Title

Understanding and Predicting Geomagnetic Secular Variation via Data Assimilation

ISBN

9781009180405

Authors

Kuang, Weijia

Gwirtz, Kyle

Tangborn, Andrew

et al.

Publication Date

2023-06-30

DOI

10.1017/9781009180412.022

21

Understanding and Predicting Geomagnetic Secular Variation via Data Assimilation

Weijia Kuang, Kyle Gwirtz, Andrew Tangborn, and Matthias Morzfeld

Abstract: Geomagnetic data assimilation is a recently established research discipline in geomagnetism. It aims to optimally combine geomagnetic observations and numerical geodynamo models to better estimate the dynamic state of the Earth's outer core, and to predict geomagnetic secular variation. Over the past decade, rapid advances have been made in geomagnetic data assimilation on various fronts by several research groups around the globe, such as using geomagnetic data assimilation to understand and interpret the observed geomagnetic secular variation, estimating part of the core state that is not observable on the Earth's surface, and making geomagnetic forecasts on multi-year time scales. In parallel, efforts have also been made on proxy systems for understanding fundamental statistical properties of geomagnetic data assimilation, and for developing algorithms tailored specifically for geomagnetic data assimilation. In this chapter, we provide a comprehensive overview of these advances, as well as some of the immediate challenges of geomagnetic data assimilation, and possible solutions and pathways to move forward.

21.1 Introduction

The observed Earth's magnetic field, measured on the ground and in orbits, is the sum of contributions from various magnetic sources, both within the Earth and external to it. Among them is the field originating in the Earth's fluid outer core, which accounts for more than 99% of the observed magnetic energy as described by the Mausberger–Lowes power spectra (Langel and Estes, 1982). This part of the field, called the 'core field' and the 'geomagnetic field' interchangeably in this chapter, is generated and maintained by turbulent convection in the outer core via dynamo action (Larmor, 1919), and was first modelled numerically a quarter century ago (e.g. Glatzmaier and Roberts, 1995; Kageyama and Sato, 1997; Kuang and Bloxham, 1997). Thus, geomagnetic observations and geodynamo simulations are powerful tools for understanding the dynamical processes in the Earth's outer core, the thermo-chemical properties in the deep Earth, and the interactions between the outer core and other components of the Earth system.

Both geomagnetic observations and geodynamo models can provide independent glimpses of the core dynamic state. Surface and orbital measurements can determine the geomagnetic field up to degree $L_{obs} \leq 14$ in spherical harmonic expansion (Langel and Estes, 1982), and its slow-time variation, called the secular variation (SV), at higher degrees with the data collected from the current Swarm satellite constellations (e.g. Finlay et al., 2020; Sabaka et al., 2020). The observed SV can be used to infer core flow beneath the core–mantle boundary (CMB) via the 'frozen flux' approximation (Roberts and Scott, 1965) and additional constraints on the core flow properties (e.g. Holme, 2007). The observed field and the inferred flow provide (observational) pieces of the core dynamics puzzle (Schaeffer et al., 2016; Aubert and Finlay, 2019; Kloss and Finlay, 2019), but the dominant part of the core state remains opaque.

On the other hand, numerical geodynamo simulations can provide self-consistent approximations of the core dynamic state, by numerically solving the dynamo equations with given boundary and initial conditions. These dynamo equations are the non-linear partial differential equations derived from first principles with various simplifications (Braginsky and Roberts, 1995). Due to computational constraints, numerical geodynamo simulations could not be made with arbitrarily high spatial temporal resolutions, and are thus limited to the parameter regimes far from those appropriate to the Earth's outer core (e.g. Roberts and King, 2013; Wicht and Sanchez, 2019). But estimates of the core state are still attempted by the asymptotic limits (scaling laws) derived from numerical simulations with (computationally permitted) broad ranges of parameter values (e.g. Christensen, 2010; Yadav et al., 2013; Aubert et al., 2017; Kuang et al., 2017). Still the differences between the 'true' core state and the numerical asymptotic limits remain uncertain.

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

The capabilities and limitations of geomagnetic observations and of geodynamo simulations led to the birth and growth of geomagnetic data assimilation (GDA) for making optimal estimates of the core dynamic state, and interpreting and predicting SV, using available geomagnetic data and numerical geodynamo models (Fournier et al., 2010). The first 'proof-of-concept' studies were carried out using simplified magnetohydrodynamic (MHD) systems (Fournier et al., 2007; Sun et al., 2007; Morzfeld and Chorin, 2012), or numerical dynamo models (Liu et al., 2007). These ignited subsequent efforts in GDA for understanding observational constraints on core dynamics (e.g. Kuang et al., 2009; Aubert and Fournier, 2011; Fournier et al., 2011; Aubert, 2014; Kuang and Tangborn, 2015); GDA error statistics and model developments (e.g. Kuang et al., 2008; Canet et al., 2009; Hulot et al., 2010; Fournier et al., 2011; Li et al., 2011, 2014; Sun and Kuang, 2015; Tangborn and Kuang, 2015, 2018; Sanchez et al., 2016, 2019; Gwirtz et al., 2021); and geomagnetic predictions (Kuang et al., 2010; Fournier et al., 2015, 2021b; Morzfeld et al., 2017; Minami et al., 2020; Sanchez et al., 2020; Tangborn et al., 2021). Despite these advancements, many fundamental questions still remain on geodynamic approximations, assimilation algorithms and model/observation error statistics, such as model and observation bias corrections, forecast covariance matrix convergences, and non-linear assimilation algorithms. However, it is expected that GDA will continue to rapidly advance and will be broadly used in geodynamo and geomagnetic field modelling, and in studies of the Earth's deep interior.

The goal of this chapter is to present the reader with an overview of GDA which is easy to comprehend and can provide a first step into geomagnetic data assimilation research. This chapter is organised as follows: the geomagnetic field and geodynamo modelling are reviewed in Section 21.2; the assimilation algorithm is given in Section 21.3. The current research results are presented in Section 21.4, followed by the challenges and future development in Section 21.5. Conclusions and Discussions are in Section 21.6.

21.2 Geomagnetic Field and Geodynamo Modelling

The world (global) magnetic maps can be constructed from observatory and satellite magnetic measurements, historical navigation data, and archeo- and paleomagnetic data (e.g. Jackson et al., 2000; Lesur et al., 2010; Panovska et al., 2019; Finlay et al., 2020; Huder et al., 2020; Sabaka et al., 2020; Alken et al., 2021b; Brown et al., 2021). For details of geomagnetic observations and field models, see, for example, Mandea and Korte (2011) and Sanchez (this volume). Despite subtle differences in the algorithms utilised to produce these field models, they all share the same objective: optimally separating the different sources that contribute to the magnetic measurements. Among those contributions is the magnetic field originated from the Earth's core, called the Earth's intrinsic magnetic field or simply the core field. In this section, this field is also called the geomagnetic field, and the corresponding models are called geomagnetic field models. These models provide descriptions of the spatial and temporal variations of the modern field, as well as lower-accuracy descriptions going as far as 100k years back in time. In a geomagnetic field model, the observed geomagnetic field $\mathbf{B}^{(o)}$ is approximated as a potential field and is described mathematically by the following spherical harmonic expansion

$$\mathbf{B}^{(o)} = -\nabla V,$$
$$V = a \sum_{0 \le m \le l}^{L_{obs}} \left(\frac{a}{r}\right)^{l+1} (g_l^m \cos m\phi + h_l^m \sin m\phi) P_l^m(\theta), \qquad (21.1)$$

where (g_l^m, h_l^m) are called the Gauss coefficients; P_l^m are the Schmidt normalised associated Legendre polynomials of degree l and order m; L_{obs} is the highest degree resolved with the data; a is the mean radius of the Earth's surface; θ and φ are the co-latitude and longitude, respectively. The highest degree L_{obs} in (21.1) depends on the quality of the data available, and thus varies over time. By (21.1), $\mathbf{B}^{(o)}$ can be continued downward from the surface to any location r^* in the interior, as long as the region $r^* \le r \le a$ is electrically insulating. In a GDA system, r^* is typically the mean radius r_c of the CMB if the entire mantle is assumed electrically insulating in the geodynamo model; or, if there is an electrically conducting D″-layer at the base of the mantle, it is the mean radius r_d of the top of the layer. For example, the top row of Fig. 21.1 are the mean observed radial component B_r and its SV $\dot{B}_r$ B_r in 2010–15 at r_d = 3,520 km (i.e. assuming a 20 km thick D″-layer). Notice that B_r reverses its sign in the areas around the tip of South America and south of Africa, which coincide with the South Atlantic Anomaly (SAA), a region with exceptionally low field intensity (see Fig. 21.2).

The working of the geodynamo can be described simply as follows. The secular cooling and differentiation through the Earth's evolution have provided the buoyancy force which drives convection in the outer core. Since the core fluid is an iron-rich liquid alloy and is therefore highly electrically conducting (e.g. Nimmo, 2007; Hirose et al., 2013), an additional magnetic field is generated by the core convection given any seed (background) magnetic field. A self-consistent dynamo is achieved if the generated magnetic field can be maintained without the presence of the seed field (Larmor, 1919). Braginsky and Roberts (1995) provide the full set of the partial differential equations for the geodynamo. But the earliest dynamo models, first by Glatzmaier and Roberts (1995), and later by Kageyama and Sato (1997) and Kuang and Bloxham (1997), were

