

FULL PAPER

Open Access



# Geomagnetic core field models and secular variation forecasts for the 13th International Geomagnetic Reference Field (IGRF-13)

I. Wardinski<sup>1\*</sup> , D. Saturnino<sup>2</sup>, H. Amit<sup>2</sup>, A. Chambodut<sup>1</sup>, B. Langlais<sup>2</sup>, M. Mandaia<sup>3</sup> and E. Thébaud<sup>4</sup>

## Abstract

Observations of the geomagnetic field taken at Earth's surface and at satellite altitude are combined to construct continuous models of the geomagnetic field and its secular variation from 1957 to 2020. From these parent models, we derive candidate main field models for the epochs 2015 and 2020 to the 13th generation of the International Geomagnetic Reference Field (IGRF). The secular variation candidate model for the period 2020–2025 is derived from a forecast of the secular variation in 2022.5, which results from a multi-variate singular spectrum analysis of the secular variation from 1957 to 2020.

**Keywords:** The geomagnetic field, Geomagnetic secular variation, Geomagnetic field models, Forecasts of the geomagnetic field

## Introduction

The International Association of Geomagnetism and Aeronomy (IAGA) regularly releases the International Geomagnetic Reference Field (IGRF), which is a mathematical description of Earth's main magnetic field and its secular variation. Previous versions (e.g., Finlay et al. 2010; Thébaud et al. 2015) are widely used in many disciplines of Earth sciences and applied for navigational purposes (Jiménez et al. 2012; Canciani and Raquet 2016) and in satellite orientation (Slavinskis et al. 2014).

In this study, we combine geomagnetic field observations taken at Earth's surface and at satellite altitude to construct continuous models of the geomagnetic core field and its secular variation between 1957 and 2020. From these models, candidate models for the IGRF (Alken et al. 2020), i.e., main field models for the epochs 2015 and 2020 and a secular variation model for the period 2020 to 2025 centered at 2022.5 are derived. We apply two modeling techniques to derive these models.

First, a method that descends from the time-dependent modeling technique developed by Bloxham and Jackson (1992); we refer to this as the classical model. The second technique is based on a method for constructing core field models that satisfy the frozen-flux radial magnetic induction equation on the core-mantle boundary (CMB) by imposing the field evolution to be entirely due to advection of the magnetic field at the core surface (Lesur et al. 2010; Wardinski and Lesur 2012), which we refer to as the kinematic field model. The latter method could be understood as a simple data assimilation approach, where the diffusion-less induction equation and assumptions about the dynamical regime of the core flow form the priors, and observations define their likelihood.

Methods of forecasting the future geomagnetic field evolution range from simple linear extrapolation to data assimilation into numerical dynamo simulations (Kuang et al. 2010; Aubert 2015; Fournier et al. 2015). Here, we devise two strategies to forecast the geomagnetic secular variation. First, a direct forecast based on a multi-variate singular spectrum analysis (MSSA) (Broomhead and King 1986; Plaut and Vautard 1994) of the magnetic field variability of past decades. Second, a kinematic forecast scheme is applied that is also based on the MSSA, but

\*Correspondence: wardinski@unistra.fr

<sup>1</sup> Institut de Physique du Globe de Strasbourg, Université de Strasbourg/EOST, CNRS, UMR 7516, Strasbourg, France

Full list of author information is available at the end of the article

of the core flow variability of past decades. The reconstruction of the past flow variability and its forecast are used to predict the future geomagnetic field by forward-modeling the diffusion-less radial induction equation on the CMB. This approach is somewhat similar to geomagnetic field forecasts using steady and time invariant flows (Beggan and Whaler 2010; Hamilton et al. 2015; Whaler and Beggan 2015). However, such forecasts are expected to fail at the occurrence of geomagnetic jerks that are sudden changes in the secular variation. Such events occurred in the past decades (Mandea et al. 2010; Brown et al. 2013), most recently in 2014 (Torta et al. 2015). The cause of these events is not fully understood. Their occurrences have been related to different types of rapid wave motion within Earth's liquid core (Bloxham et al. 2002; Aubert and Finlay 2019), temporal changes of the core flow (Wardinski et al. 2008) and Earth's rotation variation (Holme and de Viron 2005, 2013).

This paper is organized as follows. “**Geomagnetic field modeling**” section outlines the two techniques to derive the parent geomagnetic field model for the IGRF candidates. In the third section, we develop the methodology to predict future geomagnetic secular variation. “**Results and discussion**” section provides results of the geomagnetic field modeling, secular variation forecasts and the derivation of the candidate models. The last section discusses the results and concludes the study.

### Geomagnetic field modeling

In this section, we summarize the derivation of a parent geomagnetic main field model from which we deduce an IGRF candidate model. The parent model, hereafter C<sup>3</sup>FM3, covers the period from 1957 to 2020. The model derivation follows that of Wardinski and Lesur (2012) and consists of two branches, a classical model without the kinematic constraint applied (see “**Classical modeling**” section), and a kinematic field model based on the tangential geostrophic flow assumption (see “**Kinematic field modeling**” section). Like in the previous model, C<sup>3</sup>FM2, we use order 6 B-splines to parameterize field and flow coefficients in time. The spline knot spacing is set to be roughly 1.5 years. Both model branches are constrained to fit a main field model for the epoch 2015. This main field model is based on magnetic measurements taken by the Swarm satellite mission (Lesur, priv. comm.). We choose 2015, as it is the epoch of the last IGRF, with a good data coverage provided by geomagnetic observatories and the Swarm satellite mission. This data coverage decreases towards the model endpoint.

### Data

In this work, we use two types of data, measurements taken at a network of ground-based geomagnetic observatories and satellite data taken at satellite virtual

geomagnetic observatories (Mandea and Olsen 2006). The idea of combining ground-based and virtual observatory data to perform a geomagnetic field modeling was already carried out by Barrois et al. (2018). However, here we derive secular variation estimates to avoid leakage of sub-annual external field variations into the description of the core field and to obtain a sufficient representation of the short-term secular variation, that may be used to forecast geomagnetic field changes.

### Ground-based geomagnetic observatories

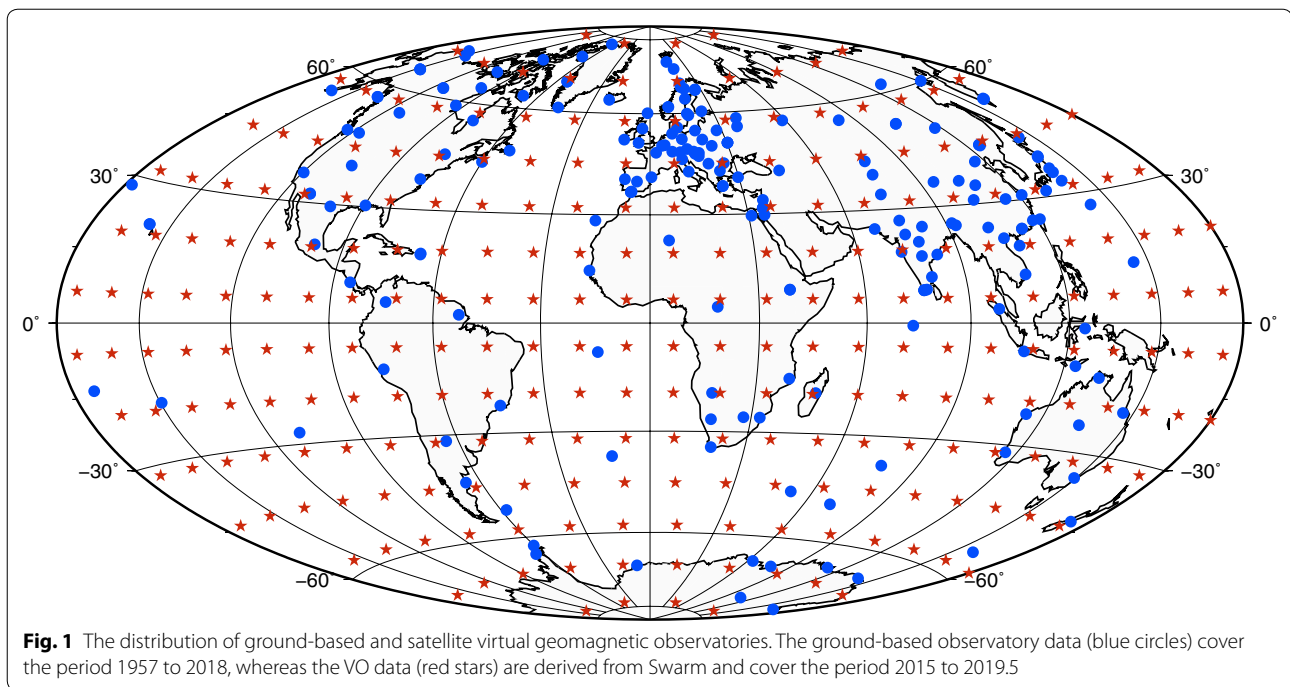
A large portion of the data used in this study comes from ground-based observatories. Like in previous studies (Wardinski and Holme 2006; Wardinski and Lesur 2012; Lesur et al. 2018), we derive estimates of secular variation by annual differences from observatory monthly means, where these monthly means are averages of observatory hourly means. Also, annual means are used for observatories for which hourly mean values are not available from the World Data Centre for Geomagnetism - Edinburgh (2019). These observatory annual means are part of a compilation that is provided by the British Geological Survey - Edinburgh (2020). Over the period 1957–2018 the number of geomagnetic observatories simultaneously in operation that have been providing vectorial hourly means of North, East and downward components ranges between 72 and 155. Data errors were removed when encountered and data gaps were not filled by interpolations. Figure 1 maps locations of the ground-based geomagnetic observatories used in this study.

### Satellite virtual geomagnetic observatories

We use vector magnetic field measurements from Swarm Level-1b data product, version 0505 (0506 for some data files). All three Swarm satellites are considered for the period between January 2014 to June 2019. Data is screened for quality flags defined in the Level-1b Product Definition Document (Tøffner-Clausen and Nielsen 2018). We keep only measurements identified as nominal, and also Swarm C vector measurements after 4th November 2014.

We select only data where the Sun is below the horizon (Chambodut et al. 2002). Additionally, we retain only data showing moderate magnetic activity. Sectorial magnetic activity index  $a\sigma$  (Chambodut et al. 2013), provided by the International Service of Geomagnetic indices (2020), were used and we only select data corresponding to  $a\sigma < 25$  nT.

