the parameters used are those of model 2 in this study. From this basic setup, the CE model is then obtained through a few key refinements related to the couplings at the fluid core boundaries. These are described and discussed in Aubert *et al.* (2013). The first is indirect coupling between the mantle and outer core, via the inner core. This is ensured through a gravitational torque linking the mantle and inner core and a magnetic torque linking the inner and outer core. Practically, this amounts to specifying no-slip and electrically conducting boundary conditions at the inner-core boundary, free-slip and insulating boundary conditions at the core–mantle boundary. The second aspect is the specification of heterogeneous mass anomaly flux boundary conditions at both boundaries of the outer core. The outer boundary condition models the effects of thermal control from the lower mantle, while the inner condition models the effects of thermochemical buoyancy release associated with a longitudinal hemispheric modulation in inner-core growth rate. The buoyancy distribution in the CE model mimics a convection of predominant chemical origin originating at the inner-core boundary, the upper outer core being neutrally buoyant on average. The model thus misses potential effects related to stratification at the top of the core (Buffett 2014). This assumption is rationalized by the inability of the Boussinesq approximation to render strong levels of stratification (see Aubert *et al.* 2013).

The CE dynamo uses a finite-differencing scheme of second order in the radial direction, with $NR = 184$ nodes on an irregular grid refined near the fluid core boundaries ($NG = 25$ of which belong to inner core). In the lateral direction, the model uses the previously introduced spherical harmonic basis for decomposition in the lateral directions, up to spherical harmonic degree and order $l_{\mathrm{max}} = 133$. The poloidal–toroidal

decomposition is used for the magnetic field **B**, with respective complex-valued potentials $W_l^m(r_j), Z_l^m(r_j), j = 1, \ldots, NR, l \leq l_{\mathrm{max}}$. The velocity field is decomposed into a toroidal potential $T_l^m(r_j), j = NG, \ldots, NR, l \leq l_{\mathrm{max}}$ and a poloidal potential, from which an equivalent spheroidal potential $S_l^m(r_j), j = NG, \ldots, NR, l \leq l_{\mathrm{max}}$ can be obtained, in order to facilitate the formulation of the core-surface flow inverse problem (see Aubert 2013). The density anomaly field C is described with a single set of complex scalars $C_l^m(r_j), j = NG, \ldots, NR, l \leq l_{\mathrm{max}}$.

The CE dynamo supplies the core magnetic, density anomaly and velocity fields in dimensionless forms. Given the large parameter space gap between this numerical model and the Earth's core and as in our previous studies (Aubert & Fournier 2011; Fournier *et al.* 2011; Aubert 2013, 2014a; Aubert *et al.* 2013), we do not rescale **u**, **B** and C to the dimensional world by using canonical units, except for the shell gap $D = r_{\mathrm{CMB}} - r_{\mathrm{ICB}} = 2260$ km. We rather rationalise this parameter gap by expressing the model output in units underlain by scaling principles thought to hold in both the model and the Earth's core. Relevant details and discussion on the procedure can be found in Aubert (2014a). For the present work, suffice it to say that these basic principles are the equivalence between the secular variation timescale in the model and in the core (Lhuillier *et al.* 2011), the power-based scaling of the magnetic field amplitude B and the link between convective power and density anomaly C (Christensen & Aubert 2006).

The main strengths of the CE dynamo model in rendering the static and kinematic properties of the geomagnetic field and underlying core flow have been recalled in Section 1. In addition, the implications of two important equilibria are covered here in more details. The first one concerns the magnetic induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B}, \tag{10}$$

where η is the magnetic diffusivity of the core fluid. The magnetic Reynolds number $Rm = [u]D/\eta$, where $[u]$ is a typical velocity in the shell, characterises the relative magnitude of the induction and diffusion terms (respectively first and second terms on the right-hand side of eq. 10). Using the outer core rms velocity to evaluate $[u]$ yields $Rm = 942$ in the CE dynamo, comparable to expected values in the Earth's core (though admittedly possibly on the low side given recent revisions of core electrical conductivity, see Aubert *et al.* 2013). This means that the CE dynamo model realistically renders the magnetic field dynamics, provided that **u** is realistic. This last condition is usually the main weakness of numerical dynamos, because the Navier–Stokes equation determining **u** involves quantities that cannot be set in Earth-like ratios, for computational reasons. Relevant non-diffusive dynamic equilibria can be attained in limited timescale ranges, though. For instance, the thermal wind balance is characterised by the equilibrium between buoyancy forces, the Coriolis force and the pressure gradient. This balance is usually investigated in a curled form where the pressure gradient vanishes:

$$\nabla \times (2\rho\Omega\mathbf{e}_z \times \mathbf{u}) \approx \nabla \times (\mathbf{g}C). \tag{11}$$

Here $\mathbf{e}_z$ is the unit vector in the axial (rotation axis) direction, **g** is the (radial) gravity vector and ρ is the fluid density. Eq. (11) has been shown to describe the long-term behaviour of numerical dynamos (Aubert 2005) such as the CE model, where the thermal wind balance can be assessed by an order one buoyancy number $Bu = g[C]/\rho\Omega[u]$, $[C]$ being a typical density anomaly. This last condition is also satisfied in the Earth's core (Aubert *et al.* 2013). At first order, the thermal wind balance can be expected to hold on timescales distinct from those involving the perturbative terms in

the Navier–Stokes equation, specifically the contributions of inertia and of the Lorentz force. When in balance with the non-gradient part of the Coriolis force, inertia causes fluctuations on timescales of the order of a day. When in balance with magnetic forces, inertia may cause fluctuations on the Alfvén timescale $\tau_A = \sqrt{\mu_0 \rho} D/[B]$, where $[B]$ is a typical field intensity and μ_0 is the vacuum magnetic permeability. The recent record of the internal geomagnetic field is supportive of Alfvén taking the form of axisymmetric torsional oscillations at a period close to 6 yr (Gillet *et al.* 2010). Finally, magnetic forces may be in balance with the non-gradient part of the Coriolis force (and with buoyancy forces), causing MC/MAC waves at periods either much smaller than a year, or longer than a century (see e.g. Finlay *et al.* 2010a). From that discussion it appears that the decadal to secular timescale range may be favourable for the thermal wind balance. However, this is only a crude description as a non-linear interaction between the fundamental timescales outlined here and convective modes may give rise to other periodicities that could have a signature within this range. Furthermore, MAC waves at decadal periods may exist within a stratified layer occupying a thin region at the top of the core (Buffett 2014).

Another point to consider when comparing the output of the CE dynamo model to the geomagnetic record is that the model itself renders some of the equilibria above at incorrect timescales, due to computational limitations. For instance, its magnetic Ekman number $E_\eta = \eta/\Omega D = 1.2 \ 10^{-5}$ is too large by 3.5 orders of magnitude, meaning that the length of the numerical day is too long. This can cause spurious signals in the numerical simulations (see Section 3). Moreover, the Alfvén number $A = \tau_A[u]/D$, measuring the ratio between inertial and magnetic forces, is $A = 0.76$ in the CE dynamo versus expected values $A = 1 - 3 \times 10^{-2}$ in the core (Aubert 2014a). In the CE model, this implies that magnetic forces may cause inertial fluid motions at incorrect timescales (and with excessive damping). In particular, length-of-day variations caused by magnetic torques applied on the inner core and gravitationally transferred to the mantle will also be too slow (though the long-term amplitude can be about correct; Aubert *et al.* 2013). All the above effects may naturally contribute to the forecast errors reported here.

In the following section, an ensemble of states from the CE dynamo will be evolved in time, all compatible with the surface geomagnetic data at the start of the time integration, but with a certain variability in their internal structure. Fig. 2 shows that during the numerical integration, the ensemble average (which is the best estimator of the core state, see the next section) satisfies the thermal wind balance (11). Forecasts can thus be considered to be driven by thermal wind dynamics, an incomplete, but arguably realistic part of Earth's core dynamics. As will be seen in the following section, this owes to the specification of the inverse problem solved initially, which leverages on the strong linear correlations existing between $\mathbf{u}$ and C and to the subsequent dynamics and the ensemble averaging, which tend to reveal the thermal wind balance in the same way as a time averaging. Although inertial contributions are subdominant in the Navier–Stokes equation, thermal wind dynamics should not be considered steady: $\mathbf{u}$ and C are kept in equilibrium, but both forced to evolve since the C is also advected by $\mathbf{u}$.