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

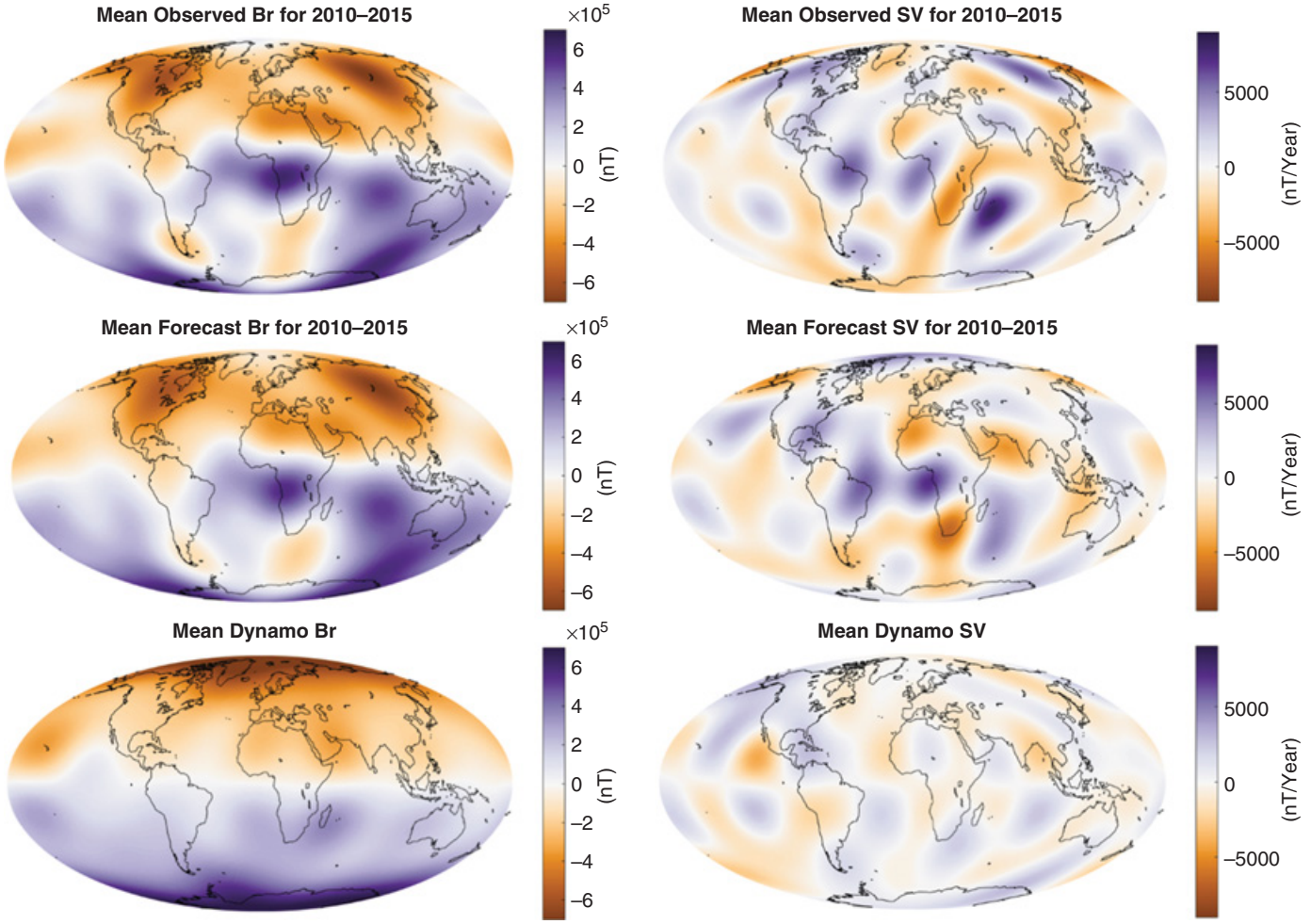

Figure 21.1 Snapshots of the mean radial component B_r of the geomagnetic field (left column) and its mean SV $\dot{B}_r$ (right column) at the top of the D″-layer for 2010–15. The top row are downward continued from that of CM6 field model at the mean surface (Sabaka et al., 2020), the centre row are from NASA GEMS assimilation solutions (Sun and Kuang, 2015; Tangborn et al., 2021), and the bottom row are from MoSST geodynamo simulation solutions (Kuang and Chao, 2003; Jiang and Kuang, 2008).

developed with simplified versions of the Braginsky–Roberts equations. Since then, many more dynamo models have been developed with different physics and/or numerical algorithms implemented. Details of these dynamo models can be found in the past community dynamo benchmark efforts (Christensen et al., 2001; Jones et al., 2011; Matsui et al., 2016).

To better explain the core state defined in geodynamo models and its correlation to geomagnetic observations, we use the MoSST core dynamics model (Kuang and Bloxham, 1999; Kuang and Chao, 2003; Jiang and Kuang, 2008) as an example. Formulations can be easily adapted to other dynamo models. In MoSST, the core fluid is assumed Boussinesq; and the core state is described by the velocity field $\mathbf{v}$, the magnetic field $\mathbf{B}$, and the temperature anomaly Θ (from the background conducting state). Since $\mathbf{v}$ and $\mathbf{B}$ are solenoidal (divergence-free), they are decomposed into the poloidal and the toroidal components, for example,

$$\mathbf{B} = \nabla \times (T_B \hat{\mathbf{r}}) + \nabla \times \nabla \times (P_B \hat{\mathbf{r}}), \tag{21.2}$$

where $\hat{\mathbf{r}}$ is the unit radial vector, and T_B and P_B are the toroidal and poloidal scalars, respectively. In MoSST, all scalar fields are described by spherical harmonic expansions, with the spectral coefficients discretised in radius, for example,

$$P_B = \sum_{0 \le m \le l}^{L_d} b_l^m(r_i)\, Y_l^m(\theta, \phi) + C.C., \text{ for } i = 0, 1, \ldots N_r, \tag{21.3}$$

where L_d is the highest degree, N_r is the number of the radial grid points r_i, Y_l^m are the orthonormal spherical harmonic

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

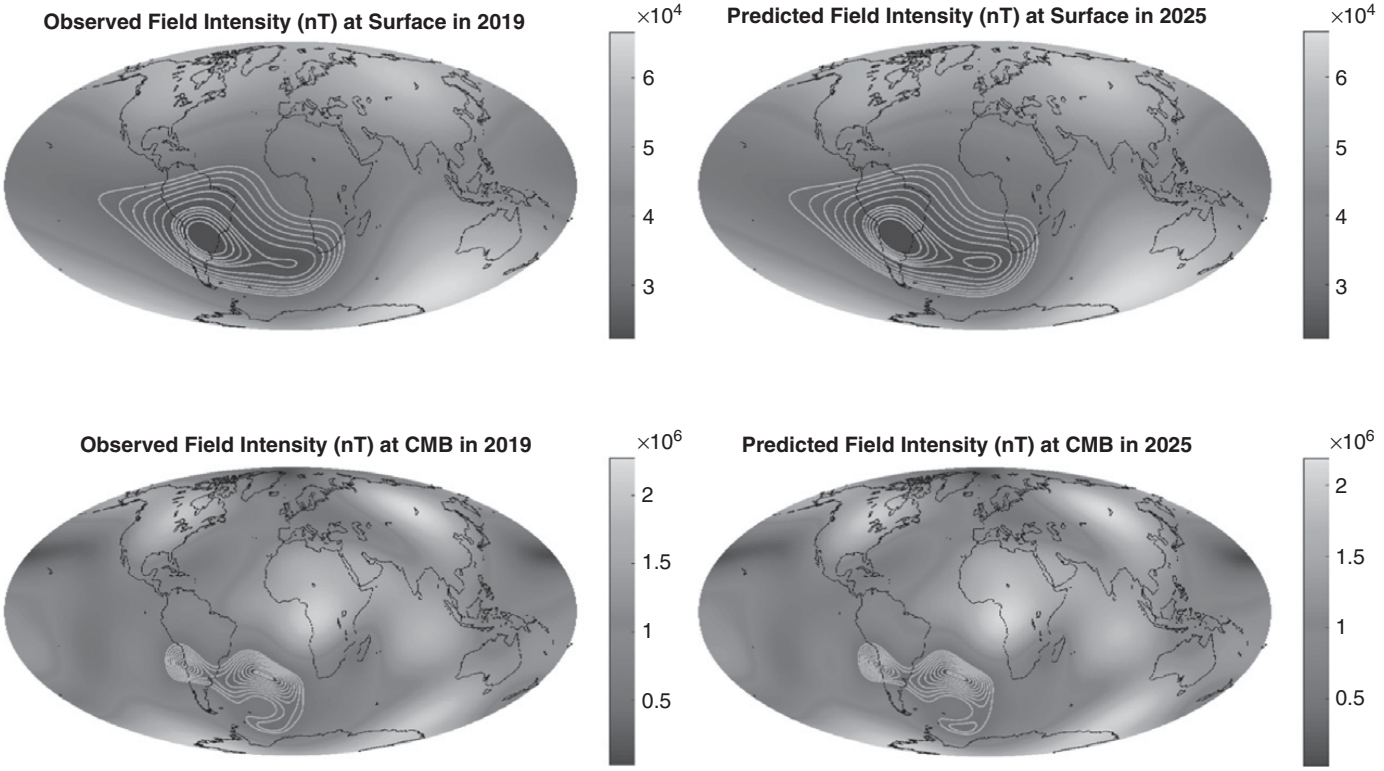

Figure 21.2 The South Atlantic Anomaly observed in 2019 (left column) and predicted in 2025 (right column). The top are the magnetic intensity at the Earth's mean surface, and the bottom are those at the top of the D″-layer. The intensity levels decrease from the outermost contour (30,000 nT on the surface) to the innermost contour (22,800 nT on the surface). The observations are from CM6 (Sabaka et al., 2020), and the forecasts are made by Tangborn et al. (2021).

functions, and C.C. stands for the complex conjugate. With these definitions, the core state vector is

$$\mathbf{x} = [\{b_l^m(r_i)\}, \{j_l^m(r_i)\}, \{v_l^m(r_i)\}, \{\omega_l^m(r_i)\}, \{\zeta_l^m(r_i)\}]^T, \tag{21.4}$$

where j_l^m, v_l^m, w_l^m and ζ_l^m are the complex spectral coefficients of the toroidal magnetic scalar, the poloidal and toroidal velocity scalars, and the temperature anomaly, respectively. Thus the dimension of $\mathbf{x}$ is $N_x = 5(N_r+1)L_d(L_d+2)$, which can be very large, for example, $N_x \approx 10^6$ if $N_r, L_d \sim 100$. The time evolution of $\mathbf{x}$ is evaluated via given discretised schemes, which are of the form

$$\mathbf{x}^{k+1} = \mathbf{L}\mathbf{x}^k + \mathbf{f}(\mathbf{x}^k), \tag{21.5}$$

where $\mathbf{L}$ is the linear matrix describing, for example, the dissipative effects; $\mathbf{f}$ is the vector describing the quadratic interactions among the fields, for example, motional induction; and the superscripts k and $k+1$ indicate the state vector at the time steps t_k and t_{k+1}, respectively. In MoSST, the entire dynamo system is non-dimensionalised, and is therefore defined with a set of non-dimensional parameters, namely the Rayleigh number R_{th} (for the buoyancy force), the magnetic Rossby number R_o (for the fluid inertia), the Ekman number E (for the fluid viscosity), and the modified Prandtl number q_κ (ratio of the thermal conductivity to the magnetic diffusivity). They are referred to as the dynamo parameters in this chapter, and are embedded in $\mathbf{L}$ and $\mathbf{f}$ of (21.5). It should be pointed out that the mathematical description (21.5) applies to all dynamo models, though the state vector and the dynamo parameters may differ. It should also be noted that most of the currently available dynamo models are non-dimensional (e.g. Christensen et al., 2001; Jones et al., 2011; Matsui et al., 2016), thus requiring appropriate rescaling to compare with the observed geomagnetic field.