With the selected data we then construct a global mesh of virtual geomagnetic observatories (VO) following Saturnino et al. (2018), with some small changes. An approximately equal area mesh is obtained with the VO centers separated in latitude by 12.8° and defining 14 latitudinal bands, with



$\theta_{vo} = \pm 6.40^\circ, \pm 19.20^\circ, \pm 32.0^\circ, \dots, \pm 83.20^\circ$ . In each band, the longitude  $\phi_{vo}$  of each VO and the number of longitudinal divisions,  $N_{\phi_{vo}}$  (rounded up to the nearest integer), are chosen so that:

$$N_{\theta_{vo}} = \frac{360}{12.8} \cos \theta_{vo}. \quad (1)$$

The resulting mesh contains 258 VOs. Figure 1 displays the locations of all VOs in the mesh. The data set of each VO consists of selected data acquired inside a cylinder of  $3.0^\circ$  radius centered around each VO and during a 30-day period, i.e., leading to nearly monthly values. The Equivalent Source Dipole (ESD) technique is then used, following closely Saturnino et al. (2018). For each month (30-day period) the ESD inversion is applied to each VO vector data for the equivalent magnetization of dipoles placed at 2900 km depth inside Earth's interior, by a least squares fit in an iterative, conjugate gradient, inversion scheme (Purucker et al. 1996). Then, the forward calculation is used to estimate a magnetic field value at the VO center location and for a given time period. In this way, time series of magnetic field values at the center of each VO and at a constant altitude of 500 km, are obtained.

The distributions of ground-based and virtual observatories differ as can be seen in Fig. 1. Overall, geomagnetic observatory locations cluster, which may lead to a higher spatial resolution of the model in some parts of the world, whereas in other parts the resolution may be lower than that of the virtual observatories.

### Secular variation estimates

We derive secular variation estimates as input for the geomagnetic field modeling. The technique is applied to monthly means of VOs and to ground-based geomagnetic monthly means. The secular variation of the  $X$ -component at a given observatory is estimated as

$$dX/dt|_{\tau} = (X(\tau + 6) - X(\tau - 6))/dt, \quad (2)$$

where  $\tau$  denotes a particular month. These are annual differences of observatory monthly means. Likewise, observatory annual means are treated using

$$dX/dt|_{t+1/2} = (X(t + 1) - X(t))/dt, \quad (3)$$

where  $t$  is in calendar years and  $dt$  is 1 year. Then, secular variation estimates derived by (2) and (3) are given in nT/year.

### Classical modeling

Conventional geomagnetic field modeling approaches rely on the assumption that Earth's magnetic field  $\mathbf{B}(r, \theta, \phi, t)$  is a potential field without magnetic sources in the mantle and in the vicinity of satellite virtual observatories. Because of this, the geomagnetic field is determined as a gradient of a scalar potential, i.e.,

$$\mathbf{B}(r, \theta, \phi, t) = -\nabla V(r, \theta, \phi, t). \quad (4)$$

Then in a spherical geometry, the scalar potential of the geomagnetic field can be represented as

$$V(r, \theta, \phi, t) = a \sum_{l=1}^{l_{\max}} \sum_{m=0}^l (g_l^m(t) \cos(m\phi) + h_l^m(t) \sin(m\phi)) \left(\frac{a}{r}\right)^{l+1} P_l^m(\cos \theta), \quad (5)$$

where  $a$  is Earth's radius (6371.2 km),  $(r, \theta, \phi)$  the geocentric spherical radial, co-latitude and longitude coordinates and  $P_l^m(\cos \theta)$  are the Schmidt quasi-normalized associated Legendre functions, with their normalization defined by

$$\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} (P_l^m(\cos \theta) \cos(m\phi))^2 \sin \theta \, d\theta \, d\phi = \frac{4\pi}{2l+1}, \quad (6)$$

with degree  $l$  and order  $m$ . The maximum spherical harmonic degree in (5) is chosen to be  $l_{\max} = 14$ , to minimize the contamination by the crustal field. The Gauss coefficients  $\{g_l^m, h_l^m\}$  are expanded in time using order six B-splines  $M_n(t)$ :

$$g_l^m(t) = \sum_{n=1}^N g_l^{mn} M_n(t), \quad h_l^m(t) = \sum_{n=1}^N h_l^{mn} M_n(t). \quad (7)$$

The objective function  $\Phi(\mathbf{m})$  to be minimized in the inversion is:

$$\Phi(\mathbf{m}) = (\mathbf{y} - \mathbf{A}\mathbf{m})^T \mathbf{C}_e^{-1} (\mathbf{y} - \mathbf{A}\mathbf{m}) + \mathbf{C}_m^{-1} + \mathbf{C}_F^{-1}, \quad (8)$$

where  $\mathbf{A}$  is an operator which relates the model vector  $\mathbf{m}$  containing the Gauss coefficients to the data  $\mathbf{y}$ .  $\mathbf{y} - \mathbf{A}\mathbf{m}$  is the misfit between data and model, subject to the regularization.  $\mathbf{C}_e$  and  $\mathbf{C}_m$  are the error and the prior model covariance matrix, respectively.  $\mathbf{C}_m$  is an expression of the model priors.

Here, we report solutions that are adjusted to minimize the integral of  $B_r^2$  over the core surface to obtain a spatial smooth model:

$$\int_{S(c)} B_r^2 \, d\Omega = \mathbf{m}^T \mathbf{S}^{-1} \mathbf{m}. \quad (9)$$

The matrix  $\mathbf{S}$  has the diagonal elements

$$s_{ll}, t_{ll} = \frac{(l+1)^2}{2l+1} \left(\frac{a}{c}\right)^{(2l+4)} \quad \text{for } l = 1 \dots, l_{\max}. \quad (10)$$

Like in some previous studies (Lesur et al. 2010; Wardinski and Lesur 2012), we seek a reliable estimate of the secular acceleration. Therefore, the temporal model

constraint is to minimize the integral of the third time derivative of the radial field component over the core surface and in the model period between  $t_S$  and  $t_E$

$$\frac{4\pi}{(t_E - t_S)} \int_{t_S}^{t_E} \int_{S(c)} \left(\frac{\partial^3 B_r}{\partial t^3}\right)^2 \, d\Omega \, dt = \mathbf{m}^T \mathbf{T}^{-1} \mathbf{m}, \quad (11)$$

where  $S(c)$  is the spherical surface of the core at radius  $c = 3485$  km. The diagonal elements of the matrix  $\mathbf{T}$  are the same as for  $\mathbf{S}$ , and the time integral is computed using a Newton–Cotes formula of a closed type, e.g., Bode's rule (Abramowitz and Stegun 1973). Minimization of the third time derivative requires placing further conditions on the second time derivatives of the radial field at the model end-points; best results are obtained when these are set to zero.

The model prior covariance matrix  $\mathbf{C}_m$  is then given by:

$$\mathbf{C}_m^{-1} = \lambda_S \mathbf{m}^T \mathbf{S}^{-1} \mathbf{m} + \lambda_T \mathbf{m}^T \mathbf{T}^{-1} \mathbf{m}. \quad (12)$$

The last term of (8) is the constraint to fit a given satellite field in 2015, which is

$$\begin{aligned} F &:= \int_{r=a} (B - {}^0B)^2 \, dS|_{t=2015} \\ &= \sum_{l=1}^{l_{\max}} \sum_{m=0}^l (l+1) [(g_l^m - {}^0g_l^m)^2 + (h_l^m - {}^0h_l^m)^2]_{t=2015}, \end{aligned} \quad (13)$$

where  ${}^0g_l^m$  and  ${}^0h_l^m$  are the Gauss coefficients of the satellite geomagnetic main field model. This constraint is necessary, as our model is based on secular variation data and it needs a main field model at a given epoch in order to provide also description of the main field at all times. Traditionally, we use a main field model that is derived in an independent study from satellite data. The constraint is then written

$$\mathbf{C}_F^{-1} = \lambda_f \mathbf{m}^T \mathbf{F}^{-1} \mathbf{m}, \quad (14)$$

with the damping parameter  $\lambda_f$ .

Solutions are sought iteratively in a very similar manner as for the previous model, C<sup>3</sup>FM2, by deriving an initial model to re-weight the observatory data by their residuals to this initial model. Then, the strength of external field variation is reduced by a noise-removal scheme (Wardinski and Holme 2011). From this data set the final model is derived.

### Kinematic field modeling

In this section, we describe our method to invert geomagnetic observations for field and flow at the core surface, which extends our previous study (Wardinski and

Lesur 2012), where we imposed the core flow to be purely toroidal. The inversion of secular variation data for field and flow at the core surface formulates to a Bayesian inference. Assuming a Gaussian distribution, then this leads to an objective function similar to (8) that is minimal for the preferred solution  $\mathbf{m}$ .

In the following, we provide details of the prior information (constraints) used to derive a preferred solution of the non-unique and non-linear inverse problem. The constraints are applied on the portion of the model vector that represents the core surface flow.

Here, we use the data set with the reduced external field noise obtained in “Classical modeling” section to jointly invert for the field and flow at the core surface. This eases the joint-inversion process, as it avoids the iterative solving scheme to re-weight the data for the non-linear problem. Different assumptions of the flow dynamic could be applied (Holme 2007). Among them: purely toroidal (Whaler 1980), tangential geostrophic (LeMouél 1984) and quasi-geostrophic flow (Pais and Jault 2008). However, we focus on the tangential geostrophic flow assumption, as it is more comprehensive than a purely toroidal flow, but less restrictive than a quasi-geostrophic flow. (Note that the term flow refers to its horizontal part only, as the radial part vanishes at the core surface).

The flow is decomposed into toroidal and poloidal components:

$$\mathbf{u} = \mathbf{u}_{\text{tor}} + \mathbf{u}_{\text{pol}} = \nabla_h \times (\hat{\mathbf{r}}\mathcal{T}) + \nabla_h \mathcal{P}. \quad (15)$$

$\mathcal{T}$  and  $\mathcal{P}$  are scalars which are expanded in Schmidt-normalized real spherical harmonics in space and B-splines of order 6 in time, represented by  $t_l^m(n)$ ,  $s_l^m(n)$ .