2.3 Single-epoch data assimilation framework

The large-scale state of Earth's core is described with a set of two vectors, respectively the magnetic and hydrodynamic state vectors

$$\mathbf{b} = [W_l^m(r_j), Z_l^m(r_j)]^T, j = 1, \dots, NR, l \leq l_{\mathrm{asm}}, \quad (12)$$
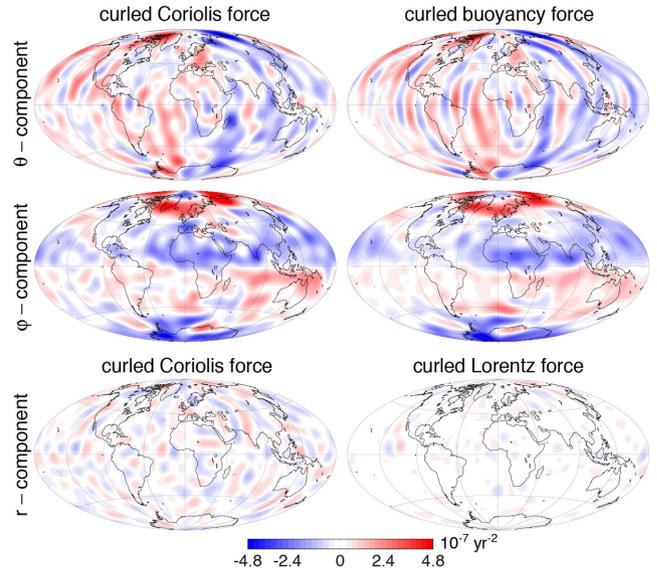

Figure 2. Evaluation of the thermal wind balance (11) on the large scales (spherical harmonic degree and order below 20) of the ensemble average in 1964 of an assimilation run started in 1914 (same as Fig. 1). First two rows: comparison of the lateral components of the Coriolis and buoyancy contributions to the thermal wind balance at mid-depth in the core; last row: comparison of the radial components of the Coriolis and Lorentz contributions at mid-depth (the buoyancy contribution identically vanishes).

$$\mathbf{x} = [S_l^m(r_j), T_l^m(r_j), C_l^m(r_j)]^T, j = NG, \dots, NR, l \leq l_{\mathrm{asm}}. \quad (13)$$

Here $l_{\mathrm{asm}} = 30$ is the spherical harmonic degree up to which the analysis of the core state is performed. Generally the resolution of the analysis should exceed at least twice that of the observations ($l_{\mathrm{asm}} > 2l_{\mathrm{obs}}$), in order to enable a description of the non-linear couplings between small-scale fields that have an observable contribution at large scales.

We wish to use the CE dynamo model in order to obtain an estimate of the true state $\mathbf{b}^t$, $\mathbf{x}^t$ underlying the observations $\mathbf{d}$. To this end, the error distribution of the model states around the truth is represented through the model error covariance matrices

$$\mathbf{P} = E\left((\mathbf{b} - \mathbf{b}^t)(\mathbf{b} - \mathbf{b}^t)'\right), \quad (14)$$

$$\mathbf{Q} = E\left((\mathbf{x} - \mathbf{x}^t)(\mathbf{x} - \mathbf{x}^t)'\right). \quad (15)$$

Recall here that the prime denotes the transpose complex conjugate, as already introduced in eq. (7). The EnKF (Evensen 2003) approximates $\mathbf{P}$ and $\mathbf{Q}$ using an ensemble of states $\mathbf{b}_i$, $\mathbf{x}_i$ of finite size N. As the true state $\mathbf{b}^t$, $\mathbf{x}^t$ is not known, the approximate error covariance matrices are computed around the ensemble averages $\overline{\mathbf{b}}$, $\overline{\mathbf{x}}$:

$$\overline{\mathbf{b}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{b}_i, \quad (16)$$

$$\overline{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i, \quad (17)$$

$$\mathbf{P} \approx \frac{1}{N-1} \sum_{i=1}^{N} \left((\mathbf{b}_i - \overline{\mathbf{b}})(\mathbf{b}_i - \overline{\mathbf{b}})'\right), \quad (18)$$

$$\mathbf{Q} \approx \frac{1}{N-1} \sum_{i=1}^{N} \left((\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})'\right). \quad (19)$$

In this formulation, the ensemble average is interpreted as the best estimator of the true state, and the spread of this ensemble defines the associated estimation error. A broadly distributed initial ensemble can be obtained by sampling a free run (a run unconstrained by the data) of the CE dynamo model. Then, an analysis step balances model and data errors to narrow down the ensemble around an optimal state. The initial ensemble comprises $N = 746$ snapshots of the CE model, stored every 90 yr. This interval is three times longer than the model e-folding time $\tau_e \approx 30$ yr (Aubert & Fournier 2011), thus ensuring statistical decorrelation between adjacent samples. The matrices obtained with (18) and (19) then contain strong cross-correlations arising from linear relationships such as (11) and long-range correlations associated with the dominance of the Coriolis force in the system (Aubert & Fournier 2011; Aubert 2013, 2014a). For a visualisation of the resulting structures of $\mathbf{P}$ and $\mathbf{Q}$ in the case of the CE dynamo, the reader may consult Fournier *et al.* (2015), figs 7 and 8, respectively.

A magnetic observation operator $\mathbf{H}$ is defined, connecting the magnetic state vector $\mathbf{b}$ to the observations $\mathbf{d}$ at the analysis epoch:

$$\mathbf{H}\mathbf{b} = \mathbf{d}. \tag{20}$$

As described in our previous studies (see e.g. Aubert 2013, eq. 22), $\mathbf{H}$ has a simple form consisting of ones in entries corresponding to an observed quantity (the poloidal magnetic field at the core surface) and zeros otherwise. For each ensemble member $\mathbf{b}_i$, the analysed state $\mathbf{b}_i^a$ is defined as a solution to the inverse problem associated to (20), after a reformulation in terms of increments:

$$\mathbf{H}\left(\mathbf{b}_i^a - \mathbf{b}_i\right) = \mathbf{d} - \mathbf{H}\mathbf{b}_i. \tag{21}$$

More specifically $\mathbf{b}_i^a$ is the solution minimising the functional

$$\mathcal{I}^i = \left(\mathbf{d} - \mathbf{H}\mathbf{b}_i^a\right)' \mathbf{R}^{-1} \left(\mathbf{d} - \mathbf{H}\mathbf{b}_i^a\right) \tag{22}$$

$$+ \left(\mathbf{b}_i^a - \mathbf{b}_i\right)' \mathbf{P}^{-1} \left(\mathbf{b}_i^a - \mathbf{b}_i\right), \tag{23}$$

corresponding to the most probable (in the Bayesian sense), minimum variance state equally balancing the model and data errors, both assumed Gaussian (this assumption is usually well verified in numerical dynamos such as the CE model; Fournier *et al.* 2011). The equation through which the ensemble members are updated by the analysis then writes

$$\mathbf{b}_i^a = \mathbf{b}_i + \frac{\mathbf{P}\mathbf{H}'}{\mathbf{H}\mathbf{P}\mathbf{H}' + \mathbf{R}} \left(\mathbf{d} - \mathbf{H}\mathbf{b}_i\right). \tag{24}$$

It should be noted that consistently with the canonical formulation of the EnKF, the inversion proceeds here by minimising the deviation of the analysis to the former value of each ensemble member, whereas earlier studies (Aubert 2013, 2014a) favoured a minimisation of the deviation to the model time average. One can also immediately note that there is no physics involved in the procedure (20)–(24). The analysis only takes advantage of the statistical description contained in $\mathbf{P}$ in order to make an inference for the internal magnetic model structure. The strong information content in $\mathbf{P}$ nevertheless yields coherent images of the interior (Aubert 2014a). In particular, this step enables estimation of a distribution of core-surface magnetic field small scales for $l_{obs} < l \leq l_{asm}$, together with a distribution of estimates of the core-surface magnetic diffusion term $\mathbf{D}_i = \eta \nabla^2 \mathbf{b}_i^a(r_{NR})$. In a second step, and for each ensemble member, the induction equation (10) is linearized at the core surface, and at the analysis epoch, defining a second observation operator:

$$\mathbf{M}_i \mathbf{x} = \dot{\mathbf{d}} - \mathbf{D}_i. \tag{25}$$

The operator $\mathbf{M}_i$ connects the hydrodynamic state vector $\mathbf{x}$ to surface magnetic field and secular variation data. This operator contains non-vanishing values only at entries corresponding to the description of the core-surface induction term $\nabla \times (\mathbf{u}(r_{CMB}) \times \mathbf{B}(r_{CMB}))$. These correspond to the Elsasser and Adams–Gaunt integrals coupling $\mathbf{x}(r_{NR})$ and $\mathbf{b}_i^a(r_{NR})$ (see Aubert 2013, for details). Recall that the core-surface flow problem embedded into (25) can be formulated exactly at the core–mantle boundary, by virtue of the specification of a stress-free external boundary condition in the CE dynamo. Note also again that the analysed core-surface field $\mathbf{b}_i^a(r_{NR})$, provided by (24) up to spherical harmonic degree $l_{asm} = 30$, is included in $\mathbf{M}_i$ and $\mathbf{D}_i$, such that this formulation handles both the effects of magnetic diffusion in the induction equation, together with truncation effects up to spherical harmonic degree l_{asm}. The desire to assimilate secular variation data and solve for core-surface flow by linearising eq. (10) rationalises the choice of a two-state vector formalism and two-step inversion procedure, contrasting with classical EnKF formulations where time derivatives are usually not considered part of the data (e.g. Fournier *et al.* 2013). For each ensemble member, the analysed state $\mathbf{x}_i^a$ is a solution to the inverse problem associated with (25), again after a reformulation in terms of increments:

$$\mathbf{M}_i \left(\mathbf{x}_i^a - \mathbf{x}_i\right) = \dot{\mathbf{d}} - \mathbf{D}_i - \mathbf{M}_i \mathbf{x}_i, \tag{26}$$

and obtained by minimising the functional

$$\mathcal{J}^i = \left(\dot{\mathbf{d}} - \mathbf{D}_i - \mathbf{M}_i \mathbf{x}_i^a\right)' \dot{\mathbf{R}}^{-1} \left(\dot{\mathbf{d}} - \mathbf{D}_i - \mathbf{M}_i \mathbf{x}_i^a\right) \tag{27}$$

$$+ \left(\mathbf{x}_i^a - \mathbf{x}_i\right)' \mathbf{Q}^{-1} \left(\mathbf{x}_i^a - \mathbf{x}_i\right), \tag{28}$$

leading to the most probable, minimum variance state balancing model and data errors. The ensemble analysis equation then writes

$$\mathbf{x}_i^a = \mathbf{x}_i + \frac{\mathbf{Q}\mathbf{M}_i^t}{\mathbf{M}_i \mathbf{Q} \mathbf{M}_i' + \dot{\mathbf{R}}} \left(\dot{\mathbf{d}} - \mathbf{D}_i - \mathbf{M}_i \mathbf{x}_i\right). \tag{29}$$

In contrast with the magnetic inversion steps (20)–(24), the hydrodynamic inversion steps (25)–(29) benefit from the physical content of the magnetic induction equation (10) at the core surface. However, two inversions of statistical nature are also implicitly embedded into this step: the determination of the deep flow and the deep density anomaly field from the core-surface flow, which take advantage of strong linear correlations induced in $\mathbf{Q}$ by the Coriolis force and the thermal wind balance (11), respectively.

Relatively to a previous implementation of inverse geodynamo modelling (Aubert 2014a), several improvements are worth mentioning here. First, the ensemble approach now handles the uncertainties in the estimation of the small-scale content of the core-surface magnetic field and magnetic diffusion, which propagate into eq. (29) through the distribution of $\mathbf{M}_i$ and $\mathbf{D}_i$. Second, the total number of inversion steps has been reduced from three to two, enabling a correct propagation of the surface flow estimation uncertainties into the determinations of the deep flow and density anomaly fields. Third, the estimation of density anomaly is now instantaneous, while it was previously considered valid only for a time average. Indeed, this estimation requires the linear, thermal-wind-driven connection between the velocity and density anomaly fields to hold. The former limitation is now alleviated by the fact that the ensemble average is in thermal wind balance at any given point in the simulation (Fig. 2).

The analysed states determined above are used to initialise numerical integrations of the CE dynamo model. The smallest scale content of these states ($l_{asm} < l \leq l_{max}$) is left unpopulated at the start of the time integration, and subsequently populates through
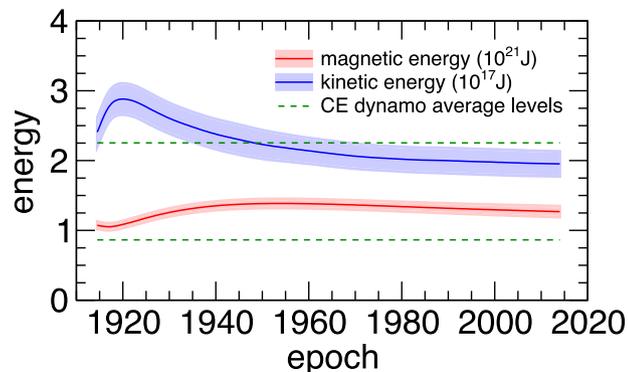
Figure 3. Evolution of the average fluid core kinetic and magnetic energies in the ensemble (solid blue and red lines, respectively), in an assimilation run started in 1914 from geomagnetic data provided by the COV-OBS model (same as Figs 1 and 2). The shaded regions represent the ±1 standard deviation spread. The dashed green lines represent the average energies in the initial ensemble obtained by a free run of the CE model. These show that the assimilation slightly overestimates the magnetic energy but correctly estimates the kinetic energy, a consequence of the statistical versus physical estimations of **B** and **u**.

the non-linear couplings present in the model. At a given subsequent epoch, the numerical integrations provide an ensemble of forecasts $\mathbf{b}_i^f$, $\mathbf{x}_i^f$ again described by an ensemble average (the best estimator) as defined in (16) and (17), and a spread given by the covariance matrices updated using (18) and (19).

In principle, the EnKF formulation makes it possible to use data at multiple epochs and perform a sequential assimilation consisting of several analyses, each one benefiting from the product of the updated covariance matrices from the previous forecast. While this has potentially powerful implications on the reduction of analysis and forecast errors (Evensen 2003), this possibility is not used here yet. This should indeed be seen as a topic for a future study because the capability of the physically limited CE model to correctly represent model errors as they evolve in time should first be assessed—only at this condition can multiple EnKF analyses efficiently narrow down the spread of the ensemble around an unbiased estimate of the truth. Furthermore, the single-analysis approach is the most tractable from a numerical standpoint, thus enabling an extensive exploration of the geomagnetic data set. The matrices **P** and **Q** can indeed be computed once for all using the initial ensemble (unconstrained by the data) of size N, and robust ensemble averages can subsequently be obtained by evolving only about $N/10$ analysed ensemble members (80 members are used throughout this study). In contrast, the task of updating the covariance matrices and performing analyses that efficiently account for the information gained through time integration would require to evolve all N ensemble members, and to significantly raise the assimilation truncation level l_{asm}. Finally, a single-epoch approach is sufficient to obtain interior states that are geophysically consistent (Aubert 2013, 2014a), and also dynamically consistent with the thermal wind balance (see again Fig. 2), with little subsequent spin-up when numerically advanced forward in time (Fig. 3).

3 RESULTS

Fig. 1 presents the secular evolution of the large-scale, ensemble-averaged radial magnetic field at the core–mantle boundary, in an assimilation forecast run started in 1914. The much improved coherence of magnetic field structures predicted until 2014 by the assimilation immediately stands out, in contrast to the physically un-
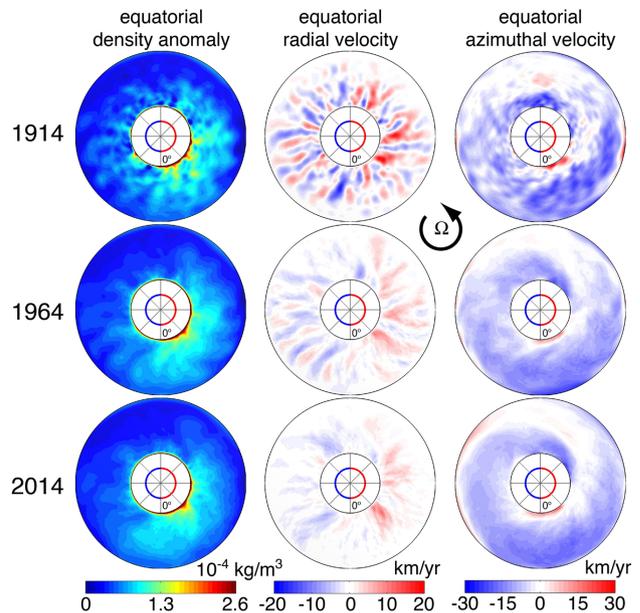

Figure 4. Ensemble-averaged internal core-state evolution in an assimilation run started in 1914 from geomagnetic data provided by the COV-OBS model. Equatorial cuts of the ensemble-averaged density anomaly field (left column, red denotes lighter fluid), radial velocity (middle column, red denotes an outwards flow) and azimuthal velocity (right column, blue denotes a westward flow). The half-circles inside the inner core indicate the orientation of the inner-core mass anomaly flux heterogeneity imposed in the CE model (red is excess flux).

realistic picture given by a mathematical linear extrapolation started at the same time. In particular, the ensemble average correctly renders the localisation of equatorial field patches of high intensity as they drift westward. Other features such as inverse flux under South America, and the relatively standing high-latitude patches of normal polarity are also well rendered (though the amplitude of these last patches appears to fade out more as the assimilation proceeds).