Geodynamo simulations, which are also called the 'free-running models' in data assimilation as they are not constrained by observations, can provide axial-dipolar–dominant magnetic fields that change in both space and time. But they are also expected to differ from observations, mainly because the dynamo parameter values used in numerical simulation are far from those appropriate for the Earth's core (see Section 21.5 for more discussions). For example, in the bottom row of Fig. 21.1, are the typical mean B_r and mean SV at the top of the D″-layer over a 100-year period from MoSST dynamo simulation results with the parameters $R_o = E = 1.25 \times 10^{-6}$ and $q_\kappa = 1$. Compared to the observations (the top row in Fig. 21.1), the axial dipole component of the simulated field is too

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

strong, and the simulated SV is too weak. These significant differences show that geodynamo simulations cannot provide accurate estimates of the core state on their own.

Aiming at improving core state estimates, one may attempt to constrain geodynamo simulations with observations, and data assimilation is an optimal choice. The observational constraints can be made through connections between the Gauss coefficients defined in (21.1), and the poloidal field spectral coefficients defined in (21.3). In MoSST, an electrically conducting D″-layer at the base of the mantle is implemented in the model. With this feature, the connection is made at the top of the D″-layer $r = r_d$:

$$b_l^m = b_l^{m(o)} \equiv \frac{(-1)^{\mathrm{m}}}{l\mathcal{B}}\left(\frac{a}{r_d}\right)^{l+2}\sqrt{\frac{2\pi(1+\delta_{m0})}{2l+1}}(g_l^m - ih_l^m) \tag{21.6}$$

for $l \le L_{obs}$. In (21.6), δ_{m0} is the Kronecker delta ($\delta_{m0} = 1$ if $m = 0$, and 0 otherwise), and $\mathcal{B}$ is a magnetic scaling factor used to match the dimensional Gauss coefficients (from geomagnetic field models) and the non-dimensional poloidal spectral coefficients (from geodynamo models). The matching condition (21.6) shows also that only part of the poloidal magnetic field can be observed, leaving the rest of the core state unobserved. It also defines the projection of $\mathbf{x}$ onto the observational subspace:

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \tag{21.7}$$

where $\mathbf{y}$ [of the dimension $N_y = L_{obs}(L_{obs} + 2)$] is the observed part of the core state, and $\mathbf{H}$ is an $N_y \times N_x$ matrix, called the observation operator in data assimilation (see Section 21.3). Since the observation $\mathbf{y}$ is defined only at r_d (the top of the D″-layer), $\mathbf{H}$ is very simple: each row of $\mathbf{H}$ contains a single non-zero entry of 1 corresponding to the observed $b_l^{m(o)}$. The relations (21.6) and (21.7) are needed for GDA. It should be noted that the methodology is applicable to other geodynamo models, though the state vector $\mathbf{x}$ and therefore the relation (21.6) may need to be modified accordingly.

21.3 Mathematics of Data Assimilation

In this section, we provide a brief review of data assimilation (DA) while highlighting details relevant to geomagnetism. We begin with a short introduction to the fundamental framework of DA. Specific methods of DA are then outlined and the computational limitations they are subject to in practice are briefly discussed.

21.3.1 The Basic Framework of Data Assimilation

Data assimilation merges a computational model of a process with observations, in order to produce an improved estimate of the state of a system. Methods of DA are typically constructed within a Bayesian framework as follows. Let $\mathbf{x}^t$ be a vector of N_x elements representing the true state of a system at a particular time (the geodynamo in GDA). Knowledge of the system state is recorded in $\mathbf{y}$, a vector of N_y observations which is related to $\mathbf{x}^t$ according to

$$\mathbf{y} = \mathbf{H}\mathbf{x}^t + \boldsymbol{\varepsilon}, \tag{21.8}$$

where $\mathbf{H}$ is the observation operator and $\boldsymbol{\varepsilon}$ is the observation noise which is frequently assumed to be Gaussian with mean zero and covariance $\mathbf{R}$. In GDA, $\mathbf{x}^t$ and $\mathbf{y}$ may consist of spherical harmonic coefficients defining the state of the geodynamo and knowledge of the poloidal magnetic field near the CMB, respectively (see Section 21.2), in which case the observation operator is the $N_y \times N_x$ matrix $\mathbf{H}$. Equation (21.8) defines the likelihood $p(\mathbf{y}|\mathbf{x}^t)$, that is, the probability distribution of the observations given a system state. In sequential DA systems, numerical simulations can be used to define a prior distribution $p_0(\mathbf{x}^t)$(see Section 21.3.2 for details) and by Bayes' rule, a posterior distribution

$$p(\mathbf{x}^t|\mathbf{y}) \propto p_0(\mathbf{x}^t)p(\mathbf{y}|\mathbf{x}^t), \tag{21.9}$$

is defined by the product of the likelihood and the prior. The ultimate objective of various approaches to DA is the approximation of this posterior distribution.

21.3.2 The Ensemble Kalman Filter

The ensemble Kalman filter (EnKF), is a widely used method that has been employed in multiple GDA systems (see, e.g., Fournier et al., 2013; Sun and Kuang, 2015; Sanchez et al., 2020; Tangborn et al., 2021). It approximates the posterior distribution of (21.9) by combining a Monte Carlo approach with the Kalman filter (see, e.g., Evensen, 2006). Specifically, the distributions in Section 21.3.1 are estimated by sampling them through multiple, simultaneous runs of a numerical model. In an EnKF, a *forecast ensemble* of N_e unique forecasts $\mathbf{X}^f = \{\mathbf{x}_1^f, \ldots, \mathbf{x}_{N_e}^f\}$ is produced at a time when observations are to be assimilated. This ensemble is taken to be a sample of the prior distribution $p_0(\mathbf{x}^t)$. The purpose of the EnKF is to adjust these forecasts by merging them with information contained in the observations $\mathbf{y}$. This collection of 'adjusted' forecasts forms an *analysis ensemble* which is taken to be the desired sampling of the posterior distribution $p(\mathbf{x}^t|\mathbf{y})$. Typically, the mean of the analysis ensemble is used as an estimate of the true state of the system, with the ensemble variance indicating the estimate's uncertainty.

An EnKF can be implemented in the following way. The initial collection of forecasts are used to determine the *forecast covariance*

$$\mathbf{P}^f = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}(\mathbf{x}_i^f - \overline{\mathbf{x}})(\mathbf{x}_i^f - \overline{\mathbf{x}})^T, \tag{21.10}$$

where $\overline{\mathbf{x}} = (1/N_e)\Sigma_{i=1}^{N_e}\mathbf{x}_i^f$. An *analysis ensemble* can then be determined by

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}[\mathbf{y} - (\mathbf{H}\mathbf{x}_i^f + \boldsymbol{\varepsilon}_i)], \tag{21.11}$$

for $i = 1, \ldots N_e$, where $\boldsymbol{\varepsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, and

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1} \tag{21.12}$$

is the estimate of the Kalman gain. Under appropriate conditions, the analysis ensemble is a sampling of the posterior and thus provides the desired approximation of (21.9). The analysis ensemble members can then be propagated forward in time by the numerical model, to the next instance when observations are available for assimilation and the process is repeated. The particular EnKF algorithm outlined here is known as the *stochastic EnKF*. Other implementations exist (see, e.g., Tippett et al., 2003; Hunt et al., 2007; Buehner et al., 2017) and differ in their details; however, all rely on a Monte Carlo approximation of the Kalman gain and are designed such that, under certain conditions, the analysis ensemble they produce is distributed according to the posterior.

21.3.3 Variational and Hybrid Methods

Variational methods produce an estimate of the system state by seeking the maximum of the posterior distribution of (21.9). This approach effectively transforms the assimilation of observations to an optimisation problem. Assume that $\mathbf{m}$ is the numerical model for advancing the state vector in time

$$\mathbf{x}(t_k) = \mathbf{m}[\mathbf{x}(t_{k-1})], \tag{21.13}$$

where $\mathbf{x}(t_k)$ is the state vector of a numerical simulation at time t_k. Assume also that at t_k, the observations $\mathbf{y}(t_k)$ are made with error covariance $\mathbf{R}$. For simplicity, we denote $\mathbf{x}_k = \mathbf{x}(t_k)$ and $\mathbf{y}_k = \mathbf{y}(t_k)$ in the rest of the discussion. With this notation, the posterior distribution of the true initial state $\mathbf{x}_0^t$ at time t_0, given observations $\mathbf{y}_1$ at time t_1 is

$$p(\mathbf{x}_0^t|\mathbf{y}_1) \propto p(\mathbf{x}_0^t)p(\mathbf{y}_1|\mathbf{x}_0^t). \tag{21.14}$$

Under the assumption that the prior and likelihood are Gaussian, maximising (21.14) is equivalent to minimising

$$J(\mathbf{x}_0) = (\mathbf{x}_0 - \boldsymbol{\mu})^T\mathbf{B}^{-1}(\mathbf{x}_0 - \boldsymbol{\mu}) + [\mathbf{H}\mathbf{m}(\mathbf{x}_0) - \mathbf{y}_1]^T\mathbf{R}^{-1} [\mathbf{H}\mathbf{m}(\mathbf{x}_0) - \mathbf{y}_1], \tag{21.15}$$

where $\boldsymbol{\mu}$ and $\mathbf{B}$ are the mean and background covariance of the prior distribution, respectively. This particular approach to variational DA is referred to as 4D-Var (see, e.g., Courtier, 1997). Determining the minimiser of the cost function $J(\mathbf{x}_0)$ is an iterative process that can be computationally challenging, in part because optimization requires computation of the gradient of the cost function. Note the dependence of Eq. (21.15) on $\mathbf{m}(\mathbf{x}_0)$, indicating that evaluations of $J(\mathbf{x}_0)$ require runs of the numerical model. Methods for improving the efficiency of the optimisation process are known, however, many require code for a tangent linear model $\mathbf{M}$ of the full numerical model $\mathbf{m}$ (see, e.g., Talagrand and Courtier, 1987), that is, approximating $\mathbf{x}_1$ as $\mathbf{M}\mathbf{x}_0$. But constructing $\mathbf{M}$ for large, non-linear, numerical models can be a significant challenge.