Following Lesur et al. (2010), the objective function of the joint inversion for the field and the flow at the core surface reads

$$\Phi(\mathbf{m}) = (\mathbf{y} - \mathbf{A}\mathbf{m})^T \mathbf{C}_e^{-1} (\mathbf{y} - \mathbf{A}\mathbf{m}) + \lambda_1 \Phi_1(\mathbf{m}), \quad (16)$$

where the model vector  $\mathbf{m}$  now contains the sets of Gauss and flow coefficients. The functional  $\Phi_1$  is related to the kinematic constraint and defined by

$$\Phi_1(\mathbf{m}) = \sum_t (\mathbf{A}_g(t) \cdot \mathbf{u} - \dot{\Psi}(t) \cdot \mathbf{g})^T \mathbf{C}^g (\mathbf{A}_g(t) \cdot \mathbf{u} - \dot{\Psi}(t) \cdot \mathbf{g}). \quad (17)$$

$\dot{\Psi}(t)$  is a design matrix based on the radial induction equation in the kinematic assumption, i.e.,

$$\partial_t B_r = -\nabla_h \cdot (\mathbf{u} B_r). \quad (18)$$

According to (18), the secular variation of the radial field component at the core surface,  $\partial_t B_r$ , can be expressed in terms of a core field,  $B_r$ , advected by the core fluid flow,  $\mathbf{u}$ , where  $\nabla_h$  is the horizontal divergence. Then

$$\dot{\Psi}(t) \cdot \mathbf{g} = \mathbf{A}_g \cdot \mathbf{u}, \quad (19)$$

where  $\mathbf{g}$  is a vector that contains the Gauss coefficients and  $\mathbf{u}$  contains the flow coefficients. The elements of the diagonal weight matrix  $\mathbf{C}^g$  are defined as  $w_l^g = \frac{4\pi(l+1)^2}{(2l+1)}$ . Minimizing the mean square difference, integrated over the core surface  $S(c)$  and time, between the observed secular variation and the secular variation generated by the flow, then the functional  $\Phi_1$  is equivalent to the integral and we can write this (similarly to (9) and (11)) as

$$\int_{\mathcal{T}} \int_{S(c)} (\partial_t B_r(t) + \nabla_h \cdot (\mathbf{u} B_r))^2 d\Omega dt = \mathbf{g}^T \mathbf{K}^{-1} \mathbf{g}, \quad (20)$$

where  $\mathbf{m}$  is now the model vector that holds the flow and Gauss coefficients. The diagonal elements of the field part of  $\mathbf{K}$  are given by  $\mathbf{C}^g$ . The parameter  $\lambda_1$  controls the conformity of the model to the kinematic constraint.

Because  $\mathbf{A}_g(t)$  in the functional  $\Phi_1$  depends on the Gauss coefficients  $g_l^m(t)$ ,  $h_l^m(t)$  and is multiplied by  $\mathbf{u}$ , this optimization problem (inversion) is clearly non-linear and has to be solved iteratively. However, the iterative process is unlikely to converge unless some constraints are applied on the flow model, as finding a flow model is an ill-posed inverse problem (Holme 2007). In order to obtain the optimal field model and simultaneously reduce the null space for the flow, two types of constraints are considered. The flow model is forced to have a convergent spectrum, i.e., to be large scale, and to minimize Bloxham’s “strong norm” (Bloxham 1988; Jackson 1997),

$$\int_{\mathcal{T}} \int_{S(c)} (\nabla_h(\nabla_h \cdot \mathbf{u}))^2 + (\nabla_h(\hat{\mathbf{r}} \times \nabla_h \cdot \mathbf{u}))^2 d\Omega dt = \mathbf{u}^T \mathcal{N} \mathbf{u}. \quad (21)$$

$\mathcal{N}$  has the diagonal elements  $(l^3(l+1)^3)/(2l+1)$ . The damping parameter  $\lambda_S$  controls to what extent the flow follows this constraint. Minimizing (21) constrains efficiently the secular variation. Secondly the flow model is chosen such that it varies smoothly in time,

$$\int_{\mathcal{T}} \int_{S(c)} (\mathbf{u})^2 d\Omega dt = \mathbf{u}^T \mathcal{V} \mathbf{u}, \quad (22)$$

with  $\lambda_T$  as the associated control parameter of the flow temporal evolution that efficiently regularizes the inverse problem. The constraints (21) and (22) are similar to the temporal constraint of the classical model, i.e., (11), as they involve a temporal integration to be minimized. Finally, it is required that the flow acceleration at starting and ending epochs is minimized by

$$\int_{S(c)} \left( \frac{\partial \mathbf{u}}{\partial t} \right)^2 d\Omega \Big|_{t_1, t_2} = \mathbf{u}^\top \mathcal{E} \mathbf{u}. \quad (23)$$

$t_1$  and  $t_2$  are the epochs 1957 and 2020.0, respectively. This becomes necessary, as, if the flow acceleration is unconstrained at the endpoint, it may exceed realistic values. The factor  $\lambda_E$  controls this constraint.

These are the basic settings for the joint inversion for the magnetic core field and the core flow. We impose a further constraint, in order to derive models that are based on different dynamical assumptions of the flow. One possible assumption commonly used is a tangential geostrophic flow (Hills 1979; LeMouél 1984) which is established by minimizing

$$\int_{S(c)} (\nabla_h \cdot (\mathbf{u} \cos \theta))^2 d\Omega = \mathbf{u}^\top \mathcal{G} \mathbf{u}, \quad (24)$$

where  $\theta$  is the co-latitude. Elements of  $\mathcal{G}$  are given by Pais et al. (2004). The constraint is controlled by setting  $\lambda_{TG}$ .

We apply two measures to discriminate models: the rms secular variation misfit is measured by differences between model and the observed secular variation, i.e.,

$$M = \sqrt{\frac{1}{(N_{\text{obs}} - 1)} \sum_{i=1}^{N_{\text{obs}}} (\text{Obs} - \text{Mod})^2} \quad (25)$$

where  $N_{\text{obs}}$  is the number of observations. In addition, for the kinematic field models we derive the ratio

$$R(t) = \frac{\int_{S(c)} \{ \partial_t B_r(t) + \nabla_h \cdot (\mathbf{u} B_r) \}^2 d\Omega}{\int_{S(c)} \{ \partial_t B_r(t) \}^2 d\Omega} \quad (26)$$

that specifies to what extent the frozen-flux radial induction equation is satisfied. Values similar to 1 and larger mean that this constraint is not fulfilled by the model, and conversely values significantly smaller than 1 indicate that the frozen-flux radial induction equation is approximately satisfied.

### Forecasting schemes

In this study, we aim to obtain a description of (a) the temporal dynamics of the secular variation and (b) the temporal variability of the advective motion within the liquid outer core. These should not differ largely, but remaining external signals in the secular variation estimates may pose problems to identify clearly signals from the core. To forecast states of the physical system that lead to the secular variation it is necessary to analyze the observed time series. Our strategy relies on the derivation of multi-variate time-series models, where individual secular variation and flow coefficients are treated as time series.

Our development of time-series models and forecasts is based on the multi-variate singular spectrum analysis (MSSA) (Broomhead and King 1986; Plaut and Vautard 1994, and see “Appendix A” section) of field and core flow variability. Other methods could be chosen to find a sufficient description of the field and flow variability, such as ARIMA-models, or vector auto-regressions models; however, these model types are based on assumptions like stationarity of the time series and normality of the residuals (Box and Jenkins 1976; Brockwell and Davis 2002, and references therein), which are unlikely for the temporal variability of the geomagnetic field (Hulot et al. 2010). Furthermore, the singular spectrum analysis does not rely on assumptions about the linearity or non-linearity of the process that is generating the time series, whereas auto-regression models implicitly assume a linear behavior of the observed data. If  $1, \dots, T$  serves as the time range for a training set, then the model parameters are estimated from these observations. The (out-of-sample) prediction is generated over the time range  $T + 1, \dots, T + h$  according to the generation mechanism of the model. In this study we consider 10 years as the time range of the prediction. This sets  $h = 120$  months. Prediction models are derived using the first 22 eigenmodes obtained by the MSSA of the temporal flow variability (see “Appendix A” section).

### Direct secular variation forecast

The time series of the Gauss coefficients are taken from the classical model branch of C<sup>3</sup>FM3. The analysis considers discrete series of secular variation coefficients

$$y_t := \dot{g}_l^m(n), n = 1, \dots, N \quad (27)$$

consisting in  $N$  observations regularly sampled in time. The sampling time  $\tau_s$  is chosen to be one month.

Time-series models are derived using a multi-variate singular spectrum analysis, where the temporal variability of the series is decomposed in eigenmodes (Vautard et al. 1992; Golyandina et al. 2001; Ghil et al. 2002). The time-series models are then reconstructions based on superposition of these eigenmodes, and similarly forecasts. For the technical details of the model selection, we refer the reader to “Appendix A” section.

### Kinematic secular variation forecast

The kinematic forecasting scheme to predict future secular variation is motivated by results that the observed secular variation can be mostly explained by the advection of magnetic field (Roberts and Scott 1965; Wardinski and Lesur 2012). Additionally, results of a previous kinematic field model show that the spatial complexity imposed by a satellite magnetic field model of 2004 is maintained backward in time over decades (Wardinski and Lesur 2012). Here, we

want to explore to what extent the approach can be used to forecast geomagnetic field changes.

The scheme consists of two steps. First, we derive time-series models that give a sufficient accurate description of the temporal variability of the fluid motion in the liquid outer core. Therefore, we consider toroidal and poloidal flow terms, which are treated as discrete multi-variate time series from which we derive time-series models for a given 'learning' phase. The learning phase is set to cover the interval 1957 to 2020.

The second step, the kinematic forecasting step, employs forward computation of the secular variation by using the diffusion-less induction equation and is initialized by

$$\partial_t B_r(t_1) = -\nabla_h \cdot (\tilde{\mathbf{u}}(t_0) B_r(t_0)), \quad (28)$$

where  $\tilde{\mathbf{u}}(t_0)$  represents the flow ( $t_l^m(n), p_l^m(n)$ ), and  $B_r(t_0)$  the radial magnetic field at an initial epoch  $t_0$ . Here, we use the flow coefficients of the kinematic branch of  $C^3$  FM3. The forecasts of  $\tilde{\mathbf{u}}(t_i)$  for  $i \neq 0$  are obtained by deriving forecasts of the flow coefficients from their time-series models, similar to the operation in "Direct secular variation forecast" section. Generally, the forecasts of the secular variation are computed by

$$\begin{aligned} \partial_t B_r(t_n) &= -\nabla_h \cdot (\tilde{\mathbf{u}}(t_n) B_r(t_n)), \\ B_r(t_{n+1}) &= B_r(t_n) + \partial_t B_r(t_n) \delta t, \end{aligned} \quad (29)$$

where  $\delta t$  represents the sampling interval of 1 month, and  $\tilde{\mathbf{u}}(t_n)$  is the flow forecast vector that contains forecasts of the individual flow coefficients for the first month after 2020.0. Equation (28) is computed on a spherical grid, and then transformed to spherical harmonic secular variation coefficients ( $\dot{g}_l^m, \dot{h}_l^m$ ) similar to the scheme of Lloyd and Gubbins (1990). The forecast period is 120 months, and ends in 2030.0.