The secular evolution of the ensemble-averaged internal dynamic structure of the fluid core is illustrated in Fig. 4. There is a striking consistency between the images obtained here through numerical integration of the CE model, and those obtained previously through a sequence of single-epoch, static inversions of the geomagnetic data (Aubert 2014a, figs 7 and 8). This confirms that the core-state morphology requested by the data is the one naturally sustained by the CE model dynamics, a key component of the success of the method. As described earlier, this morphology includes a longitudinal hemispheric density anomaly pattern, with larger buoyancy release in the Eastern hemisphere ($0°E − 180°E$). This pattern is in equilibrium with an eccentric, roughly axially columnar (see also core-surface flow signatures in Fig. 5), planetary-scale azimuthal flow gyre reaching the core–mantle boundary below the Atlantic, where the equatorial westward drift is maximal, and touching the inner-core boundary below the Pacific. The ensemble evolution appears to smooth out details in the internal velocity and density anomaly patterns as the assimilation proceeds. These patterns are also slowly advected by the background westward drift (about 0.25 deg yr^{-1}). Compared to Fig. 1, the unfiltered radial magnetic field patterns in Fig. 5 highlight the intense focusing of equatorial magnetic flux patches. Note again the small-scale content brought by the analysis, and the general westward drift of these patches.

Forecast prediction errors at ranges up to 100 yr are quantified in Fig. 6. At any given range t of the numerical integration past the

core surface radial magnetic field   flow below the core-mantle boundary

1914

1964

2014

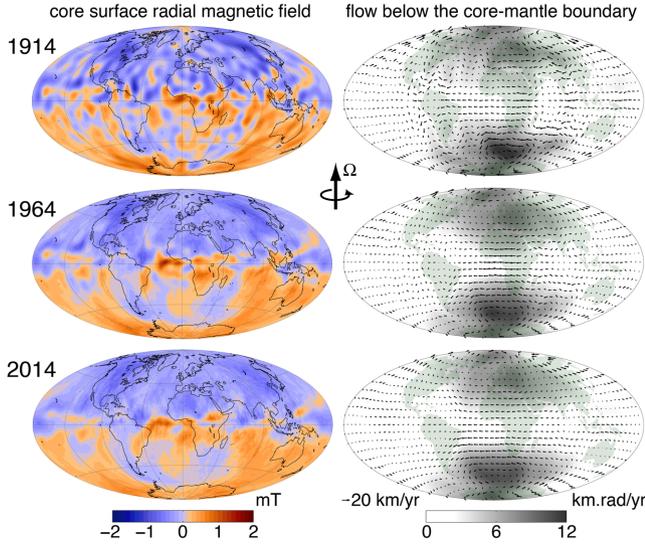mT   −20 km/yr   km.rad/yr

−2 −1 0 1 2   0   6   12

Figure 5. Ensemble-averaged core-surface evolution in an assimilation run started in 1914 from geomagnetic data provided by the COV-OBS model. Core-surface radial magnetic field (left column, same as Fig. 1 but this time presented at full resolution $l_{max} = 133$) and core-surface velocity field (right column, arrows) superimposed on a grey-scale map of the absolute toroidal velocity scalar $|T(r_{CMB}, \theta, \varphi)|$ measuring the amount of local surface rotation.
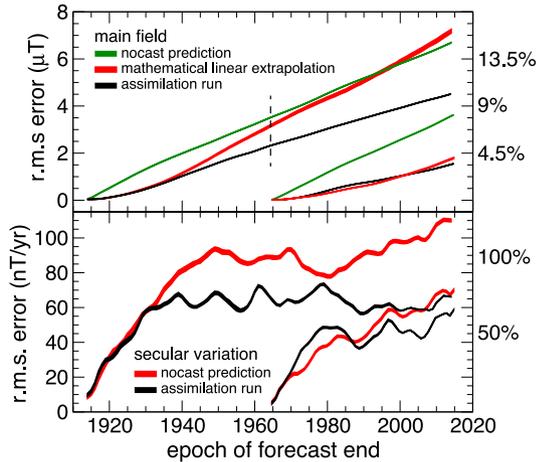

COV-OBS   −9.4 −4.7 0 4.7 9.4   μT

assimilation run

mathematical linear extrapolation

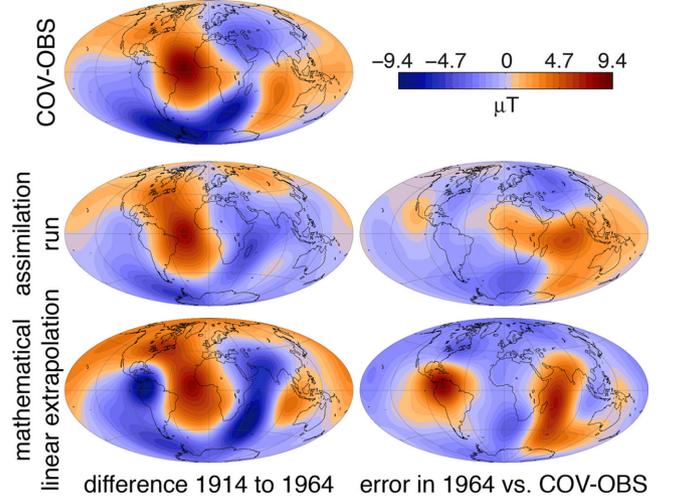difference 1914 to 1964   error in 1964 vs. COV-OBS

Figure 7. Left column: 1914 to 1964 variation in Earth-surface radial magnetic field, for the geomagnetic field model COV-OBS (top), the ensemble average of the assimilation started in 1914 (middle) and the mathematical linear extrapolation (bottom) from 1914. Right column: Earth-surface radial magnetic field error patterns in 1964 for the ensemble average of the assimilation started in 1914 (middle) and the mathematical linear extrapolation from 1914 (bottom).

Figure 6. Prediction errors $\epsilon_B(t)$, $\epsilon_{SV}(t)$ (black lines) plotted versus the prediction range t, in forecasts started in 1914 and 1965 from COV-OBS data. The errors are evaluated following eqs (30) and (31), as the distance between the numerically integrated ensemble average and COV-OBS/CHAOS-5. Also reported are errors obtained in a mathematical linear extrapolation of the field, corresponding to a no-cast of the secular variation (both assuming a constant secular variation, red lines) and errors obtained in a no-cast of the field itself (green lines). The right axis reports the errors relative to the magnetic field and secular variation intensities in 2014 (respectively 43.7 μT and 89 nT yr^{-1}) and the vertical dashed line marks epoch 1964. Line thicknesses represent the ± 1 std. dev. related to the uncertain estimation of the true core field by COV-OBS and CHAOS-5 (see Section 2.1).

analysis time, these errors are defined as

$$\epsilon_B(t) = \sqrt{\frac{1}{\mathcal{S}_E} \int_{\mathcal{S}_E} \left(\overline{\mathbf{B}}^f(t) - \mathbf{B}^t(t) \right)^2 dS}, \tag{30}$$

$$\epsilon_{SV}(t) = \sqrt{\frac{1}{\mathcal{S}_E} \int_{\mathcal{S}_E} \left(\dot{\overline{\mathbf{B}}}^f(t) - \dot{\mathbf{B}}^t(t) \right)^2 dS}. \tag{31}$$

Here $\mathcal{S}_E$ is the surface of the Earth, $\overline{\mathbf{B}}^f$ is the ensemble averaged magnetic field forecast and $\mathbf{B}^t$ is the truth, here defined by COV-OBS and CHAOS-5 (note though that Figs 6 and 8 account for the fact that the truth may differ from COV-OBS and CHAOS-5 according to their data error covariance matrices). The errors generated by the assimilation are comparable to those of a mathematical linear extrapolation at ranges up to approximately 30 yr (but see Fig. 8 for further analysis). This confirms that the secular variation rate is correctly assimilated. This also underlines the weak benefits that can be expected from the current assimilation until a range of approximately three times the large-scale geomagnetic secular acceleration timescale τ_{SA}. At ranges beyond 30 yr, the prediction of $\mathbf{B}$ by the assimilation clearly outperforms the linear extrapolation, with an error reduced by 40 per cent in the 2015 forecast performed from 1915. Given the observed rates of error increase, this gain presumably amplifies at ranges greater than 100 yr. It should be noted that the linear extrapolation error is not bounded in principle. In contrast, an upper bound exists for the assimilation error, corresponding to the difference between a free-run time average of the CE dynamo, which the ensemble should ultimately reach at infinite prediction ranges (see Fig. 9) and the truth at the final epoch. For epoch 2015 this upper bound corresponds to $\epsilon_B(\infty) = 13.8$ μT and $\epsilon_{SV}(\infty) = 91$ nT yr^{-1}. The assimilation prediction error of $\mathbf{B}$ is far below this upper bound at a secular range, while that of $\dot{\mathbf{B}}$ appears to reach a plateau also significantly below the upper bound. The detailed Earth-surface error patterns on $\mathbf{B}$ are further analysed in Fig. 7, showing that the dominant error is carried by low-order multipoles, with coefficients of degree 2, order 0 and degree 1, order 1 roughly representing half of the mean squared error. The assimilation clearly performs better than the linear extrapolation, especially so in the Western hemisphere. Most of the remaining error appears to be concentrated in the Eastern hemisphere, a location which is typically problematic because this is where core–mantle boundary flux patches are chaotically generated in the geodynamo (Jackson *et al.* 2000; Finlay *et al.* 2012).
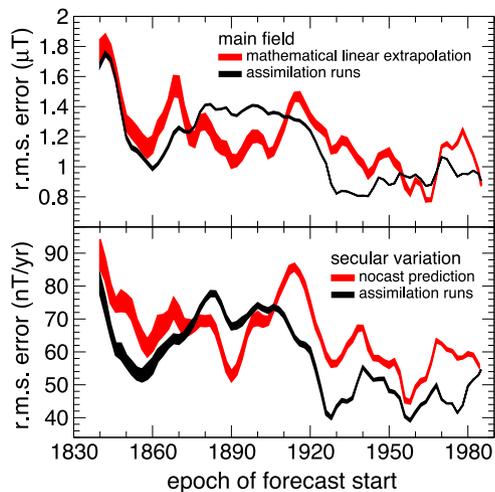
Figure 8. Prediction errors $\epsilon_B(30)$, $\epsilon_{SV}(30)$ (black lines) in 30-yr fixed-range forecasts started in [1840, 1842, ..., 1984, 1985] from COV-OBS data. Also reported are errors obtained in mathematical linear extrapolations of the field and in no-casts of the secular variation (red lines). Line thicknesses represent the ±1 std. dev. related to the uncertain estimation of the true core field by COV-OBS and CHAOS-5 (see Section 2.1).