Hybrid techniques have been developed which combine variational and ensemble-based approaches to DA. For example, one may run an ensemble of 4D-Var systems (Bonavita et al., 2012) or couple an EnKF to a 4D-Var setup (Zhang and Zhang, 2012), and use the ensemble mean and covariance to define $\boldsymbol{\mu}$ and $\mathbf{B}$, the statistics of the prior distribution appearing in Eq. (21.15). Recently, for the first time, a hybrid method (4DEnVar) was employed in GDA to propose an SV candidate model for IGRF-13 (Minami et al., 2020).

21.3.4 The Role of Ensemble Size

Under appropriate conditions, approximations of the posterior distribution produced by EnKF implementations or hybrid methods will converge as $N_e \to \infty$. But in practice, one aims to use sufficiently large, but finite ensemble sizes. However, what constitutes a 'sufficiently large' ensemble depends on individual models. It depends, among other things, on the dimension N_x of the state space $\mathbf{x}$, and the quality and extent of the observations $\mathbf{y}$ (Chorin and Morzfeld, 2013). Typically, the larger the state space and the sparser the observations, the larger the ensemble size required. In GDA, a typical 3-D geodynamo model with modest numerical resolutions has a dimension $N_x \sim 10^6$, while observations are limited to $N_y \sim 10^2$ on a 2-D spherical surface (e.g. the outer boundary of a geodynamo model). In many DA applications, the computational expense of the numerical model limits the ensemble size with which it is practical to implement EnKF/hybrid methods. This is the case in GDA where the computational demands of numerical geodynamos have typically restricted ensemble sizes to the hundreds. Efforts towards making GDA with ensembles of limited size more effective are discussed in Sections 21.4.3 and 21.5.

21.4 Geomagnetic Data Assimilation: Current Results

The geomagnetic field varies on timescales ranging from sub-annual to geological time scales, as found from paleomagnetic data (e.g. Panovska et al., 2019), historical magnetic navigation data (e.g. Jackson et al., 2000), and ground observatory and satellite magnetic measurements (e.g. Finlay et al., 2020; Huder et al., 2020; Sabaka et al., 2020). Over the past century, details of small-scale (high spherical harmonic degree) changes in the geomagnetic field morphology have been discovered with observatory and satellite magnetic measurements, such as persistent localised

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

magnetic fluxes (e.g. Jackson, 2003; Finlay and Jackson, 2003); geomagnetic acceleration and geomagnetic jerks (e.g. Mandea et al., 2010; Chulliat and Maus, 2014); high-degree geomagnetic variations (e.g. Hulot et al., 2002; Olsen and Mandea, 2008; Kloss and Finlay, 2019); and the South Atlantic Anomaly (SAA), a localised region of extra low magnetic field intensity at the Earth's mean surface (e.g. Finlay et al., 2020). Historical, archaeo- and paleomagnetic data can provide information on large-scale (low spherical harmonic degree) geomagnetic field variations, such as the strong decay of the axial dipole moment in the past century (e.g. Brown et al., 2018; Jackson et al., 2000), persistent westward drift in the northern hemisphere over millennial time scales (e.g. Nilsson et al., 2020), and the well-known polarity reversals of the geomagnetic field over the last 150 Myr (Cande and Kent, 1995; Lowrie and Kent, 2004; Ogg, 2012).

These observations provide the data used in GDA and highlight the primary purposes of GDA: forecasting, hindcasting, and interpreting geomagnetic SV. The GDA-based SV studies provide one of the few windows we have into the dynamics of Earth's deep interior. Additionally, geomagnetic forecasts have broad scientific and societal applications, such as their contributions to time-varying global geomagnetic models which are widely used in various scientific communities (e.g. Doglioni et al., 2016), for navigation and survey applications (e.g. Kaji et al., 2019), and for space exploration (e.g. Heirtzler et al., 2002).

In this section, we provide a brief overview of some of the current results in using GDA to understand and predict SV. The discussion is organised according to the types of the dynamic models used in the assimilation system. We begin with a review of GDA relying on self-consistent 3-D dynamo models, followed by a discussion of alternative physical and statistical models, and finally, simplified models used in, for example, GDA algorithm development and the prediction of very long-term geomagnetic variations.

21.4.1 GDA with Self-consistent Dynamo Models

The first attempt to predict SV through GDA was carried out a little more than a decade ago (Kuang et al., 2010) and made use of a 3-D numerical geodynamo model. The resulting forecast contributed as a candidate SV model to the 11th generation International Geomagnetic Reference Field (IGRF-11, Finlay et al., 2010). While this demonstrated clearly the value of GDA using Gauss coefficients from geomagnetic field models (not direct geomagnetic measurements) and geodynamo models (with non 'Earth-like' dynamo parameters, see Section 21.5), it was also limited in estimating model uncertainties and biases. For example, the forecast errors in the study by Kuang et al. (2010) are approximated by a simple, time-invariant mathematical description (analogous to the Optimal Interpolation scheme in data assimilation), not by the covariance matrix of a forecast ensemble (as in an EnKF scheme). Fournier et al. (2015) continued the effort to make GDA-based SV forecasts as a candidate model for IGRF-12, with a major improvement in utilising the dynamo solution covariances for the model error statistics. However, the covariances in their assimilation were based only on free-running models, not updates from assimilations (Aubert, 2014).

In the most recent IGRF release (IGRF-13, Alken et al., 2021a), of the fourteen candidate SV models included, four are products of GDA systems using 3-D numerical geodynamos (Minami et al., 2020; Sanchez et al., 2020; Fournier et al., 2021b; Tangborn et al., 2021). Each of these systems make use of an ensemble-based method for assimilation, including the first-ever use of a hybrid variational method (Minami et al., 2020). In addition, observations over the past decades (Minami et al., 2020) and longer (Sanchez et al., 2020; Tangborn et al., 2021) were assimilated. These help produce assimilation solutions which are dynamically consistent over time, as suggested by earlier work with sequential GDA systems (Tangborn and Kuang, 2018; Sanchez et al., 2019). For a more recent description of SV forecasts made by GDA using full dynamo models, we refer the reader to, for example, Fournier et al. (2021a).

The quality of geomagnetic data decreases rapidly when looking back in time, and thus may present the ultimate limitation to the forecast accuracy and dynamic consistency of assimilation solutions from sequential data assimilation systems. But the variational approach by Minami et al. (2020) has the potential to reduce such limitations, as it provides an opportunity to improve earlier geomagnetic data with much more accurate satellite magnetic measurements. All GDA systems face astronomical computational expense if they are to assimilate all available geomagnetic and paleomagnetic data (see Section 21.5 for further discussion).

GDA with self-consistent geodynamo models presents an opportunity to predict the future geomagnetic field over several decades – much longer than the five-year IGRF period. For example, Aubert (2015) showed that forecasts using the dynamic models of GDA systems can outperform linear extrapolations on decadal time scales. This is possible because the estimated 'unobserved' part of the core state is utilised by the dynamic models to predict the future magnetic field. Such forecasts currently predict a continuation of the weakening of the axial dipole, and the expansion and weakening of the SAA over the coming 50–100 years (Aubert, 2015; Sanchez et al., 2020). In particular, as shown in Fig. 21.2, the forecasts predict that a second minimum will grow to split the SAA region in 2025. This highlights a significant contribution of GDA besides simply forecasting the magnetic field: GDA systems are powerful tools for obtaining insight into Earth's deep interior and the origins of SV.

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

21.4.2 GDA with Alternative Dynamic Models

While the use of full numerical dynamo models in GDA has grown over the last decade, approaches employing alternative models based on various physical, mathematical, and statistical descriptions of the magnetic field continue to be successfully developed and employed. These models are, in general, much simpler than the full dynamo models in both mathematical formulation and numerical simulation. However, they have unique advantages in, for example, allowing for assimilation runs on million-year time scales that would be prohibitively expensive with GDA systems using a full dynamo model. Alternative models can also be useful in testing and validating specific physical and/or mathematical approximations which are applicable to geodynamo models. Such models can be particularly useful as testbeds for advancing GDA methods.

The first such approach to GDA with an alternative model is the geomagnetic forecasts made with magnetic induction via core surface flows (e.g. Maus et al., 2008; Beggan and Whaler, 2009, 2010). The flows used for the induction are also derived from geomagnetic observations. It has been shown that a steady core flow model is capable of producing accurate SV forecasts over five-year periods, and are, in principle, improved 'linear extrapolations' of the SV from current and past geomagnetic observations. These results agree also with studies of Aubert (2015).

Improvement could be made if time-varying core flows are considered, such as the quasi-geostrophic flows in the studies of Pais and Jault (2008), Canet et al. (2009) and Aubert (2015), and the 'inertia-free' core flow of Li et al. (2014). In the former approach, the full dynamo model is replaced by a quasi-geostrophic flow which can be determined without the Lorentz force (i.e. the magnetic field in the outer core). In the latter, the momentum equation becomes 'diagnostic', making it equally numerically stable for both forecast and hindcast, a welcome simplification in, for example, a variational data assimilation approach. In summary, these models allow one to avoid solving the full dynamo equations for GDA, greatly reducing computational demands.