### Assessing the accuracy of forecasts

Choosing an appropriate error measure in forecasting is problematic, because no single measure gives an unambiguous indication of forecasting performance, while the use of multiple measures makes comparisons between forecasting methods difficult and unwieldy (Mathews and Diamantopoulos 1994). We apply a derivative of the widely used mean absolute percentage error, i.e., the symmetric absolute percentage error (sAPE) to assess the difference between the observation, i.e., the actual field and flow coefficient  $y(t)$  and its forecast  $\hat{y}(t)$  as:

$$\text{sAPE}(t) = \frac{2 |y(t) - \hat{y}(t)|}{|y(t)| + |\hat{y}(t)|}. \quad (30)$$

This follows the definition by Chen and Yang (2004), which differs from those of Armstrong (1985), Flores

(1986), as it does not become singular and has a maximum value of two when either  $y(t)$  or  $\hat{y}(t)$  is zero, but is undefined when both are zero. The range of (30) is (0, 2), i.e., the maximum value corresponds to 200% percentage error. We define a prediction length, for which the sAPE( $t$ ) becomes larger than 10%. This is a rather cautious definition and other limits could also be considered.

Another measure that is widely used (e.g., Aubert 2015; Whaler and Beggan 2015) is the rms-difference between models

$$\sqrt{dP} = \sqrt{\sum_{l=1}^{l_{max}} \sum_{m=0}^l (l+1) [{}^A g_l^m - {}^B g_l^m]^2 + [{}^A h_l^m - {}^B h_l^m]^2}, \quad (31)$$

where  ${}^A g_l^m, {}^A h_l^m$  and  ${}^B g_l^m, {}^B h_l^m$  are Gauss coefficients of the compared field models A and B.  $\sqrt{dP}$  represents

$$\sqrt{dP} = \int B d\Omega, \quad (32)$$

the total (difference) field integrated over the surface of the Earth, and similarly is the secular variation. We compute it in order to allow comparisons to other studies which use this measures as primary diagnostic. However, we must note, that  $\sqrt{dP}$  is biased towards large-scale contributions, when computed at Earth's surface where the spatial energy spectrum of the field is decreasing, and is biased towards small scales at the core surface, where the spectrum is increasing. This is not the case for the sAPE( $t$ ), as it is relative to the amplitude of the actual coefficient.

## Results and discussion

In this section, we present results of the core field modeling and the secular variation forecasts. At the end of this section, we also present and discuss our candidates for the definitive geomagnetic reference field model (DGRF) in 2015, and IGRF candidate model for 2020.

We focus on two models: Model 1 was derived with the classical modeling approach and Model 2 is a kinematic field model based on the tangential geostrophic flow assumption. Table 1 compiles the set of model parameters and related model characteristics. Both models provide almost equal fits to the data in terms of their residual standard deviation.

We note that, by strongly imposing the tangential geostrophic constraint, the fit to the radial induction equation deteriorates, when compared to solutions with a weaker tangential geostrophic constraint, see Appendix Table 3. The reason for this is not clear, but, perhaps, could be explained by the penalization of the poloidal  $p_n^0$  terms when the tangential geostrophic constraint is imposed.

**Table 1 Model diagnostics: residuals standard deviation in [nT/year] and damping parameters**

	Standard deviation [nT/year]			Damping	
	X	Y	Z	$\lambda_S$	$\lambda_T$
Model 1	13.69	11.07	14.47	3.0d-10	3.0d-2
Model 2	12.42	10.89	13.89	see Model x3d Table 3	

### Comparison of modeled and observed secular variation

In Fig. 2, we show the rms differences (31) between the two models during the model period. The rms differences of the main field (Fig. 2, left) and secular variation coefficients (Fig. 2, right) are continuously decreasing during the period, except close to the model endpoint in 2020, where they increase. These increases can be explained by the constraint on the flow acceleration in Model 2. We argue that the decrease in differences during the model period is not directly related to the growth of the available ground-based geomagnetic observatory data, because this affects both models. Rather, the decrease in differences could be explained by the different modeling techniques. As it was found in a previous study (Wardinski and Lesur 2012), the kinematic field modeling seems to project the field morphology of a spatially high-resolved satellite field model imposed in 2010 backward in time. Therefore, the kinematic model is less subject to a varying data distribution, whereas the classical field model is.

Figure 3 compares the observed and modeled secular variation at different observatory sites. The model curves are derived from our preferred classical and kinematic field models. The observed secular variation series from ground-based geomagnetic observatories stop in 2017.5, as more recent data were not available at the time of this study.

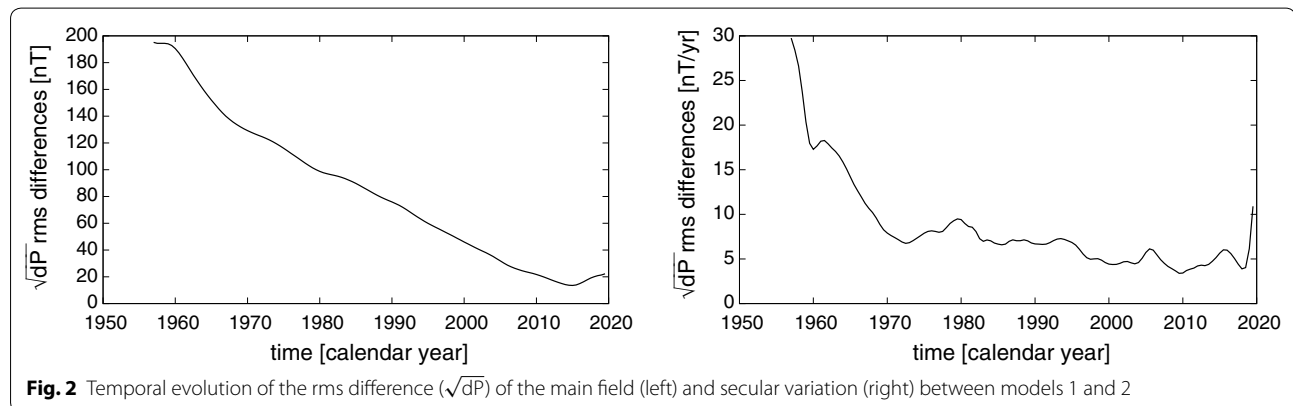
The overall impression is that the classical and kinematic field models fit the observed secular variation

equally well. An exception is, at the model endpoints where the two curves deviate. Most likely, this is due to the endpoint constraint of the kinematic field models. The known jerk occurrences (A–J) are all clearly visible in the Y-component of Chambon-la-Forêt, whereas in other observatory data only some jerks are detectable, e.g., the series of Sitka shows only events in 1969 and 2011. Data of the Sitka observatory show large changes of the secular variation of Y and Z components in the past decade, that may be related to the rapid movement of the magnetic North Pole.

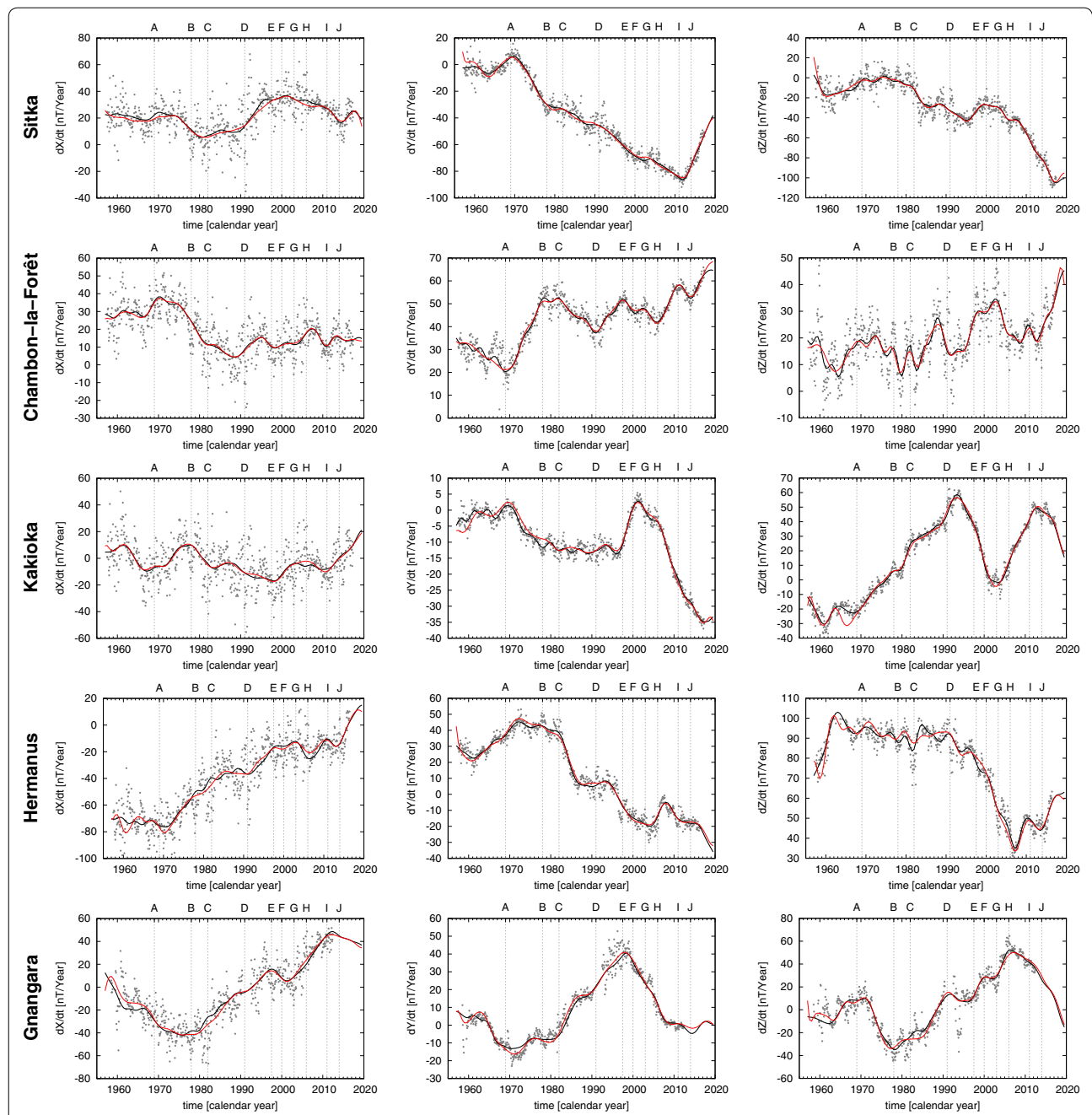
The same large similarity between the models can be observed in Fig. 4, which shows six secular variation coefficients with the largest amplitude of the two different models. These coefficients largely determine the morphology of the secular variation at Earth's surface. Particularly, the coefficient  $h_2^1$  is closely related to the prominent patch of reverse magnetic flux in the southern hemisphere (Terra-Nova et al. 2017). Overall, differences between the model coefficients rarely exceed 2 nT/year, except close to model endpoints, where models disperse largely.