Fig. 6 shows that at prediction ranges below 30 yr, the two forecasts started in 1914 and 1965 do not behave in a completely similar manner relatively to the linear predictions. A more detailed analysis of this effect is presented in Fig. 8, where a comprehensive set of fixed-range 30-yr forecasts are started between 1840 and 1985. At this range, the average assimilation prediction errors $\overline{\epsilon_B}(30) = 1.14\ \mu$T, $\overline{\epsilon_{SV}}(30) = 58$ nT yr^{-1} outperform those of the linear extrapolation (respectively 1.18 μT, 65.1 nT yr^{-1}), but there are also significant error fluctuations with respect to the epoch of forecast start. These comprise a long-term decreasing trend, of about 0.3 μT century^{-1} for the main field and faster variations, most visible in the results of linear extrapolation, corresponding to short-timescale geomagnetic acceleration phenomena not rendered by either prediction method. Forecasts started after epoch 1920 are generally better than those started earlier, and consistently outperform the linear extrapolations by a larger margin. Among core dynamics phenomena accounting for the unpredicted geomagnetic acceleration, one can presume that the 60-yr geomagnetic signal (Roberts *et al.* 2007; Buffett 2014) plays a role, as appears to be confirmed by a rough similar periodicity in the linear prediction results (though this periodicity remains elusive in the assimilation results).

Interestingly, Fig. 6 shows that the assimilation prediction results remain well below, and (in the case of the secular variation) do not appear to converge towards their asymptotes at infinite prediction range. This suggests that the ensemble is able to preserve some of the initial deviations introduced by the data at long range despite the smoothing effect observed in the ensemble average (Figs 4 and 5). This smoothing is caused by the chaotic divergence of the ensemble members, and should ultimately bring the ensemble average back to the CE model time average. This conjecture is investigated in more details in Fig. 9. We first define the mean m and spread σ of the core–mantle boundary radial magnetic field B_r in the initial (free run) ensemble of CE dynamo states:

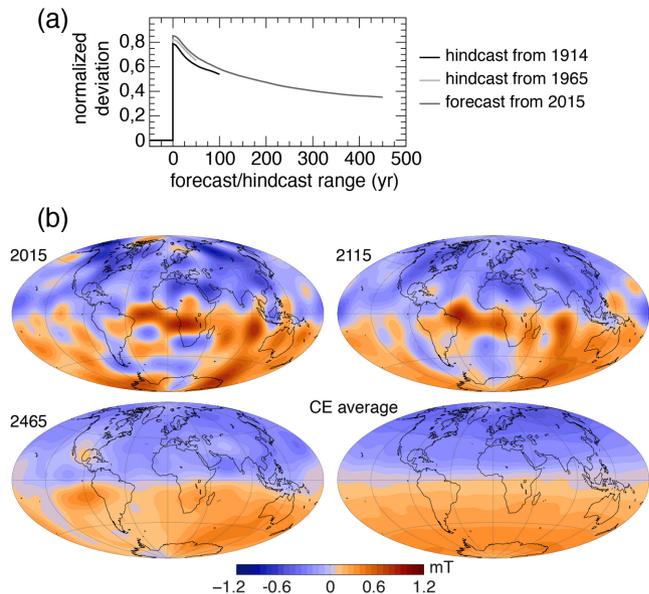$$m = \frac{1}{N} \sum_{i=1}^{N} B_r^i, \qquad (32)$$

Figure 9. (a) Normalized deviation $\mathcal{D}(t)$ brought by the analysis, as defined by eq. (34), plotted versus prediction range t, in forecasts started in 1914 and 1965 from COV-OBS data and in 2015 from CHAOS-5 data. (b) Ensemble-averaged core–mantle boundary radial magnetic field predicted in 2015, 2115, 2465 (red is outwards, field is truncated after spherical harmonic degree $l_{obs} = 13$) in the forecast started in 2015, compared with the free-run time average of the CE dynamo that should be reached at infinite prediction range.

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \int_{\mathcal{S}_{CMB}} \left(B_r^i - m \right)^2 dS}, \qquad (33)$$

where $\mathcal{S}_{CMB}$ is the core–mantle boundary surface. The deviation of the ensemble average to the asymptote defined by the CE model free-run time average, normalized by the free-run spread, is then:

$$\mathcal{D}(t) = \frac{1}{\sigma} \sqrt{\int_{\mathcal{S}_{CMB}} \left(\overline{B_r}^f(t) - m \right)^2 dS}, \qquad (34)$$

Prior to analysis, the initial ensemble is the free-run ensemble, hence $\mathcal{D}(0^-) = 0$, as illustrated in Fig. 9(a). Right after the analysis, $\mathcal{D}(0^+)$ is close to 1, meaning that the analysis has produced an ensemble average with a deviation from the free-run average typical of the free-run spread, as should be expected. As the forecast range increases, $\mathcal{D}(t)$ decreases, as a result of the increasing dispersion of the ensemble leading to smoothing (see Figs 4 and 5), and presumably reaches zero again at infinite prediction range. This decrease is quite slow, though, and some deviation can strikingly be preserved even after several hundred years, as illustrated in Fig. 9(b) where a forecast of epoch 2465 from epoch 2015 still shows significant deviations from the CE model free-run time average. This result rationalises the prediction error behaviour obtained in Fig. 6, and shows that even though ensemble members diverge from each other with an *e*-folding time constant $\tau_e \approx 30$ yr, the ensemble average retains some predictive power well beyond τ_e. Fig. 9(b) also illustrates the assimilation predictions for the coming century. High-latitude flux lobes are predicted to remain roughly stable (though these appear to fade, an artefact presumably related to the CE model that is also observed in Fig. 1). The westward drift of low-latitude equatorial flux patches of normal polarity is predicted to continue just as observed over the past 400 yr (Jackson *et al.* 2000). The assimilation
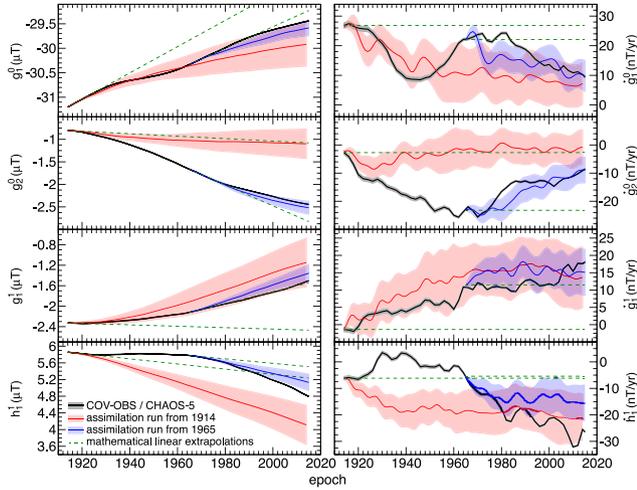
Figure 10. Temporal evolution of selected Gauss coefficients (as defined for instance in Finlay *et al.* 2010b) for the magnetic field (left) and its secular variation (right), in the geomagnetic field models COV-OBS and CHAOS-5 (black lines), in ensemble averages of forecasts started in 1914 (red lines) and 1965 (blue lines) from COV-OBS data and in mathematical linear extrapolations started at the same epochs (dashed green lines). Colour shadings represent the ± 1 standard deviation uncertainty range, as provided by the geomagnetic field models (grey) and by the ensemble spread in the assimilations (pink and light blue). Selected coefficients include the dipole g_1^0 (top row), together with the three low-order multipoles g_2^0, g_1^1 and h_1^1 (bottom three rows) most significantly contributing to the forecast error in Fig. 7.
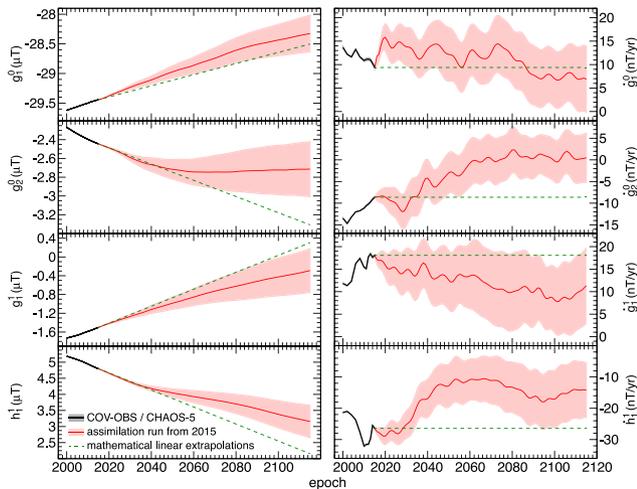

Figure 11. Temporal evolution of selected Gauss coefficients (same set as Fig. 10), for a forecast started in 2015 from CHAOS-5 data (red line is ensemble average, pink shading is the ± 1 standard deviation uncertainty range, dashed green line marks the linear mathematical extrapolation).

also predicts a similar drift of the South-Atlantic region of inverse flux.