Another approach relies on purely statistical models, such as those of Barrois et al. (2018) and Bärenzung et al. (2020). While these models may not capture the detailed physics of a numerical dynamo, it makes the consideration of multiple magnetic sources outside of the core (e.g. the lithosphere and magnetosphere), computationally tractable, which, in turn, allows for the direct assimilation of geomagnetic measurements (as opposed to geomagnetic field models). Computationally less demanding dynamic models can also make ensemble sizes attainable which would otherwise be impractical with numerical dynamos, resulting in more reliable uncertainty estimates. For example, the statistical model of Bärenzung et al. (2018) permits an EnKF with an ensemble size of N_e = 40,000, nearly two orders of magnitude larger than those used in EnKFs with numerical dynamos.

From long-term archeo- and paleomagnetic data, one can find some persistent global scale magnetic features over very long time periods (e.g. Amit et al., 2011; Constable et al., 2016). Thus, an interesting approach is to use low-dimensional models based on either stochastic PDEs (see, e.g., Morzfeld and Buffett, 2019; Pétrélis et al., 2009) or deterministic ODEs (Gissinger, 2012) for long-term SV, such as the behaviour of the axial dipole component of Earth's magnetic field, and for the occurrence of reversals. These models have been used to investigate the predictability of reversals (see, e.g., Gwirtz et al., 2020) including an effort involving the assimilation of paleomagnetic data (Morzfeld et al., 2017). Results from the latter work indicate that assimilations with simplified models of the axial dipole may be useful for anticipating reversals within a window of a few millennia. Limited paleomagnetic data however, makes the validation of reversal prediction strategies a challenge. But it is expected that, as more paleomagnetic data and better low-degree models become available, GDA will also become a powerful tool for predicting very slow secular variations and geomagnetic reversals.

21.4.3 Proxy Models for GDA Development

Simplified models have also played a role in the development of GDA outside of directly being used to make predictions about the Earth. The earliest works concerning GDA involved demonstrating its viability through observing system simulation experiments (OSSEs) with simplified MHD systems (Fournier et al., 2007; Sun et al., 2007; Morzfeld and Chorin, 2012). The dynamic models of those works consisted of 1-D scalar fields intended to represent the magnetic field and fluid velocity of the outer core. While these 'proxy models' were significantly simpler than the geodynamo, they enabled extensive numerical studies of some of the challenges of GDA. This approach of using proxy models has been widely used in the successful development of DA in other applications, such as numerical weather prediction and oceanography. Surprisingly, the pursuit of proxy models for GDA was discontinued until a two-dimensional geomagnetic proxy system (TGPS) recently developed by Gwirtz et al. (2021) was used to study assimilation strategies for use with numerical dynamos.

The TGPS is a magnetoconvection system consisting of 2-D magnetic and velocity fields, on either a plane or spherical surface, which are non-linearly coupled. The right side of Fig. 21.3 shows a snapshot of a solution to the TGPS in a spherical geometry.

The magnetic field and the velocity field are defined by the scalar fields A and ω, respectively (shading in the image of the fields), permitting a complete description of the system state in spherical harmonics, similar to geodynamo models such as MoSST (see Section 21.2). The TGPS was designed to mimic

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

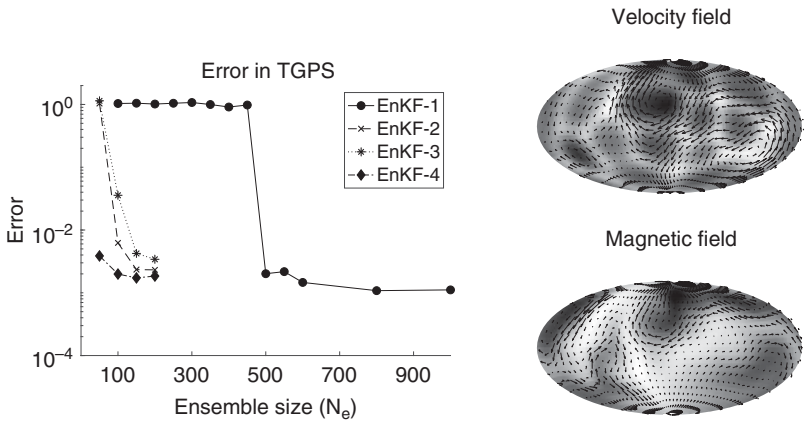

Figure 21.3 A 2-D proxy model for studying GDA. *On the left* are average forecast errors during OSSEs as a function of ensemble size, for the TGPS in a square geometry when using EnKFs with various modifications (labelled 1–4). The errors are scaled by the average forecast error of a free-running model assimilating no data. *On the top right* is a snapshot of the 2-D velocity field (vectors) and the normal component of vorticity (ω, shading) for the TGPS on a sphere. *On the bottom right* is a snapshot of the 2-D magnetic field (vectors) and the magnetic scalar (A, shading) for the TGPS on a sphere.

the scenario where spherical harmonics, determined from geomagnetic field models, are assimilated into numerical geodynamos. This is accomplished through OSSEs in which only noisy 'observations' of the large length-scale spectral coefficients of A, which defines the magnetic field of the TGPS, are assimilated. As with proxy models for other DA applications, a major advantage of the TGPS is that its computational simplicity allows for a large number of OSSEs which would otherwise be impractical with numerical geodynamos.

Extensive OSSEs with the TGPS have been used to explore and propose assimilation strategies for improving accuracy and reducing the computational demands of operational GDA systems. The left panel of Fig. 21.3 shows the average forecast errors with the TGPS during various OSSEs, as a fraction of the average forecast error of a free-running model using no assimilations. The OSSEs differ only in the particular details of the EnKFs (labelled 1–4) and the size of the ensemble (horizontal axis). The curve determined by the black circles (EnKF-1) relies on the standard stochastic EnKF described in Section 21.3.2, while the others (EnKF 2–4) use various modifications generally known as *localisation* and *inflation* (see, e.g., Kotsuki et al., 2017; Shlyaeva et al., 2019). It can be seen that while the unmodified EnKF-1 requires an ensemble of $N_e = 500$ to reduce forecast errors, similar results can be achieved at reduced ensemble sizes when the modified EnKFs (2–4) are used. These findings support the recent implementation of 'localised' EnKFs in GDA (Sanchez et al., 2019, 2020) and suggest additional EnKF modifications which might be useful in reducing the necessary ensemble size, and therefore computational demands, of GDA systems.

21.5 Geomagnetic Data Assimilation: Challenges and Developments

GDA has advanced greatly in the past decade by utilising knowledge accumulated in other Earth sciences, in particular in numerical weather prediction (NWP), and has been recognised as a unique tool for geomagnetic forecasting and core-state estimation. But many challenges still remain in areas ranging from observations and physics to mathematical and computational techniques. Future progress in GDA will rely on overcoming these hurdles. Some could be addressed by leveraging knowledge from, for example, NWP; but many others are unique to geomagnetism, the geodynamo, and core dynamics. Among these challenges are GDA system spin-up given the limited availability of high resolution geomagnetic observations from the past; the astronomical computational requirements of GDA systems; and the systematic errors (model biases) arising from large gaps between the dynamo parameter values used in numerical dynamo simulations and those appropriate to the Earth's core.

A major challenge in GDA is the differences between the observed geomagnetic field and the magnetic field from dynamo simulations. EnKF-type assimilation algorithms such as (21.11) assume that forecast errors are random with the zero mean, that is, no systematic error (bias). It also requires that the observations $\mathbf{y}$ and the forecasts $\mathbf{x}^f$ in (21.11) are defined in the same units. But both are difficult to implement in current GDA systems. First, model biases exist because of the large parameter gaps between the parameter values used in numerical dynamo simulation and those appropriate for the Earth's outer core. The dynamo parameters described in Section 21.2, such as the magnetic Rossby number R_o, the Ekman number E, and the modified Prandtl number q_κ, are very small: $R_o \sim 10^{-9}$, $E \sim 10^{-15}$, and $q_\kappa \sim 10^{-6}$ in the outer core if the molecular fluid viscosity, magnetic diffusivity and thermal conductivity are used (e.g. Braginsky and Roberts, 1995). In numerical simulations, they are at least two orders of magnitude larger (e.g. Wicht and Sanchez, 2019), simply due to computational limitations. Numerical dynamo solutions with such large parameter gaps certainly differ from the (unknown) true core dynamic state. This is particularly significant in dynamical processes that are directly related to these small parameters, such as torsional oscillations (waves) in the outer

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

core. These waves are excited to leading order by the balance between the fluid inertia and the Lorentz force (e.g. Braginsky, 1970; Wicht and Christensen, 2010). Thus their typical frequencies are $\sim\mathcal{O}(R_o^{-1/2})$, implying that two orders of magnitude differences in R_o result in one order of magnitude differences in the wave frequencies (and thus the time scales). Since these waves are conjectured to play a major role in explaining the observed sub-decadal geomagnetic SV (e.g. Bloxham et al., 2002; Cox et al., 2016; Aubert and Finlay, 2019), the impacts of the large parameter gaps must be properly addressed for accurate geomagnetic forecasts. Although continuous efforts are made to narrow the parameter gap (Aubert et al., 2017; Aubert and Gillet, 2021), numerical dynamos which are practical for GDA will continue to use parameters which differ significantly from those representative of the Earth in near future.

Model biases also arise from uncertainties in the thermo-chemical properties of the deep Earth, such as the adiabatic and the total heat fluxes across the CMB (e.g. Nimmo, 2007; Nakagawa, 2020), and the heterogeneity in the inner core and the lower mantle (e.g. Garnero, 2000; Deuss, 2014). The former directly affects the Rayleigh numbers R_{th}, and the latter affects the boundary conditions at the CMB and the inner core boundary (ICB) for numerical dynamo models. It is expected that the parameter gaps will be narrowed and the thermo-chemical uncertainties will be reduced in the coming decades, but will not vanish. The model biases in GDA systems will therefore remain in the foreseeable future.

One possible approach is to estimate the model biases with the asymptotic limits (scaling laws) derived from systematic numerical dynamo simulations with wide ranges of parameter values (e.g. Christensen, 2010; Yadav et al., 2013; Kuang et al., 2017; Petitdemange, 2018). But this could be very difficult, since the numerical asymptotes may not agree with the core state, and since the computational needs for acquiring such asymptotes could be comparable or even higher than those for GDA runs with large ensembles.