Figure 4 also shows markers of known geomagnetic jerks, to allow their identification in the evolution of the secular variation coefficients. The coefficients  $\dot{g}_1^0$  and  $\dot{h}_2^1$  represent equatorial anti-symmetric contributions of the secular variation. These coefficients carry most of the known geomagnetic jerks, apart from the event in 1969 which is either not clearly visible in  $\dot{g}_1^0$ , or appears later in 1970. The identification of geomagnetic jerks in the equatorial symmetric parts of the secular variation described by  $\dot{g}_1^1, \dot{h}_1^1, \dot{g}_2^0$  is less clear.

The temporally averaged secular variation spectra of the two models are shown in Fig. 5. On large and small scales, differences between spectra remain small at Earth's surface. The spectra also indicate a very high temporal variability of the kinematic field secular variation at degrees 9 and 10, suggesting that at some epochs the





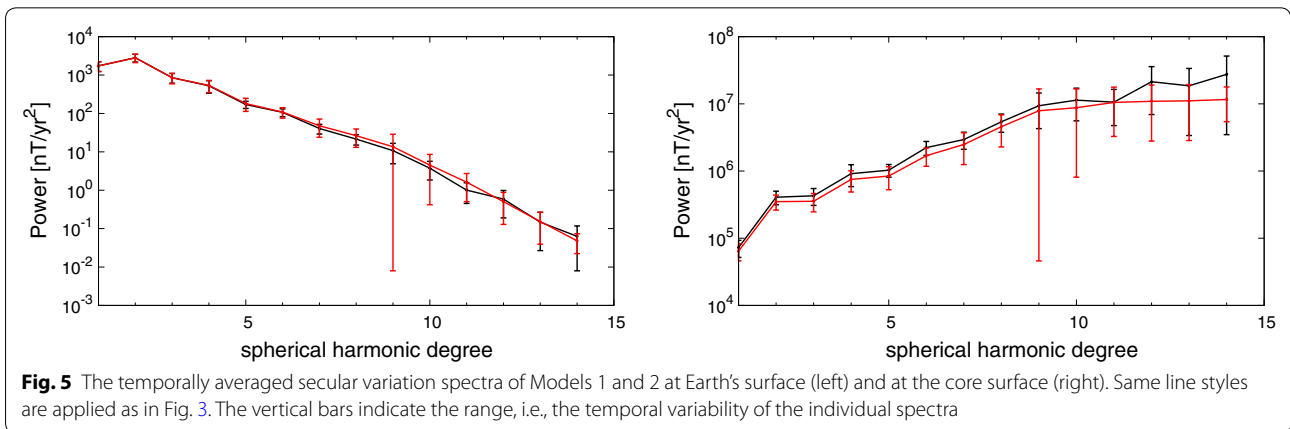
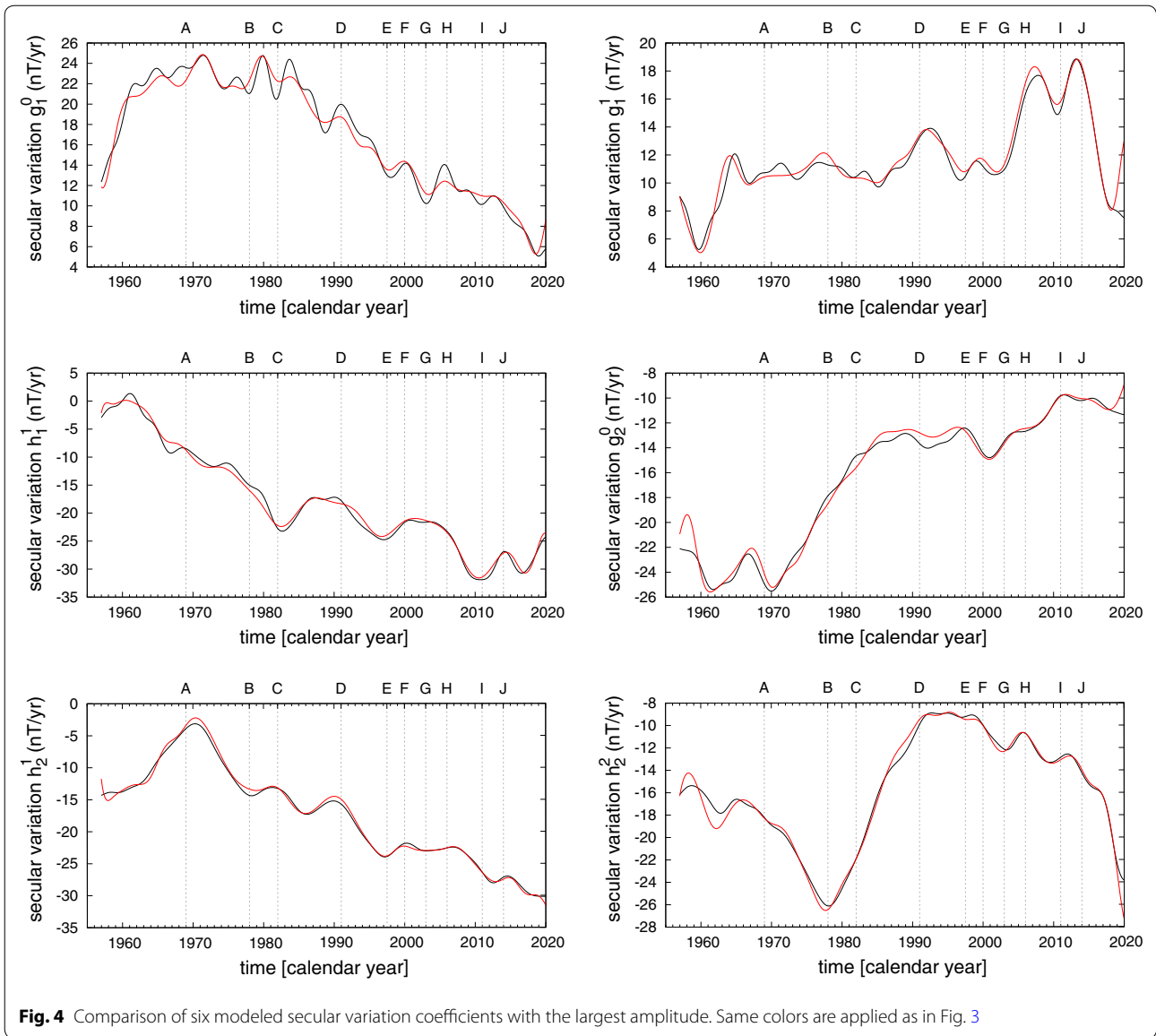


**Fig. 3** Observed and modeled secular variation at some observatory sites. From top to bottom: Sitka (Alaska), Chambon-la-Forêt (France), Kakioka (Japan), Hermanus (South Africa), Gngagara (Australia). The gray dots represent the observed monthly secular variation in X, Y and Z (from left to right). Solid curves display the modeled secular variation of Model 1 (black) and Model 2 (red). Vertical lines mark occurrences of geomagnetic jerks

power of the secular variation of these models is weaker by a few orders of magnitude than the secular variation of the classical field model.

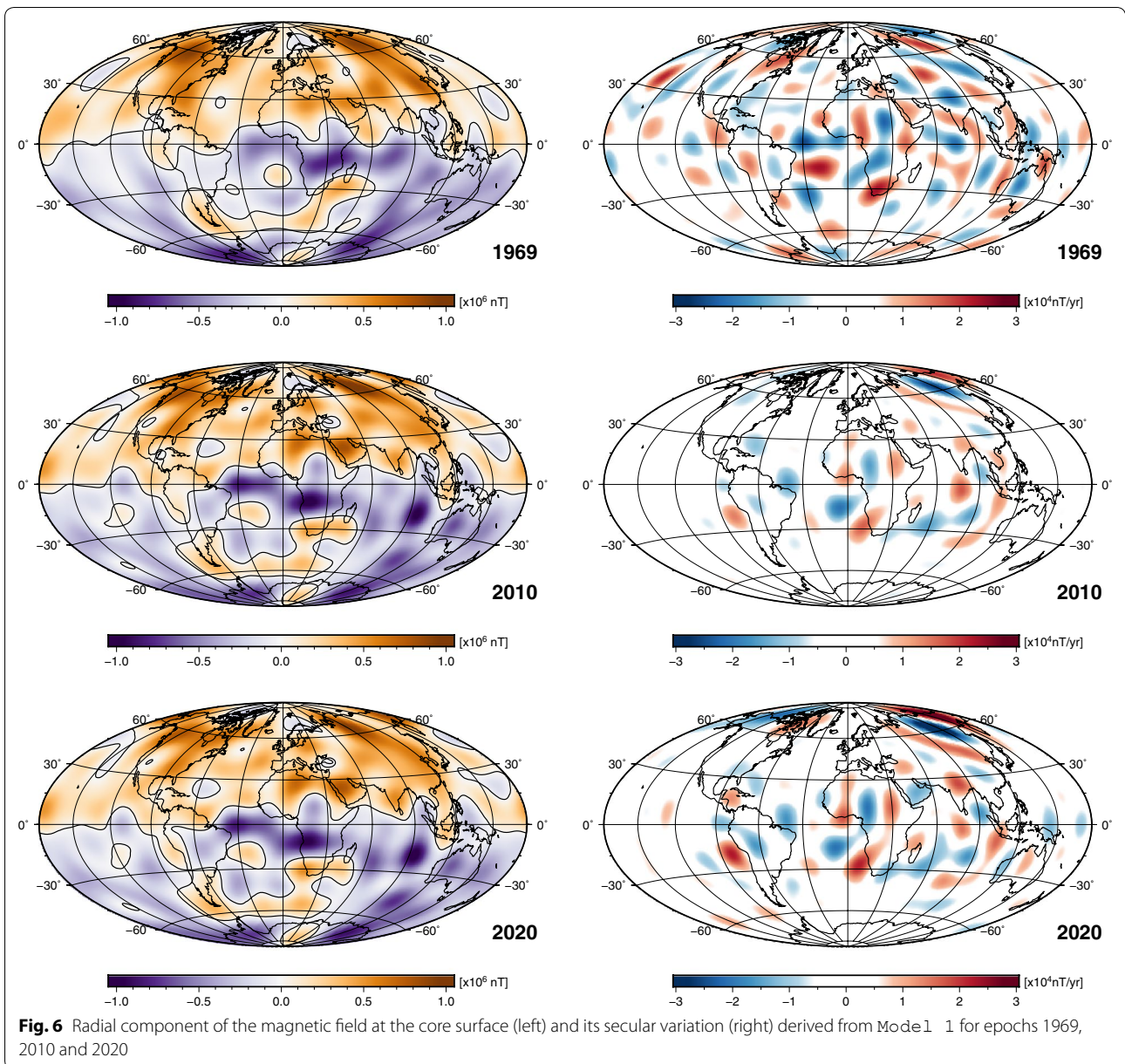
At the core surface (Fig. 5, right) the spectral power of the secular variation grows with spherical harmonic degree. The spectra of the two models start to deviate