Examining the evolution of low-order Gauss coefficients in forecasts started either in the past or in the present (Figs 10 and 11, respectively) provides another way to represent errors, quantify the chaotic divergence of the ensemble, and assess future trends in the geomagnetic field evolution. Fig. 10 confirms that most of the forecast error is contained in low-order multipoles, particularly of degree 2, order 0 and degree 1, order 1, as already illustrated in Fig. 7. At a forecast range of 50 yr, starting from data in the more recent past (i.e. 1965 rather than 1914) not only yields an ensemble

average providing a better prediction of the truth (see also Fig. 8), but also reduces the ensemble spread. This spread also provides a reasonable estimate of the prediction error on Gauss coefficients, with the truth lying within one to two standard deviations away from the ensemble average in most cases. A notable exception though is the evolution of g_2^0 in the forecast starting from 1914. At a secular range, the spread also remains typically smaller than the predicted variations, thus demonstrating the viability of thermal wind-based forecasts of the field at this range.

On decadal timescales, the dynamics produced by the assimilation highlight some of its shortcomings, when compared to the true signal. First, the true signal (Fig. 10) exhibits, for instance, a fluctuation of the dipole secular variation $\dot{g}_1^0$ with an apparent timescale close to 60 yr, which is not rendered by the assimilation (see Section 4 for a discussion on the possible origins of this signal). Second, the assimilation itself displays smaller, rapid fluctuations at a period close to 10 yr (seen in all secular variation plots of Figs 10 and 11), which are not present in the true signal. A closer analysis reveals that these occur at the period of the planetary rotation rate in the numerical model (which is too slow as the magnetic Ekman number is too large, see Section 2.2). This suggests a transient equilibrium between inertial forces and the part of the Coriolis force not balanced by buoyancy and pressure forces. Such spurious fluctuations can thus be interpreted as an inertial response to departures from the thermal wind balance. Though they remain weak respectively to the longer term variations predicted by the assimilation, they are clearly not realistic and underline the limitations of the numerical simulation and assimilation scheme.

The ensemble average evolution observed in Figs 10 and 11 is constrained both by the initial trend fitting the observations and by the long-range relaxation towards a mostly axially dipolar field with strength $|g_1^0| = 35$ μT and vanishing secular variation (see Fig. 9). In some cases, the dominant effect is the relaxation towards the asymptotic state, and the assimilation performs better at times when the true geomagnetic evolution also displays similar characteristics. For instance, the forecast of g_2^0 in Fig. 10 shows low error when started from 1965, because the true value of $|g_2^0|$ decelerates. Conversely, the same forecast is notably worse when started from 1914 because the true value of $|g_2^0|$ accelerates. Another example is the $|g_1^0|$ dipole decay predicted by the assimilation, which also tends to systematically decelerate, thus frequently leading to an underprediction of the true decay. In other cases, though, the effect of initial conditions can be felt rather strongly and can offset the natural relaxation towards the asymptotic state. For instance, all our forecasts predict a continuation of the $|g_1^0|$ decay, although the initial $|g_1^0|$ is already below the free-run average $|g_1^0| = 35$ μT. Similarly, both $|g_1^1|$ and $|h_1^1|$ in Fig. 10 accelerate, thus appearing to further depart from the asymptotic state. At long range these trends should presumably reverse, but within the first 50–100 yr the effects of initial conditions appear to dominate here.

Fig. 11 provides a forecast for the dipole evolution in the next century, with a further decay of $\Delta g_1^0 = 1.1 \pm 0.3$ μT between 2015 and 2115. Opposite to the natural tendency to deceleration outlined above, the rate of dipole decay is actually predicted to slightly increase in the next 70 yr, by 2.7 nT yr^{-1} on average. Still, it is possible that this decay remains underpredicted, if the apparent 60-yr fluctuation of the observed dipole secular variation enters a rising half-period in the next 30 yr.

As outlined in Section 1, the evolution of the magnetic field intensity at the Earth surface is of high practical interest, particularly concerning the depth and extent of the SAA. Fig. 12 presents forecasts of its evolution, and Table 1 quantifies the quality of
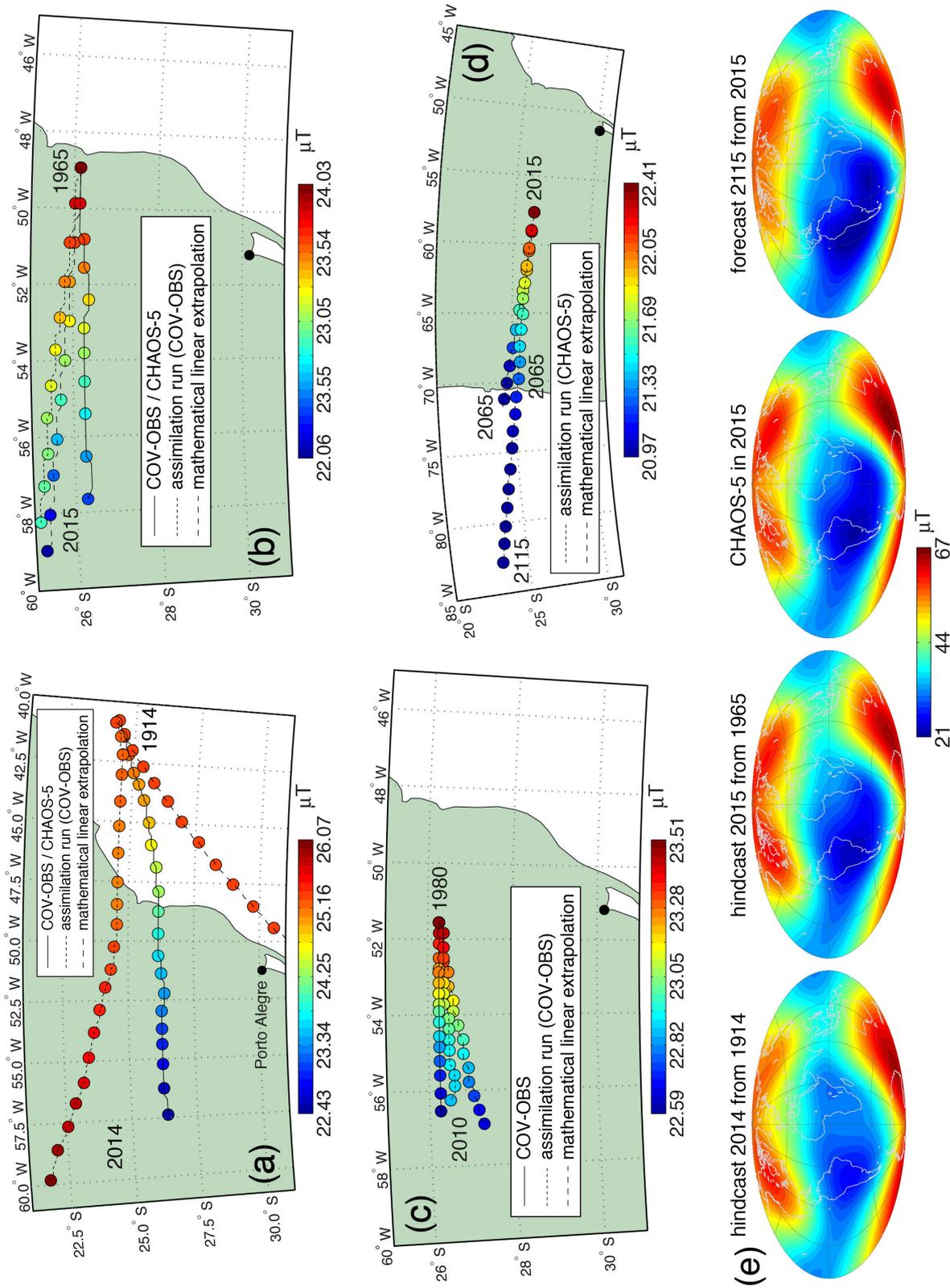
Figure 12. (a)–(d) Evolution of the focus (minimum intensity point) of the South Atlantic anomaly at Earth's surface, in geomagnetic field models COV-OBS and CHAOS-5 (solid lines), in the ensemble average of assimilation runs started at various epochs (short dashed lines) and in linear mathematical extrapolations (long dashed lines). Coloured circles mark the intensity during the temporal evolution. (e) Earth-surface intensity of the geomagnetic field for ensemble average forecasts started from 1914 and 1965 (first and second plot from the left), for the true 2015 field from CHAOS-5 (third plot) and for an assimilation forecast of the 2115 field from CHAOS-5 in 2015 (fourth plot).