Another approach is to rescale $\mathbf{y}$ (or equivalently $\mathbf{x}^f$) based on the properties of the observed field and of the dynamo model used in GDA (e.g. Kuang et al., 2010; Aubert, 2014; Fournier et al., 2015, 2021b; Tangborn et al., 2021). This is perhaps more pragmatic since such rescaling is needed if non-dimensional numerical dynamo models are used in GDA, and since the canonical scaling rules employed in dynamo modelling are inappropriate due to non-'Earth-like' dynamo parameters in simulation. In the approach of Kuang et al. (2010) and Tangborn et al. (2021), the numerical and the observed axial dipole moments are used for the magnetic field rescaling, but the time rescaling remains the same as the canonical time scale of the numerical dynamo model. In the approach of Aubert (2014) and Fournier et al. (2015), the typical time scales of the numerical dynamo solutions and of the observed SV are used for the time rescaling, but the magnetic rescaling relies on the asymptotic properties derived from independent numerical dynamo simulations. As such, both approaches may lead to inconsistencies between the magnetic rescaling and the time rescaling since, as shown in Fig. 21.4, the typical time scales and the typical intensities of the numerical magnetic fields vary with the dynamo parameters. An immediate development could be to select both the magnetic rescaling and the time rescaling to match the observed and the modelled field intensities and time scales. Consistencies of the scalings and their potential improvements may be tested and validated with various OSSEs.

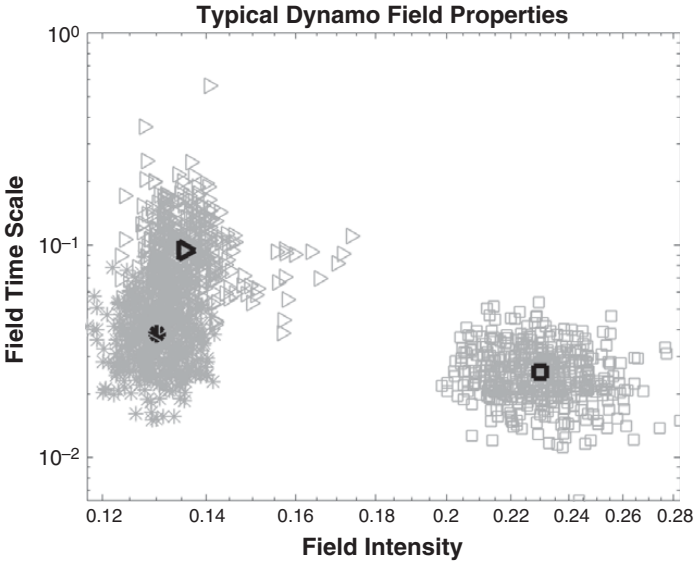

Figure 21.4 The typical intensity (x-axis) and the typical time scale (y-axis) of the ensembles of numerical poloidal magnetic fields with (1) $R_{th} = 1811$, $R_o = E = 1.25 \times 10^{-6}$ (squares); (2) $R_{th} = 1811$, $R_o = E = 6.25 \times 10^{-7}$ (stars); (3) $R_{th} = 905$, $R_o = E = 6.25 \times 10^{-7}$ (triangles). The dark bold-face symbols are the ensemble mean values. The intensities are for the poloidal magnetic field at the top of D″-layer for spherical harmonic degrees $l \le 13$. Each of the three ensembles consist of 512 snapshots of numerical dynamos selected from large pools of well-developed free-running dynamo solutions obtained with the MoSST core dynamics model (e.g. Jiang and Kuang, 2008).

The GDA forecast spin-up has always been a concern because, as discussed in Section 21.2, the geomagnetic observations are very sparse (the dimension of $\mathbf{y}$ is more than four orders of magnitudes less than that of $\mathbf{x}^f$), and are not 'in-situ' as in many other Earth systems ($\mathbf{y}$ is determined by downward continuing the surface observations to the outer boundary of the dynamo system). The spin-up can be measured by the time evolution of the forecast accuracy $(\text{O-F})^B \equiv \mathbf{y} - \mathbf{H}\mathbf{x}^f$: for an EnKF GDA system, it is expected to decrease in time as more data are assimilated, until it reaches some minimum level (thus the system is fully spun-up). Reaching the minimum level is critical for minimising the SV forecast error which, by definition, can be determined by the field forecast errors at different times:

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

$$(\text{O-F})^{SV} \equiv \dot{\mathbf{y}} - \mathbf{H}\dot{\mathbf{x}}^{f} = \frac{1}{\Delta t}[(\text{O-F})^{B}_{t+\Delta t} - (\text{O-F})^{B}_{t}]. \qquad (21.16)$$

However, previous studies suggest that $(\text{O-F})^{B}$ may not always decay monotonically in time (e.g. Tangborn and Kuang, 2015, 2018). Though SV forecasts can be improved by some additional calibration of $(\text{O-F})^{B}$ (Kuang et al., 2010; Tangborn et al., 2021), or by careful selection of the initial ensembles of the inverse dynamo solutions matching earlier geomagnetic observations (Aubert, 2014; Fournier et al., 2015), better solutions could be found by improved assimilation algorithms, such as the rescaling approach discussed in the previous paragraph, and by assimilating other geodynamic observables, such as the length of day variation on decadal time scales which is likely due to angular momentum exchange between the solid Earth and the fluid core (e.g. Jault et al., 1988). It should be noted that the angular momentum $\mathbf{M}_{oc}$ of the outer core (relative to the solid mantle) is the volume integral $\mathbf{M}_{oc} = \int \rho_{oc}\mathbf{r} \times \mathbf{v} dV$ (where ρ_{oc} is the core fluid density).

Development and validation of new assimilation algorithms needed to address the issues outlined here will add to the already very challenging computational needs of GDA. Take a Boussinesq dynamo model for an illustration. As shown in Section 21.2, there are five independent scalar fields in the state vector $\mathbf{x}$ of the dimension $5N_1N_2N_3$ (where $N_{1,2,3}$ are the numerical resolutions in a 3-D dynamo domain). In the MoSST core dynamics model (e.g. Kuang and Bloxham, 1999; Jiang and Kuang, 2008), for example, the resolution is defined by the radial grid points N_x and the truncation order L_x of the spherical harmonic expansions. Thus, in one time step, there are $\sim\mathcal{O}(50N_xL_x^3 \ln L_x)$ floating point operations (with spherical harmonic transforms). For a modest resolution $N_x = L_x = \mathcal{O}(10^2)$ and time step $\Delta t \sim 10^{-6}$(typical for dynamo simulations with $R_o, E \sim 10^{-6}$) a total of $\sim 10^{16}$ floating point operations are needed for a dynamo simulation over a magnetic free-decay time. This amounts to ~ 3 h on a tera floating point operations per second (teraflops) computing system (excluding the communication time across computing nodes). The computing needs will be N_e-fold more for GDA runs using ensembles of size N_e. Therefore, there is a need for more efficient assimilation algorithms to make the GDA computing needs bearable, such as those aiming at reducing the necessary ensemble size for EnKF based GDA systems (e.g. Sanchez et al., 2019; Gwirtz et al., 2021). Development of proxy models, such as those of Canet et al. (2009) and Gwirtz et al. (2021), are of particular importance for advancing GDA, as they can provide dynamically complex, but computationally economical platforms for at least early stage proof-of-concept studies of assimilation algorithms and physical approximations.

21.6 Discussion

In this chapter, we have provided an overview of geomagnetic data assimilation (GDA), including some basics of geomagnetic observations, geodynamo models, and assimilation methodologies. We have also presented a wide range of GDA results in understanding core dynamical processes, interpreting observed SV, and geomagnetic forecasting. In addition, we have elaborated on some of the challenges in GDA and possible pathways to move forward. As such, this chapter serves as a quick and comprehensive introduction for those who wish to learn GDA and/or work on GDA-related research and applications.

We would like to point out a particular useful application of the proxy models described in Section 21.4. Since these models are mathematically simple and computationally affordable, they are very handy for teaching/learning GDA. Compared to any full geodynamo model, these models are easy to analyse. In particular, simulation and assimilation (with these models) can be completed quickly on desktops and laptops, thus making them ideal in, for example, student projects.

While we have made an effort to include representative GDA results and developments, this chapter does not cover all GDA activities, in part due to the page limit and the rapid development of GDA in recent years. For example, the description of variational geomagnetic data assimilation is very brief in this chapter, and we refer the reader to relevant references for more details. Regardless, this should not affect the main purpose of this chapter, which is to provide a comprehensive understanding of geomagnetic data assimilation.

References

Alken, P., Thébault, E., Beggan, C. D. et al. (2021a). International Geomagnetic Reference Field: The thirteenth generation. *Earth, Planets and Space*, 73, 49. https://doi.org/10.1186/s40623-020-01288-x.

Alken, P., Chulliat, A., and Nair, M. (2021b). NOAA/NCEI and University of Colorado candidate models for IGRF-13. *Earth, Planets and Space*, 73, 44. https://doi.org/10.1186/s40623–020–01313–z.

Amit, H., Korte, M., Aubert, J., Constable, C., and Hulot, G. (2011). The time-dependence of intense archeomagnetic flux patches. *Journal of Geophysical Research: Solid Earth*, 116, B12106. https://doi.org/10.1029/2011JB008538.

Aubert, J. (2014). Earth's core internal dynamics 1840–2010 imaged by inverse geodynamo modelling. *Geophysical Journal International*, 197, 1321–34.

Aubert, J. (2015). Geomagnetic forecasts driven by thermal wind dynamics in the Earth's core. *Geophysical Journal International*, 203(3), 1738–51.

Aubert, J., and Finlay, C. C. (2019). Geomagnetic jerks and rapid hydromagnetic waves focusing at Earth's core surface. *Nature Geoscience*, 12, 393–8.

https://doi.org/10.1017/9781009180412.022 Published online by Cambridge University Press

Aubert, J., and Fournier, A. (2011). Inferring internal properties of Earth's core dynamics and their evolution from surface observations and a numerical geodynamo model. *Nonlinear Processes in Geophysics*, 18, 657–74.

Aubert, J., and Gillet, N. (2021). The interplay of fast waves and slow convection in geodynamo simulations nearing Earth's core conditions. *Geophysical Journal International*, 225(3), 1854–73.

Aubert, J., Gastine, T., and Fournier, A. (2017). Spherical convective dynamos in the rapidly rotating asymptotic regime. *Journal of Fluid Mechanics*, 813, 558–93.