from degree 11, where the spectrum of Model 2 flattens, whereas the spectrum of Model 1 continues to increase. We note the same high temporal variability at spherical harmonic degree 9 of the kinematic secular variation model.

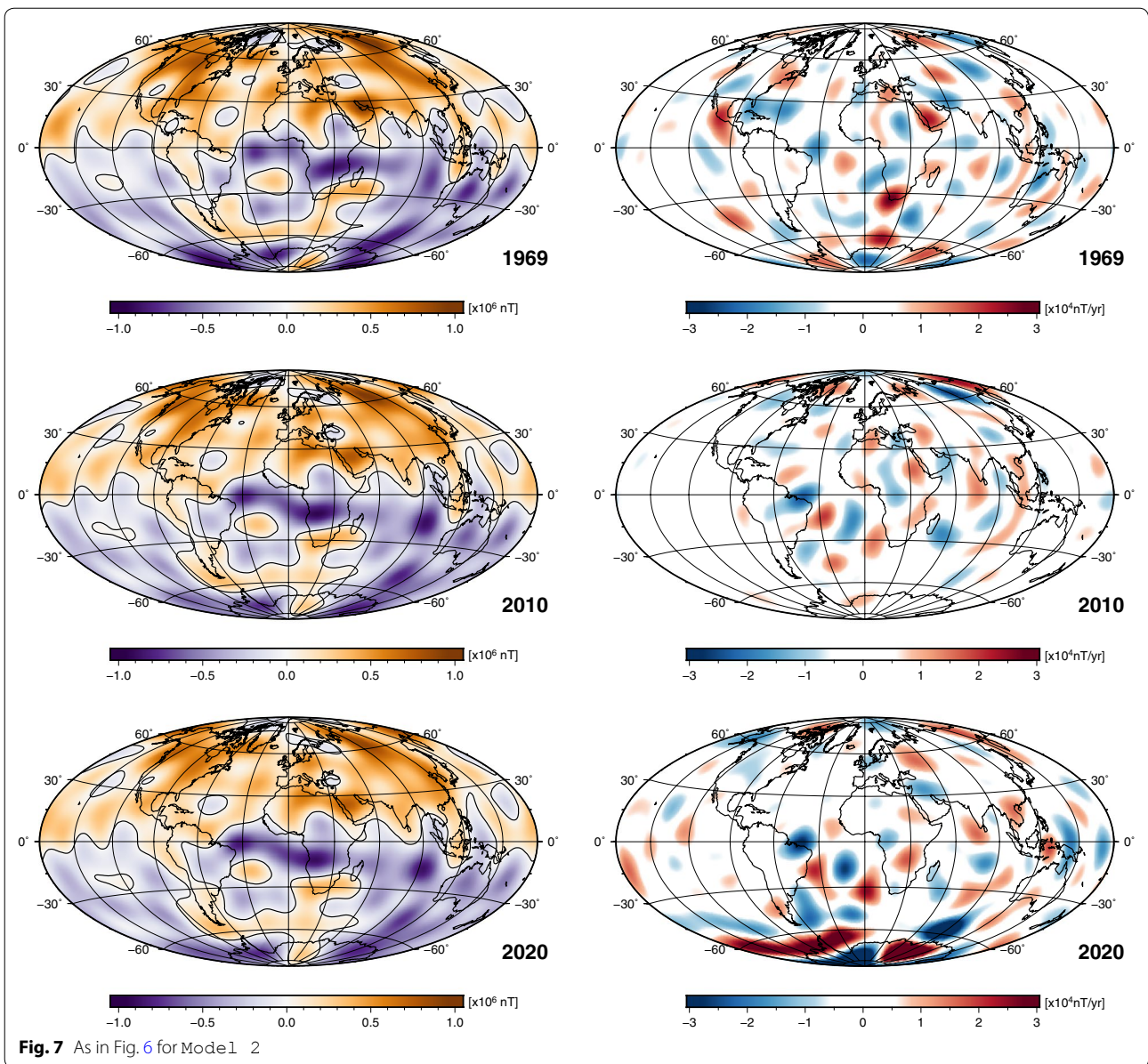


Maps of the radial component of the magnetic field and its secular variation derived from Model 1 at the top of the core are shown in Fig. 6. The radial magnetic field component shows a dipole-dominated morphology, but with considerably small-scale features, such as reversed flux patches in both hemispheres. The secular variation at the core surface tends to be dominated by small scales. Particularly notable features are low latitude pairs of opposite polarity secular variation which are typical to advection (Amit 2014). Also note larger patches of the radial secular variation in the vicinity of the magnetic North Pole, e.g., under Eastern Siberia.

Figure 7 shows maps based on Model 2. Maps of the radial component of the magnetic field at the core surface are widely in agreement with the respective maps of Fig. 6. Perhaps most apparent, is the presence of the reversed flux patch underneath the East Pacific, in both models except at 1969 in Fig. 6. Maps of the secular variation in 2020 differ substantially, which may be due to the endpoint constraint of Model 2. The secular variation map of Model 2 in 2020 contains some large-scale anomalously intense structures below the geographic South Pole, indicating an intensification of the flux in this region. Based on the strong secular variation at high



**Fig. 6** Radial component of the magnetic field at the core surface (left) and its secular variation (right) derived from Model 1 for epochs 1969, 2010 and 2020



**Fig. 7** As in Fig. 6 for Model 2

latitudes of the northern hemisphere, Livermore et al. (2017) inferred a zonal jet there. In the context of equatorial symmetry, they argued that such a zonal jet at high latitudes of the southern hemisphere would not produce detectable secular variation, because the field is oriented in the east–west direction there. Indeed in Fig. 6 in 2020 the secular variation is very weak around the South Pole (and quite strong under the North Pole). All this makes the anomalously strong secular variation around the South Pole in 2020 as seen in Fig. 7 quite suspicious. However, secular variation maps of other epochs (1969 and 2010) are very similar to maps of Fig. 6, with dominant small-scale features at mid- and low latitudes.

We conclude that classical and kinematic field models largely agree, both in spatial morphology as well as in the temporal evolution, but significantly deviate in their secular variation towards the endpoints. The cause for this deviation is the constraint of the kinematic field model onto the flow acceleration, while the classical model gets along without such constraint.

**Direct secular variation forecast**

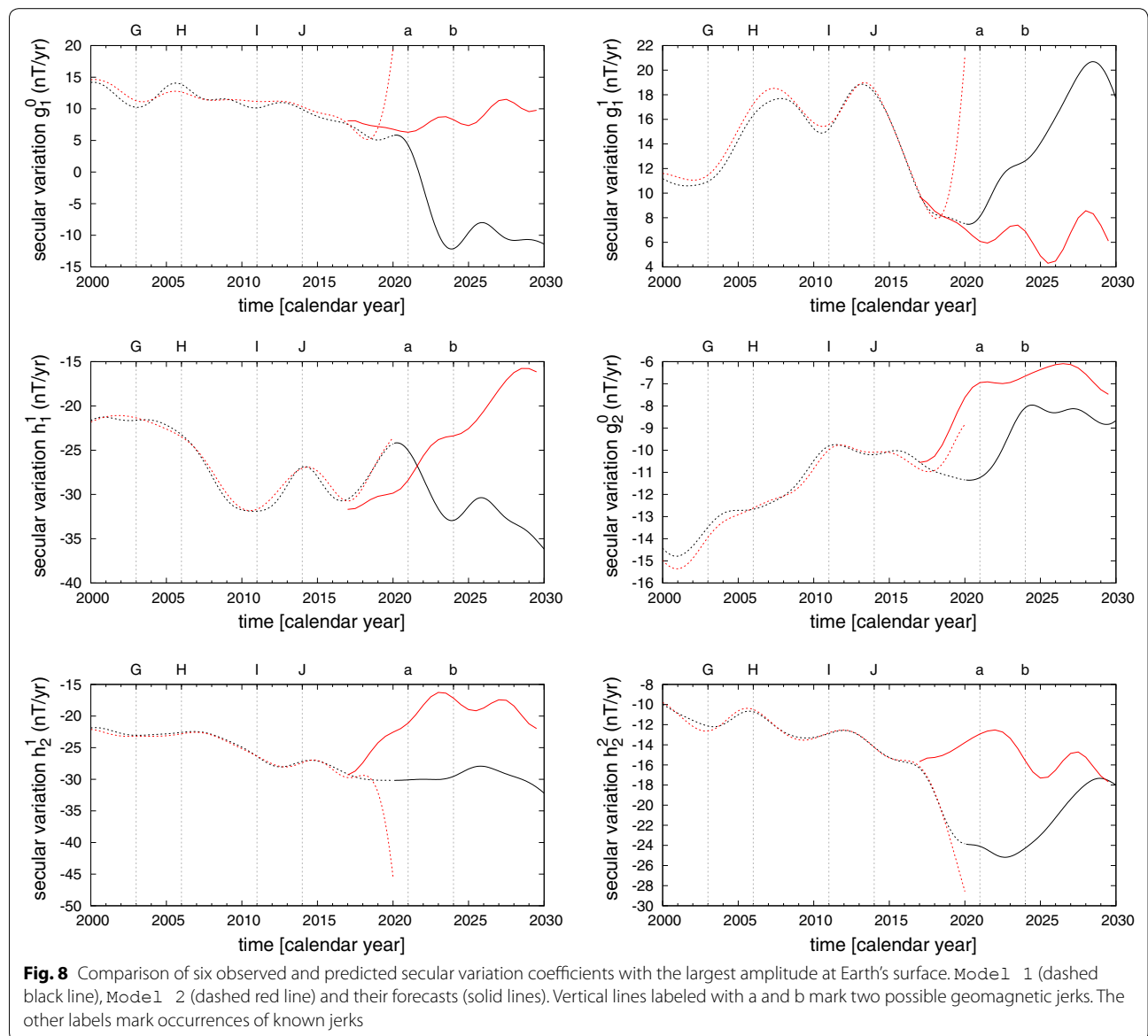
Figure 8 displays observations and predictions of secular variation coefficients with the largest amplitude at the Earth surface, i.e.,  $\dot{g}_1^0, \dot{g}_1^1, \dot{h}_1^1, \dot{g}_2^0, \dot{h}_1^2, \dot{h}_2^2$ , from 2000 to 2030. The most prominent feature of the forecast of Model 1