Table 1. Angular distance between the predicted and true focus of the South Atlantic anomaly at the Earth surface (first number, degrees), and difference between the predicted and true focal intensity (second number), in assimilation forecasts and mathematical linear extrapolations from various start epochs and at various ranges. The linear extrapolation of 2014 from 1914 (bottom right) produces an artefact with the apparition of a second intensity focus beneath the South African region.

Forecast range	Forecast start epoch	Assimilation error	Linear prediction error
30 yr	1914	1.72°, 0.33 μT	1.24°, 0.51 μT
30 yr	1965	1.15°, 0.13 μT	1.34°, 0.19 μT
30 yr	1980	0.36°, 0.27 μT	1.07°, 0.05 μT
50 yr	1914	2.25°, 1.24 μT	3.61°, 1.26 μT
50 yr	1965	1.26°, 0.48 μT	1.66°, 0.33 μT
100 yr	1914	5.21°, 3.61 μT	105°, 0.03 μT

predictions. The assimilation prediction of the SAA evolution in the secular forecast started from 1914 (Fig. 12a) is not good, both for the localisation and the intensity of the focus. The mathematical linear extrapolation also very quickly diverges from the true trajectory, suggesting that the data quality is insufficient to perform such a prediction (see also Figs 6, 8 and 10). Things improve drastically in forecasts started from 1965 and 1980 (Figs 12b and c), where the assimilation generally outperforms the linear extrapolation in pinpointing the westward drift of the focus. However, the assimilation consistently underpredicts the focus intensity decrease (as could be anticipated from the underprediction of the dipole evolution in Fig. 10). The linear extrapolation tends to overpredict it, such that both intensity predictions are found to be roughly of similar quality. The forecast for the coming century (Fig. 12d) highlights a similar evolution for the assimilation and the linear extrapolation up to 2065, with the assimilation again predicting less intensity decrease than the linear extrapolation, though the two are now very close. Averaging the two predictions, and taking the prediction errors into account (Table 1, error bars taken from epoch 1965 and a prediction at 50-yr range) leads to a predicted westward displacement of the focus of 12.8 ± 1.4 deg by 2065, with a slight northward component. This corresponds to an average westward drift rate of 0.26 deg yr^{-1}, comparable to the typical westward drift rate produced in the CE model. The average prediction for the SAA focus intensity decrease by 2065 is 1.46 ± 0.4 μT, comparable to the rate of intensity decrease observed over the past century (Finlay *et al.* 2010b). The SAA should thus also significantly broaden in the future (Fig. 12e). The mathematical linear extrapolation predictions for 2065–2115 completely diverge from those of the assimilation, as the focal point shifts to the Atlantic, southwest of Africa (same as in Heirtzler 2002). In contrast, the 2065–2115 assimilation predicts that the focus will continue its westward drift and enter the southern Pacific region. This latter prediction appears more realistic, as the appearance of a second focus in the linear prediction is presumably the same artefact as that obtained when attempting a linear prediction of 2014 from 1914 (Fig. 12a and Table 1). The assimilation finally predicts a further decrease of the focal intensity between 2065 and 2115.

4 DISCUSSION

Respectively to recent developments in geomagnetic prediction and data assimilation (Kuang *et al.* 2010; Li *et al.* 2014; Whaler & Beggan 2015), the system presented here has the novelty of allying geodynamic and statistical consistency in its analysis of the core state. Geodynamic consistency is first achieved through the use of a physically reasonable direct numerical model of a convection-driven dynamo with Earth-like output. Second, the direct model and the assimilation scheme correctly render key dynamic equilibria such as the thermal wind balance and the partition of geomagnetic secular variation into advection by the core-surface flow and magnetic diffusion. Statistical consistency is achieved by virtue of the overlap existing between the statistical data and model spaces, both realistically sampled through the specification of dense error covariance matrices. Statistical and geodynamic consistencies lead to dynamic trajectories (Figs 1, 4 and 5) that are coherent because they are both required by the data and naturally sustained by the physical model.

Forecasts provided in this framework are operational and viable at the secular range for several reasons. First, the predictions outperform mathematical extrapolations that can be made with the same amount of data (Fig. 6). Second, the ensemble spread generally quantifies well the prediction uncertainty, and this spread remains smaller than the predicted variations (Figs 10 and 11). Third, a single-epoch analysis suffices to bring a strong information content into the system, which is preserved at a range significantly exceeding the *e*-folding time $\tau_e \approx 30$ yr (Fig. 9) common to the CE model and to the geodynamo (Hulot *et al.* 2010).

The assimilation supports the remanence on secular timescales of the eccentric gyre (Fig. 4) previously imaged inside the core throughout the historical geomagnetic era (Pais & Jault 2008; Gillet *et al.* 2009, 2015; Aubert 2013). As expected, the gyre remains in thermal wind dynamic equilibrium with a longitudinal hemispheric buoyancy distribution, with stronger buoyancy release in the Eastern hemisphere. Advection of this buoyancy pattern by the westward internal core flow causes the gyre to rotate westwards, as already observed in the unconstrained CE dynamo model (Aubert *et al.* 2013) and as also inferred from a sequence of static core inverse images (Aubert 2014a). In that sense, the dynamics brought by the assimilation cross-validate the concepts put forward by the CE dynamo model and by the inverse geodynamo modelling technique.

Somewhat unsurprisingly, the assimilation predicts that the next century of geomagnetic evolution may be very similar to what has been seen over the past four centuries. At the core–mantle boundary, intense equatorial magnetic flux patches of normal polarity will continue to drift westwards below the Atlantic, together with the south Atlantic region of inverse flux, while high-latitude patches of normal polarity will undergo less motion (Fig. 1). At the Earth surface, the dipole will continue its decay at a rate similar to that of previous centuries (Fig. 11), while the SAA will continue to deepen and enlarge similarly to its past behaviour, its focus drifting westwards and entering the South Pacific region (Figs 12d and e). Much of this continuity in behaviour owes to the stability of this geomagnetic anomaly respectively to the eccentric gyre (as seen in Fig. 5). Indeed, the dipole decay is caused by meridional advection of the core–mantle boundary radial magnetic flux by the gyre surface flow

(Olson & Amit 2006). Here the imbalance provided by the SAA may be essential to tilt the budget towards a decay of the dipole. The anomaly itself is maintained by the action of the anticyclonic gyre in the south Atlantic, which disperses core-surface radial magnetic flux towards its edges (Aubert *et al.* 2008). The system will presumably remain stable for as long as, at the core surface, the magnetic anomaly and the gyre remain locked to each other while drifting westwards, as appears to be predicted for the next century by the forecast started in 2015. Beyond that range, a continuation of the dipole decay is also predicted (not shown here) but this prediction is presumably less trustworthy because the information loss is large (Fig. 9). At such long ranges, more dramatic events such as the appearance of new core-surface flux patches below the Indian ocean, which would disrupt the basic configuration, may not be well captured by the assimilation (see further below). Finally, regarding the practical impact of the predicted field variations on the future geomagnetic environment, the results give more credence to previous conclusions based on mathematical linear extrapolations (Heirtzler 2002). The deepening and enlargement of the SAA will correspondingly enlarge the problematic flight area for satellites and increase the likeliness of failure within this area.

Prediction errors of the scheme are due for some part to the limitations of the direct numerical model. For instance, thermal wind dynamics is clearly unable to account for the signal of apparent 60-yr periodicity in the dipole secular variation (Fig. 10), leading to fluctuations in the errors at 30-yr range (Fig. 8), and confirming that additional modelling is needed to account for this signal. Inclusion of a stratified region at the top of the core that can sustain magneto-Archimedes–Coriolis (MAC) waves at the correct frequency (Buffett 2014) represents a possible future improvement to the CE dynamo model. However, this may represent significant difficulties as the required level of stratification is far beyond what can typically be achieved in a Boussinesq framework, thus requiring more elaborated anelastic formulations. Furthermore, since the 60-yr signal appears to have both a geomagnetic and a geodetic signature (Roberts *et al.* 2007), it is presently not clear how such confined MAC waves could also account for the coupled variations in the length of the day. Here it appears necessary to invoke bulk mechanisms efficiently coupling the poloidal flows causing fluctuations in the dipole variations and the toroidal flows at the origin of length-of-day variations. In the hypothesis that core–mantle coupling is indeed primarily of magnetic origin (directly at the core–mantle boundary or via an additional gravitational torque linking the inner core and the mantle), such mechanisms would also need to involve Lorentz forces at the correct amplitude relative to inertia, in order to efficiently transmit angular momentum to the mantle. This motivates further research beyond thermal wind dynamics in order to elucidate this signal.