Bärenzung, J., Holschneider, M., Wicht, J., Lesur, V., and Sanchez, S. (2020). The Kalmag model as a candidate for IGRF-13. *Earth, Planets and Space*, 72(163).

Bärenzung, J., Holschneider, M., Wicht, J. Sanchez, S., and Lesur, V. (2018). Modeling and predicting the short-term evolution of the geomagnetic field. *Journal of Geophysical Research: Solid Earth*, 123(6), 4539–60.

Barrois, O., Hammer, M. D., Finlay, C. C., Martin, Y., and Gillet, N. (2018). Assimilation of ground and satellite magnetic measurements: inference of core surface magnetic and velocity field changes. *Geophysical Journal International*, 215 (1), 695–712.

Beggan, C. D., and 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.

Beggan, C. D., and Whaler, K. A. (2010). Forecasting secular variation using core flows. *Earth, Planets and Space*, 62, 821–28.

Bloxham, J, Zatman, S., and Dumburry, M. (2002). The origin of geomagnetic jerks. *Nature*, 420, 65–68.

Bonavita, M., Isaksen, L., and Hólm, E. (2012). On the use of EDA background error variances in the ECMWF 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, 138(667), 1540–59.

Braginsky, S. I. (1970). Torsional magnetohydrodynamic vibrations in the Earth's core and variation in day length. *Geomagnetism and Aeronomy*, 10, 1–8.

Braginsky, S. I., and Roberts, P. H. (1995). Equations governing convection in Earth's core and the geodynamo. *Geophysical and Astrophysical Fluid Dynamics*, 79, 1–97.

Brown, M., Korte, M., Holme, R., Wardinski, I., and Gunnarson, S. (2018). Earth's magnetic field is probably not reversing. *PNAS*, 115, 5111–16.

Brown, W. J., Beggan, C. D., Cox, G. A., and Macmillan, S. (2021). The BGS candidate models for IGRF-13 with a retrospective analysis of IGRF-12 secular variation forecasts. *Earth, Planets and Space*, 73 (42). https://doi.org/10.1186/s40623–020–01301–3.

Buehner, M., McTaggart-Cowan, R., and Heilliette, S. (2017). An Ensemble Kalman filter for numerical weather prediction based on variational data assimilation: VarEnKF. *Monthly Weather Review*, 145(2), 617–35.

Cande, S. C., and Kent, D. V. (1995). Revised calibration of the geomagnetic polarity timescale for the late Cretaceous and Cenozoic. *Journal of Geophysical Research*, 100, 6093–6095.

Canet, E., Fournier, A., and Jault, D. (2009). Forward and adjoint quasigeostrophyic models of geomagnetic secular variations. *Journal of Geophysical Research*, 114, B11101.

Chorin, A. J., and Morzfeld, M. (2013). Conditions for successful data assimilation. *Journal of Geophysical Research: Atmospheres*, 118 (20), 11522–33.

Christensen, U. R. (2010). Dynamo scaling laws and applications to the planets. *Space, Science, Reviews*, 152, 565–90.

Christensen, U.R., Aubert, J., Cardin, P. et al. (2001). A numerical dynamo benchmark. *Physics of the Earth and Planetary Interiors*, 128(1), 25–34.

Chulliat, A., and Maus, S. (2014). Geomagnetic secular acceleration, jerks, and localized standing wave at the core surface from 2000 to 2010. *Journal of Geophysical Research: Solid Earth*, 119, 1531–43.

Constable, C., Korte, M., and Panovska, S. (2016). Persistent high paleosecular variation activity in southern hemisphere for at least 10000 years. *Earth and Planetary Science Letters*, 453, 78–86.

Courtier, P. (1997). Variational methods. *Journal of the Meteorological Society of Japan*, 75(1B), 211–18.

Cox, G. A., Livermore, P. W., and Mound, J. E. (2016). The observational signature of modelled torsional waves and comparison to geomagnetic jerks. *Physics of the Earth and Planetary Interiors*, 255, 50–65.

Deuss, A. (2014). Heterogeneity and anisotropy of Earth's inner core. *Annual Review of Earth and Planetary Sciences*, 42, 103–26.

Doglioni, C., Pignatti, J., and Coleman, M. (2016). Why did life develop on the surface of the Earth in the Cambrian? *Geoscience Frontiers*, 7, 865–75.

Evensen, G. (2006). *Data assimilation: The ensemble Kalman filter*. Springer.

Finlay, C. C., and Jackson, A. (2003). Equatorially dominated magnetic field change at the surface of the Earth's core. *Science*, 300, 2084–6.

Finlay, C. C., Maus, S., Beggan, C. D. et al. (2010). International Geomagnetic Reference Field: The eleventh generation. *Geophysical Journal International*, 183(3), 1216–30.

Finlay, C. C., Kloss, C., Olsen, N. et al. (2020). The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly. *Earth, Planets and Space*, 72, 156. https://doi.org/10.1186/s40623-020-01252-9.

Fournier, A., Eymin, C., and Alboussier, T. (2007). A case for variational geomagnetic data assimilation: Insights from a one-dimensional, nonlinear, and sparsely observed MHD system. *Nonlinear Processes in Geophysics*, 14, 163–80.

Fournier, A., Aubert, J., and Thébault, E. (2011). Inference on core surface flow from observations and 3-D dynamo modelling. *Geophysical Journal International*, 186, 118–36.

Fournier, A., Nerger, L., and Aubert, J. (2013). An ensemble Kalman filter for the time-dependent analysis of the geomagnetic field. *Geochemistry, Geophysics, Geosystems*, 14(10), 4035–43. https://doi.org/10.1002/ggge.20252.

Fournier, A., Aubert, J., and Thébaut, E. (2015). A candidate secular variation model for IGRF-12 based on Swarm data and inverse geodynamo modeling. *Earth, Planets and Space*, 67. https://doi.org/10.1186/s40623–015–0245–8.

Fournier, A., Aubert, J., Lesur, V., and Thébault, E. (2021a). Physics-based secular variation candidate models for the IGRF. *Earth, Planets and Space*. https://doi.org/10.1186/s40623–021–01507–z.

Fournier, A., Aubert, J., Lesur, V., and Ropp, G. (2021b). A secular variation candidate model for IGRF-13 based on Swarm data and ensemble inverse geodynamo modeling. *Earth, Planets and Space*. https://doi.org/10.1186/s40623–020–01309–9.

Fournier, A. G., Hulot, G., Jault, D. et al. (2010). An introduction to data assimilation and predictability in geomagnetism. *Space Science Reviews*. https://doi.org/10.1007/s11214–010–9669–4.

Garnero, E. J. (2000). Heterogeneity of the lowermost mantle. *Annual Review of Earth and Planetary Sciences*, 28, 509–37.

Gissinger, C. (2012). A new deterministic model for chaotic reversals. *European Physical Journal B*, 85, 137.

Glatzmaier, G. A., and Roberts, P. H. (1995). A three-dimensional convective dynamo solution with rotating and finitely conducting inner core and mantle. *Physics of the Earth and Planetary Interiors*, 91, 63–75.

Gwirtz, K, Morzfeld, M, Fournier, A, and Hulot, G. (2020). Can one use Earth's magnetic axial dipole field intensity to predict reversals? *Geophysical Journal International*, 225(1), 277–97.

Gwirtz, K., Morzfeld, M., Kuang, W., and Tangborn, A. (2021). A testbed for geomagnetic data assimilation. *Geophysical Journal International*, 227, 2180–203.

Heirtzler, J. R., Allen, J. H., and Wilkinson, D. C. (2002). Everpresent South Atlantic Anomaly damages spacecraft. *Eos, Transactions American Geophysical Union*, 83(15), 165–9.

Hirose, K., Labrosse, S., and Hernlund, J. (2013). Composition and state of the core. *Annual Review of Earth and Planetary Sciences*, 41(1), 657–91.

Holme, R. (2007). Large-scale flow in the core. In P. Olson, ed., *Treatise on Geophysics: Vol. 8. Core Dynamics*. Amsterdam: Elsevier, pp. 107–30.

Huder, L., Gillet, N., Finlay, C. C., Hammer, M. D., and Tchoungui, H. (2020). COV-OBS.x2: 180 years of geomagnetic field evolution from ground-based and satellite observations. *Earth, Planets and Space*, 72(160). https://doi.org/10.1186/s40623–020–01194–2.

Hulot, G., Eymin, C., Langlais, B., Mandea, M., and Olsen, N. (2002). Smallscale structure of the geodynamo inferred from Oested and Magsat satellite data. *Nature*, 416, 620–3. https://doi.org/10.1038/416620a.

Hulot, G., Lhuillier, F., and Aubert, J. (2010). Earth's dynamo limit of predictability. *Geophysical Research Letters*, 37(6). https://doi.org/10.1029/2009GL041869.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, 230(1), 112–26.

Jackson, A. (2003). Intense equatorial flux spots on the surface of the Earth's core. *Nature*, 424, 760–63.

Jackson, A., Jonkers, A. R. T., and Walker, M. R. (2000). Four centuries of geomagnetic secular variation from historical records. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 358(1768), 957–90.

Jault, D., Gire, C., and LeMouël, J.-L. (1988). Westward drift, core motions and exchanges of angular momentum between core and mantle. *Nature*, 333, 353–6.

Jiang, W., and Kuang, W. (2008). An MPI-based MoSST core dynamics model. *Physics of the Earth and Planetary Interiors*, 170(1), 46–51.

Jones, C. A., Boronski, P., Brun, A. et al. (2011). Anelastic convection-driven dynamo benchmarks. *Icarus*, 216, 120–35.

Kageyama, A., and Sato, T. (1997). Generation mechanism of a dipole field by a magnetohydrodynamic dynamo. *Physical Review E*, 55, 4617–26.

Kaji, C. V., Hoover, R. C., and Ragi, S. (2019). Underwater navigation using geomagnetic field variations. *2019 IEEE International Conference on Electro Information Technology*. https://doi.org/10.1109/EIT.2019.8834192 of:

Kloss, C, and Finlay, C. C. (2019). Time-dependent low-latitude core flow and geomagnetic field acceleration pulses. *Geophysical Journal International*, 217(1), 140–68.