is the steep decrease of  $\dot{g}_1^0$  by almost 15nT/year during a very short interval of about 3 years. While this decrease is determined by the statistical properties of the time-series model, it may be difficult to accept this when considering the past secular variation. In fact, this decrease would be caused by a large secular acceleration. Advective sources and sinks of  $g_1^0$  exhibit a lot of cancellations and it is the rather small residual that gives the historical dipole decay (Olson and Amit 2006; Finlay et al. 2016), i.e., subtle change in the core flow pattern and its interaction with the field may yield considerable changes in  $g_1^0$ . However, apparently such a change did not happen since 1840. The other terms vary within ranges of previous oscillations (see Fig. 4).

Most of the secular variation terms of Model 1 indicate possible occurrences of two future geomagnetic jerks in late 2020–early 2021, and in early 2024. In Fig. 8, at these dates the predicted secular variation shows notable changes in slope.

**Kinematic secular variation forecast**

The kinematic secular variation forecast is based on forecasts of individual flow coefficients derived from a multivariate singular spectrum analysis (MSSA). Figure 9 shows the past and future temporal evolution of two toroidal flow coefficients,  $t_1^0, t_3^0$ . These coefficients are of a particular interest, as they relate to changes in Earth’s rotation (Jault et al. 1988; Jackson et al. 1993), which



coincide with geomagnetic jerks (Holme and de Viron 2005; Wardinski et al. 2008).

The MSSA of the temporal flow variability is truncated at the degree where past model and forecast of the flow coefficients becomes almost continuous, which is achieved for a truncation level  $k = 22$ , i.e., the first 22 eigenmodes. Both flow coefficients show spurious accelerations from 2017 on, which is caused the endpoint constraint of the flow acceleration, cf. (23). Therefore, we consider the kinematic field and flow coefficients to be flawed and start the kinematic forecast from 2017.

Figure 8 also provides a comparison between the direct and kinematic forecasts of secular variation coefficients. The kinematic forecast starts in 2017 to reduce the influence of the model's faulty temporal behavior close to its endpoint. These forecasts differ by their temporal evolution to those of the direct secular variation forecast. Most apparent are these differences in the forecast of  $g_1^0$ , where both forecasts have opposite signs and differ by about 20 nT/year. The kinematic forecast predicts a gently varying  $g_1^0$ . This forecast discrepancy is also clearly seen for  $g_1^0, g_1^1$  and  $h_2^1$ . However, these forecasts show also common features related to future geomagnetic jerks. The cause for the forecast discrepancy is not understood, and it is not in relation to the anomalous flow acceleration at the kinematic model endpoint, as the kinematic forecast starts a few years prior to the model's endpoint.

### Predictability

Previous studies (Aubert 2015; Whaler and Beggan 2015) used rms-based measures (31) to quantify differences between a reference model and forecasts. The reference model is a model for the epoch 2017 of the classical and kinematic model branches and labeled as M1 2017 and M2 2017, respectively.

Table 2 lists the rms-differences between different epoch models of Model 1 and Model 2. Generally,

the rms-differences of the direct secular variation forecasts (Model 1) is smaller than those of the kinematic forecasts. However, it is not clear whether it allows to conclude upon the performance of the forecasts. Globally, it suggests that the main field and secular variation of Model 1 tend to be more similar over longer time intervals than for Model 2. This may agree with results based on numerical dynamos (Hulot et al. 2010) and data assimilation using numerical dynamos (Aubert 2015). Their studies used rms-based error estimates, and their results suggested a possible predictability of several decades to a century.

We perform several prediction experiments to measure the prediction length of the first 26 secular variation coefficients for different prediction setups. In these setups, the starting of the forecast is set to 1995, 2005 and 2015, respectively. The forecast period is in all cases 10 years, for which we compute the  $sAPE(t)$ , cf. (30).

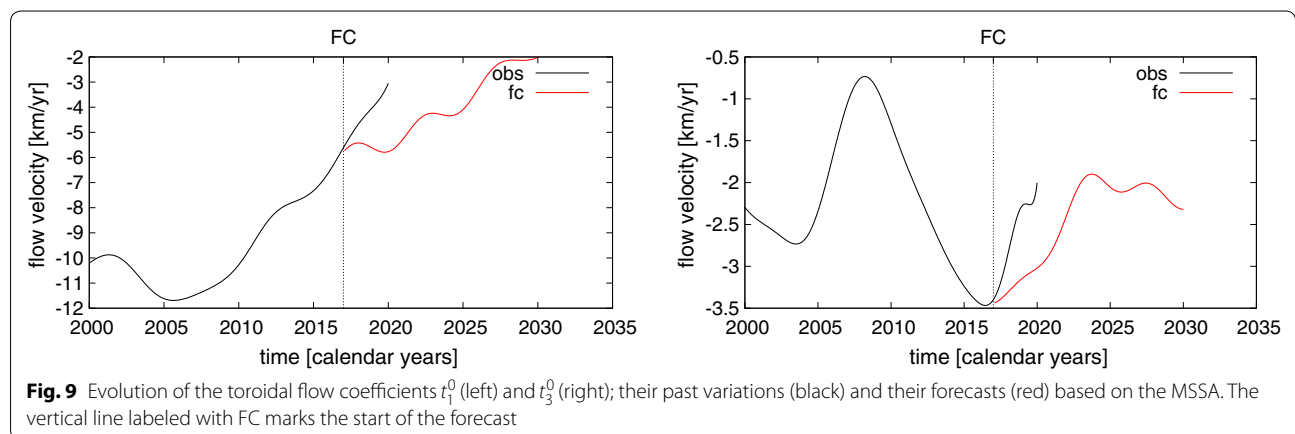
First, it is found for the direct secular variation forecasts of the classical field model that the prediction length is largely independent of the number of eigenvalues considered in the MSSA forecast. Merely, this number has to be larger than the number of significant eigenvalues.

Results of the forecast experiments are shown in Figs. 10 and 11. Generally, the prediction error ( $sAPE(t)$ ) of the zonal secular variation coefficients

**Table 2** The rms differences between models of different epochs and Model 1 and Model 2

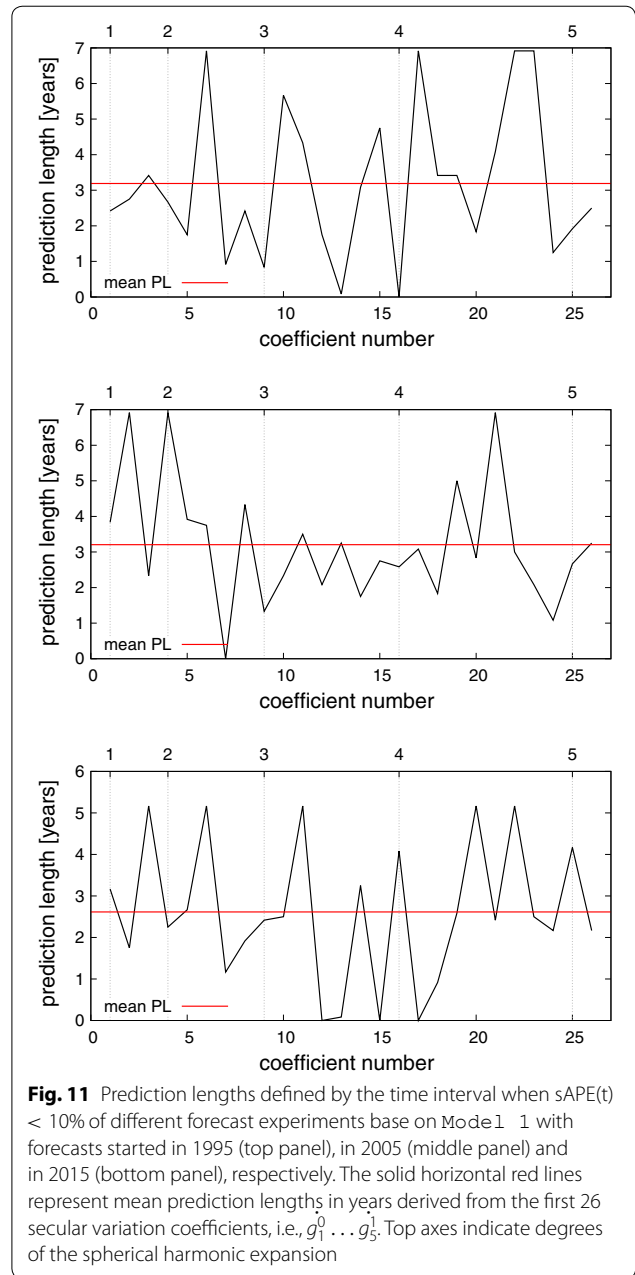
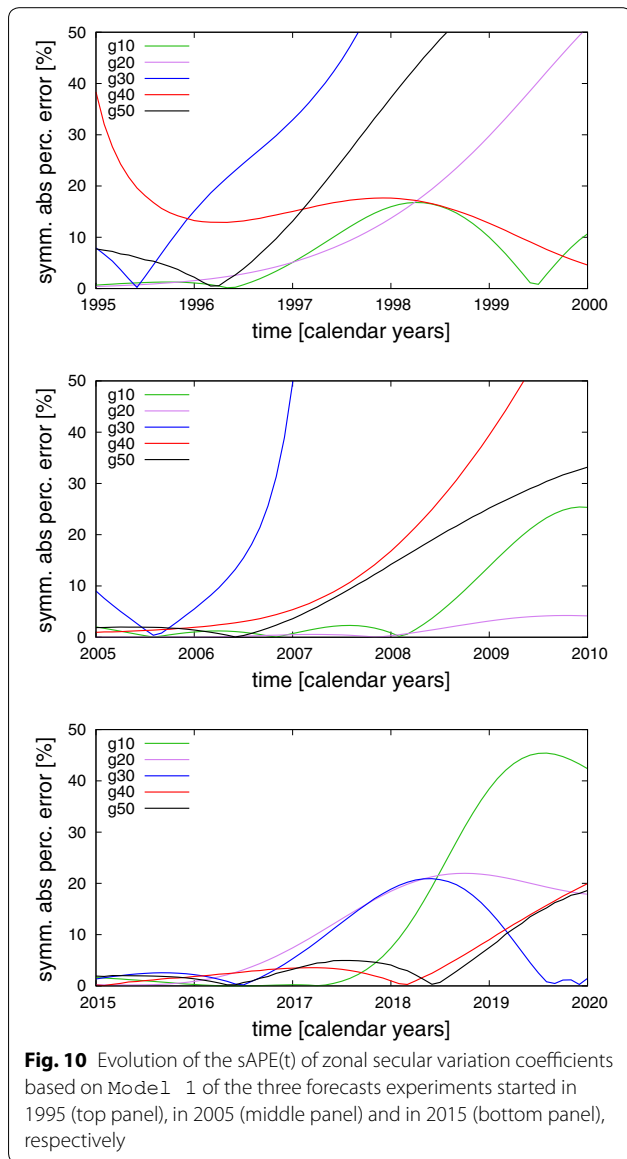
	Main field		Secular variation	
	M1 2017	M2 2017	M1 2017	M2 2017
2020	271.41	863.49	21.00	22.11
2025	306.11	1511.43	34.70	40.79
2030	306.54	359.97	35.13	90.02