Perhaps the largest part of prediction errors arises from limitations in the data assimilation scheme. In our tests, a significant part of the error pattern (Fig. 7) indeed originates from low-latitude regions below the Indian ocean, a place where, at the core surface, radial upwelling at the eastern limb of the gyre is thought to be responsible for the creation of the intense, westward drifting magnetic flux patches of normal polarity (Finlay *et al.* 2012; Aubert *et al.* 2013). The exact sequence of appearance of these patches can presumably be accurately predicted only if the deep underlying magnetic structure is reasonably well estimated. The assimilation scheme relies on a purely statistical estimation of the magnetic field, and a single-epoch analysis probably reveals its limits at this point. Here clear benefits can be expected from a multi-epoch EnKF assimilation scheme, which will help to enable a dynamically, rather than statistically consistent estimate of the deep magnetic structure. Other forthcoming improvements to the method include a higher resolution analysis (higher values of l_{asm}). The use of temporal correlations in the analysis could also help to bring the performance of the scheme closer to that of variational methods (Li *et al.* 2014). All these refinements could also produce sharper forecasts with less ensemble spread and better accuracy, also enabling a preservation of information and relevance at even longer ranges.

REFERENCES

Aubert, J., 2005. Steady zonal flows in spherical shell dynamos, *J. Fluid Mech.,* **542,** 53–67.

Aubert, J., 2013. Flow throughout the Earth's core inverted from geomagnetic observations and numerical dynamo models, *Geophys. J. Int.,* **192**(2), 537–556.

Aubert, J., 2014a. Earth's core internal dynamics 1840–2010 imaged by inverse geodynamo modelling, *Geophys. J. Int.,* **197,** 1321–1334.

Aubert, J., 2014b. Erratum to: Earth's core internal dynamics 1840–2010 imaged by inverse geodynamo modelling (vol. 197, p. 1321, 2014), *Geophys. J. Int.,* **198**(3), 1683.

Aubert, J. & Fournier, A., 2011. Inferring internal properties of Earth's core dynamics and their evolution from surface observations and a numerical geodynamo model, *Nonlinear Processes Geophys.,* **18**(5), 657–674.

Aubert, J., Aurnou, J. & Wicht, J., 2008. The magnetic structure of convection-driven numerical dynamos, *Geophys. J. Int.,* **172,** 945–956.

Aubert, J., Finlay, C.C. & Fournier, A., 2013. Bottom-up control of geomagnetic secular variation by the Earth's inner core, *Nature,* **502,** 219–223.

Buffett, B., 2014. Geomagnetic fluctuations reveal stable stratification at the top of the Earth's core, *Nature,* **507**(7493), 484–487.

Christensen, U.R. & Aubert, J., 2006. Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields, *Geophys. J. Int.,* **166,** 97–114.

Christensen, U.R., Aubert, J. & Hulot, G., 2010. Conditions for Earth-like geodynamo models, *Earth planet. Sci. Lett.,* **296**(3–4), 487–496.

Christensen, U.R., Wardinski, I. & Lesur, V., 2012. Timescales of geomagnetic secular acceleration in satellite field models and geodynamo models, *Geophys. J. Int.,* **190**(1), 243–254.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte-carlo methods to forecast error statistics, *J. geophys.. Res.,* **99**(C5), 10 143–10 162.

Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation, *Ocean Dyn.,* **53,** 343–367.

Finlay, C.C., Olsen, N. & Toffner-Clausen, L., 2015. DTU candidate field models for IGRF-12 and the CHAOS-5 geomagnetic field model, *Earth planets Space,* **67,** 114, doi:10.1186/s40623-015-0274-3.

Finlay, C.C., Dumberry, M., Chulliat, A. & Pais, M.A., 2010a. Short timescale core dynamics: theory and observations, *Space Sci. Rev.,* **155** (1–4), 177–218.

Finlay, C.C. *et al.*, 2010b. International Geomagnetic Reference Field: the eleventh generation, *Geophys. J. Int.,* **183**(3), 1216–1230.

Finlay, C.C., Jackson, A., Gillet, N. & Olsen, N., 2012. Core surface magnetic field evolution 2000–2010, *Geophys. J. Int.,* **189**(2), 761–781.

Fournier, A., Aubert, J. & Thébault, E., 2015. A candidate secular variation model for IGRF-12 based on Swarm data and inverse geodynamo modelling, *Earth planets Space,* **67,** 81, doi:10.1186/s40623-015-0245-8.

Fournier, A. *et al.*, 2010, An introduction to data assimilation and predictability in geomagnetism, *Space Sci. Rev.,* **155**(1–4), 247–291.

Fournier, A., Aubert, J. & Thébault, E., 2011. Inference on core surface flow from observations and 3-D dynamo modelling, *Geophys. J. Int.,* **186**(1), 118–136.

Fournier, A., Nerger, L. & Aubert, J., 2013. An ensemble Kalman filter for the time-dependent analysis of the geomagnetic field, *Geochem. Geophys. Geosyst.,* **14**(10), 4035–4043.

Gillet, N., Jault, D. & Finlay, C.C., 2015. Planetary gyre, time-dependent eddies, torsional waves and equatorial jets at the Earth's core surface, *J. geophys. Res.,* **120,** 3991–4013.

Gillet, N., Pais, M.A. & Jault, D., 2009. Ensemble inversion of time-dependent core flow models, *Geochem. Geophys. Geosyst.,* **10,** Q06004, doi:10.1029/2008GC002290.

Gillet, N., Jault, D., Canet, E. & Fournier, A., 2010. Fast torsional waves and strong magnetic field within the Earth's core, *Nature,* **465**(7294), 74–77.

Gillet, N., Jault, D., Finlay, C.C. & Olsen, N., 2013. Stochastic modeling of the Earth's magnetic field: inversion for covariances over the observatory era, *Geochem. Geophys. Geosyst.,* **14**(4), 766–786.

Hamilton, B., Ridley, V., Beggan, C.D. & Macmillan, S., 2015. The BGS magnetic field candidate models for the 12th generation IGRF, *Earth planets Space,* **67,** 69, doi:10.1186/s40623-015-0227-x.

Heirtzler, J., 2002. The future of the South Atlantic anomaly and implications for radiation damage in space, *J. Atmos. Sol.-Terr. Phys.,* **64**(16), 1701–1708.

Hulot, G., Lhuillier, F. & Aubert, J., 2010. Earth's dynamo limit of predictability, *Geophys. Res. Lett.,* **37,** L06305, doi:10.1029/2009GL041869.

Jackson, A. & Finlay, C.C., 2007. Geomagnetic secular variation and its applications to the core, in *Treatise on Geophysics,* Vol. 5: *Geomagnetism,* pp. 147–193, ed. Schubert, G., Elsevier.

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

Kuang, W., Wei, Z., Holme, R. & Tangborn, A., 2010. Prediction of geomagnetic field with data assimilation: a candidate secular variation model for IGRF-11, *Earth planets Space,* **62,** 775–785.

Lhuillier, F., Fournier, A., Hulot, G. & Aubert, J., 2011. The geomagnetic secular-variation timescale in observations and numerical dynamo models, *Geophys. Res. Lett.,* **38,** L09306, doi:10.1029/2011GL047356.

Li, K., Jackson, A. & Livermore, P.W., 2014. Variational data assimilation for a forced, inertia-free magnetohydrodynamic dynamo model, *Geophys. J. Int.,* **199**(3), 1662–1676.

Olsen, N. & Mandea, M., 2008. Rapidly changing flows in the Earth's core, *Nature Geosci.,* **1**(6), 390–394.

Olson, P. & Amit, H., 2006. Changes in Earth's dipole, *Naturwissenschaften,* **93**(11), 519–542.

Pais, M.A. & Jault, D., 2008. Quasi-geostrophic flows responsible for the secular variation of the Earth's magnetic field, *Geophys. J. Int.,* **173**(2), 421–443.

Roberts, P.H., Yu, Z.J. & Russell, C.T., 2007. On the 60-year signal from the core, *Geophys. astrophys. Fluid. Dyn.,* **101**(1), 11–35.

Thébault, E. *et al.*, 2015. International Geomagnetic Reference Field: the twelfth generation, *Earth planets Space,* **67,** 79, doi:10.1186/s40623-015-0228-9.

Whaler, K.A. & Beggan, C.D., 2015. Derivation and use of core surface flows for forecasting secular variation, *J. geophys. Res.,* **120**(3), 1400–1414.