Kotsuki, S., Ota, Y., and Miyoshi, T. (2017). Adaptive covariance relaxation methods for ensemble data assimilation: experiments in the real atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 143(705), 2001–15.

Kuang, W., and Bloxham, J. (1997). An Earth-like numerical dynamo model. *Nature*, 389, 371–4.

Kuang, W., and Bloxham, J. (1999). Numerical Modeling of Magnetohydrodynamic Convection in a Rapidly Rotating Spherical Shell: Weak and Strong Field Dynamo Action. *J. Comput. Phys.*, 153(1), 51–81.

Kuang, W., and Chao, B. F. (2003). Geodynamo Modeling and Core-Mantle Interactions. In V. Dehant, K. Creager, S. Karato, and S. Zatman, eds., *Earth's Core: Dynamics, Structure, Rotation, Geodynamics Series 31*. Washington, DC: American Geophysical Union (AGU), pp. 193–212.

Kuang, W., and Tangborn, A. (2015). Dynamic responses of the Earth's outer core to assimilation of observed geomagnetic secular variation. *Progress in Earth and Planetary Science*, 2. https://doi.org/10.1186/s40645–015–0071–4.

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

Kuang, W., Tangborn, A., Wei, Z., and Sabaka, T. J. (2009). Constraining a numerical geodynamo model with 100 years of surface observations. *Geophysical Journal International*, 179(3), 1458–68, https://doi.org/10.1111/j.1365–246X.2009.04376.x.

Kuang, W., Wei, Z., Holme, R., and Tangborn, A. (2010). Prediction of geomagnetic field with data assimilation: a candidate secular variation model for IGRF-11. *Earth, Planets and Space*, 62, 775–85.

Kuang, W., Chao, B. F., and Chen, J. (2017). Decadal polar motion of the Earth excited by the convective outer core from geodynamo simulations. *Journal of Geophysical Research: Solid Earth*, 122(10), 8459–73.

Langel, R. A., and Estes, R. H. (1982). A geomagnetic field spectrum. *Geophysical Research Letters*, 9, 250–3.

Larmor, J. (1919). How could a rotating body such as the Sun become a magnet? *Reports of the British Association*, 87, 159–60.

Lesur, V., Wardinski, I., Hamoudi, M., and Rother, M. (2010). The second generation of the GFZ internal magnetic model: GRIMM-2. *Earth, Planets and Space*, 62, 765–73.

Li, K., Jackson, A., and Livermore, P. W. (2011). Variational data assimilation for the initial-value dynamo problem. *Physical Review E*, 84. https://doi.org/10.1103/PhysRevE.84.056321.

Li, K., Jackson, A., and Livermore, P. W. (2014). Variational data assimilation for a forced, inertia-free magnetohydrodynamic dynamo model. *Geophysical Journal International*, 199, 1662–76.

Liu, D., Tangborn, A., and Kuang, W. (2007). Observing system simulation experiments in geomagnetic data assimilation. *Journal of Geophysical Research*, 112. https://doi.org/10.1029/2006JB004691.

Lowrie, W., and Kent, D. V. (2004). Geomagnetic polarity time scale and reversal frequency regimes. *Timescales of the paleomagnetic field*, 145, 117–29.

Mandea, M., and Korte, M., eds. (2011). *Geomagnetic Observations and Models*. Dordrecht: Springer.

Mandea, M., Holme, R., Pais, A. et al. (2010). Geomagnetic jerks: Rapid core field variations and core dynamics. *Space Science Reviews*, 155, 147–75.

Matsui, H., Heien, E., Aubert, J. (2016). Performance benchmarks for a next generation numerical dynamo model. *Geochemistry, Geophysics, Geosystems*, 17(5), 1586–607.

Maus, S., Silva, L., and Hulot, G. (2008). Can core-surface flow models be used to improve the forecast of the Earth's main magnetic field? *Journal of Geophysical Research*, 113, B08102. https://doi.org/10.1029/2007JB005199.

Minami, T., Nakano, S., Lesur, V. et al. (2020). A candidate secular variation model for IGRF-13 based on MHD dynamo simulation and 4DEnVar data assimilation. *Earth, Planets and Space*, 72, 136. https://doi.org/10.1186/s40623–020–01253–8.

Morzfeld, M., and Buffett, B. A. (2019). A comprehensive model for the kyr and Myr timescales of Earth's axial magnetic dipole field. *Nonlinear Processes in Geophysics*, 26(3), 123–42.

Morzfeld, M., and Chorin, A. J. (2012). Implicit particle filtering for models with partial noise, and an application to geomagnetic data assimilation. *Nonlinear Processes in Geophysics*, 19 (3), 365–82.

Morzfeld, M., Fournier, A., and Hulot, G. (2017). Coarse predictions of dipole reversals by low-dimensional modeling and data assimilation. *Physics of the Earth and Planetary Interiors*, 262, 8–27.

Nakagawa, T. (2020). A coupled core-mantle evolution: review and future prospects. *Progress in Earth and Planetary Science*, 7. https://doi.org/10.1186/s40645–020–00374–8.

Nilsson, A., Suttie, N., Korte, M., Holme, R., and Hill, M. (2020). Persistent westward drift of the geomagnetic field at the core-mantle boundary linked to recurrent high-latitude weak/reverse flux patches. *Geophysical Journal International*, 222, 1423–32.

Nimmo, F. (2007). Energetics of the core. In P. Olson, ed., *Treatise on Geophysics: Vol. 8. Core Dynamics*. Amsterdam: Elsevier, pp. 31–66.

Ogg, J. G. (2012). Geomagnetic polarity time scale. In F. M. Gradstein, J. G. Ogg, M. Schmitz, and G. Ogg, eds., *The Geologic Time Scale 2012*. Amsterdam: Elsevier Science, pp. 85–113.

Olsen, N., and Mandea, M. (2008). Rapidly changing flows in the Earth's core. *Nature Geoscience*, 1, 390–94.

Pais, M. A., and Jault, D. (2008). Quasi-geostrophyic flows responsible for the secular variation of the Earth's magnetic field. *Geophysical Journal International*, 173, 421–43.

Panovska, S., Korte, M., and Constable, C. G. (2019). One hundred thousand years of geomagnetic field evolution. *Reviews of Geophysics*, 57(4), 1289–337.

Petitdemange, L. (2018). Systematic parameter study of dynamo bifurcations in geodynamo simulations. *Physics of the Earth and Planetary Interiors*, 277, 113–32.

Pétrélis, F., Fauve, S., Dormy, E., and Valet, J.-P. (2009). Simple mechanism for reversals of Earth's magnetic field. *Physical Review Letters*, 102, 144503.

Roberts, P. H., and Scott, S. (1965). On analysis of the secular variation. *Journal of Geomagnetism and Geoelectricity*, 17, 137–51.

Roberts, P. H, and King, E. M. (2013). On the genesis of the Earth's magnetism. *Reports on Progress in Physics*, 76(9), 096801.

Sabaka, T. J., Tøffner-Clausen, L., Olsen, N., and Finlay, C. C. (2020). CM6: A comprehensive geomagnetic field model derived from both CHAMP and Swarm satellite observations. *Earth, Planets and Space*, 72, 80.

Sanchez, S., Fournier, A., Aubert, J., Cosme, E., and Gallet, Y. (2016). Modeling the archaeomagnetic field under spatial constraints from dynamo simulations: A resolution analysis. *Geophysical Journal International*, 207, 983–1002.

Sanchez, S., Wicht, J., Bärenzung, J., and Holschneider, M. (2019). Sequential assimilation of geomagnetic observations: Perspectives for the reconstruction and prediction of core dynamics. *Geophysical Journal International*, 217, 1434–50.

Sanchez, S., Wicht, J., Bärenzung, J., and Holschneider, M. (2020). Predictions of the geomagnetic secular variation based on the ensemble sequential assimilation of geomagnetic field models by dynamo simulations. *Earth, Planets and Space*, 72, 157. https://doi.org/10.1186/s40623–020–01279–y.

Schaeffer, N., Lora Silva, E., and Pais, M. A. (2016). Can core flows inferred from geomagnetic field models explain the Earth's dynamo? *Geophysical Journal International*, 204(2), 868–77.

Shlyaeva, A., Whitaker, J. S., and Snyder, C. (2019). Model-space localization in serial ensemble filters. *Journal of Advances in Modeling Earth Systems*, 11(6), 1627–36.

Sun, Z., and Kuang, W. (2015). An ensemble algorithm based component for geomagnetic data assimilation. *Terrestrial Atmospheric and Oceanic Sciences*, 26, 53–61.

Sun, Z., Tangborn, A., and Kuang, W. (2007). Data assimilation in a sparsely observed one-dimensional modeled MHD system. *Nonlinear Processes in Geophysics*, 14, 181–92.

Talagrand, O., and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quarterly Journal of the Royal Meteorological Society*, 113(478), 1311–28.

Tangborn, A., and Kuang, W. (2015). Geodynamo model and error parameter estimation using geomagnetic data assimilation. *Geophysical Journal International*, 200, 664–75.

Tangborn, A., and Kuang, W. (2018). Impact of archeomagnetic field model data on modern era geomagnetic forecasts. *Physics of the Earth and Planetary Interiors*, 276, 2–9.

Tangborn, A., Kuang, W., Sabaka, T. J., and Yi, C. (2021). Geomagnetic secular variation forecast using the NASA GEMS ensemble Kalman filter: A candidate SV model for IGRF-13. *Earth, Planets and Space*, 73, 47. https://doi.org/10.1186/s40623–020–01324–w.

Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S. (2003). Ensemble square root filters. *Monthly Weather Review*, 131, 1485–90.

Wicht, J., and Christensen, U. R. (2010). Torsional oscillations in dynamo simulations. *Geophysical Journal International*, 181, 1367–80.

Wicht, J., and Sanchez, S. (2019). Advances in geodynamo modelling. *Geophysical & Astrophysical Fluid Dynamics*, 113(1–2), 2–50.

Yadav, R. K., Gastine, T., and Christensen, U. R. (2013). Scaling laws in spherical shell dynamos with free-slip boundaries. *Icarus*, 225, 185–93.

Zhang, M., and Zhang, F. (2012). E4DVar: Coupling an ensemble Kalman filter with four-dimensional variational data assimilation in a limited-area weather prediction model. *Monthly Weather Review*, 140(2), 587–600.