Differences of main field models are given in [nT] and differences of secular variation models are in [nT/year]



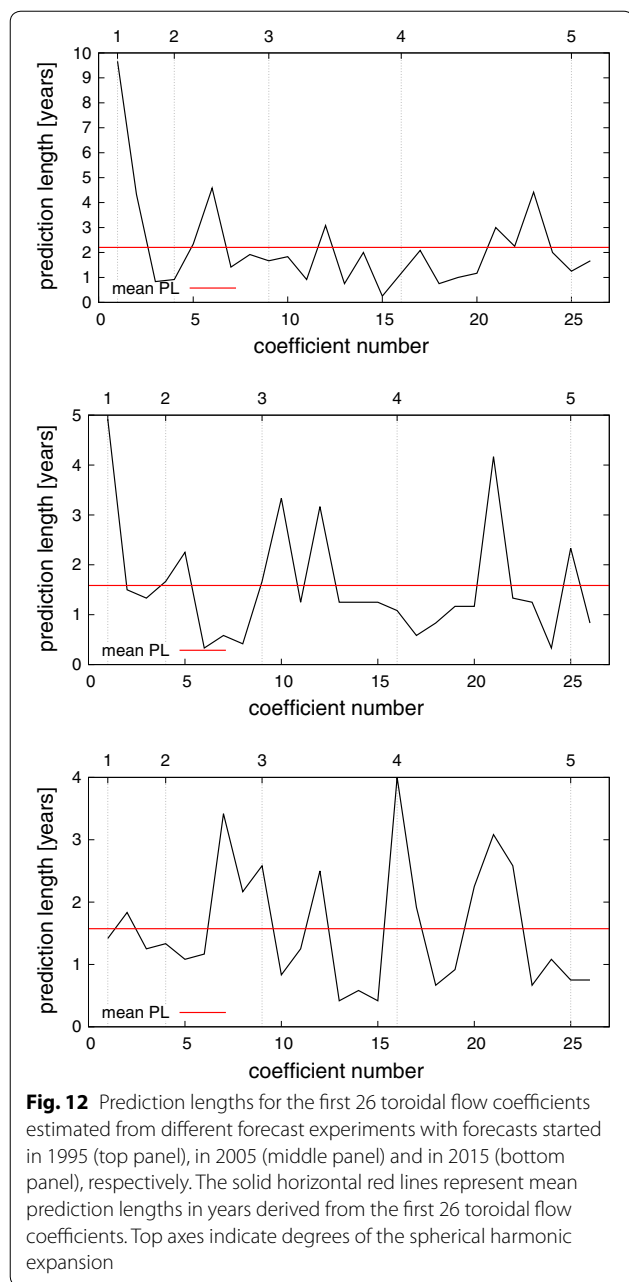
increases with time, except for the prediction of  $\dot{g}_4^0$  of the 1995-experiment which improves with time. It is also noted that prediction error of the 2015-experiment is mostly smaller than for the other experiments (Fig. 10). We focus on the prediction of the large-scale secular variation represented by the first coefficients. For practical purposes, we arbitrarily examine the first 26 coefficients. The mean prediction length for which the prediction error  $sAPE(t) < 10\%$  of the first 26 secular variation coefficients is around 3 years (Fig. 11). This is substantially shorter than results by Aubert (2015), who found a possible predictability of a century.

Similar experiments are performed for the flow forecasts of the kinematic field model (Model 2), with the



resulting prediction lengths shown in Fig. 12. The mean prediction length for all experiments is always shorter than 2.5 years, i.e., shorter than the mean prediction length of direct secular variation forecasts (Fig. 11). However, we note that the prediction length of  $t_1^0$  for experiments started in 1995 and 2005 is longer or equal to 5 years, and that only the forecast deteriorates only when the experiment is started closer to the model endpoint.

We avoided performing such experiments for the kinematic secular variation forecast, because of the faulty



temporal behavior of Model 2 close to its endpoints. This behavior is caused by the constraint that controls the flow acceleration at the model endpoints by  $\lambda_E$ . Constraining the flow acceleration too strongly leads to a very small flow acceleration, which is as unrealistic as the counter case with large flow acceleration. To this end, conclusions about the flow acceleration at the endpoints are loose, and in the future we may constrain secular variation of kinematic field models to be similar to the secular variation of the classical model.

### Derivation of candidate models

Our candidate models for the DGRF 2015, IGRF 2020, and the secular variation for the period 2020 to 2025 are derived from Model 1 and its respective forecast. This is justified by the large resemblance of the main field description of two models. The candidates for the DGRF 2015 and IGRF 2020 are given by the main field model in 2015.0 and 2020.0, respectively, truncated at spherical harmonic degree 13, in nT with two decimal places. The candidate of the secular variation model is given by the forecast for the epoch 2022.5 truncated to spherical harmonic degree 8.

### Conclusion

We derive field models for the period 1957 to 2020 from ground-based geomagnetic observatory data as well as satellite-based virtual observatory data. These models are constructed using two different techniques, a classical modeling (Model 1) that provides an optimal fit to observations, and a data assimilation to a dynamical assumption of Earth's core flow (Model 2). These strategies provide similar results for the core field and its temporal variations over the past decades.

In this study we set up two forecasting schemes that rely on analyses of multi-variate time series of secular variation coefficients. We derived time-series models of the field variability, from which forecasts of individual secular variation coefficients are obtained. These serve as direct secular variation predictions. Prediction experiments indicate a robust forecast of the secular variation of about three years. This might be a lower boundary and is determined by our cautious definition of the prediction length; the time interval when  $\text{sAPE}(t) < 10\%$ .

Forecasts of the flow coefficients are derived in the same manner and used in forward calculations of the secular variation due to advection in the core, where contributions from magnetic diffusion are neglected and a tangential geostrophy assumption couples the toroidal and poloidal flows (kinematic secular variation forecast). This approach extends beyond the approach of Beggan and Whaler (2010); Whaler and Beggan (2015) which used steady and constantly accelerated flow to predict future secular variation. However, our kinematic secular variation forecast suffers from faulty estimates of the core flow at the model endpoints due to a possibly inaccurate restriction of the flow acceleration. Therefore, the latter approach may not provide robust forecasts of the secular variation, unless the endpoint constraint can be dropped. Interestingly, both strategies consistently indicate the occurrence of future geomagnetic jerks in 2021 and 2024. The uncertainty in the exact timing of these events is related to the original temporal resolution of our data, which is optimistically smaller than  $\pm 1$  year



and realistically larger than  $\pm 6$  months, as well as to the strength of the temporal constraint.

Of course, our strategies of forecasting are limited by the lack of observations from within the (geodynamo) system (which are not available), but not by inferences made upon the geodynamo. Similar limitations may occur to approaches using a full description of the geodynamo in a data assimilation framework to forecast geomagnetic secular variation, like Aubert (2015), Fournier et al. (2015). We hypothesize that when longer time series are considered, a longer behavior of the field can be modeled; however, this will not improve the predictability of the short-term (decadal) field variations, as they may be chaotic.

To this end, results of the classical modeling (Model 1) are used to provide candidates for the definitive geomagnetic reference field model (DGRF) in 2015, and IGRF candidate model for 2020. The candidate of the secular variation model is given by the forecast for the epoch 2022.5.

#### Abbreviations

CMB: Core–mantle boundary; DGRF: Definitive geomagnetic reference field; IAGA: International Association of Geomagnetism and Aeronomy; IGRF: International Geomagnetic Reference Field; MSSA: Multi-variate singular spectrum analysis; sAPE: Symmetric absolute prediction error; VO: Virtual geomagnetic observatory.

#### Acknowledgements

The authors would like to record their gratitude to scientists working at magnetic observatories, data suppliers and involved national institutes for making geomagnetic observatory data available, to the World Data Centre for Geomagnetism (Edinburgh) for compiling these data in a database and to the International Service of Geomagnetic Indices for making data products (indices) available. The authors acknowledge ESA for providing access to Swarm Level 1b products. We also would like to thank V. Lesur for sharing his satellite-based geomagnetic field model before its submission. This work was supported by CNES. The critical reading and very valuable comments by two anonymous reviewers are appreciated.

#### Authors' contributions

IW organized the manuscript based on the analyses of the contributing co-authors. All authors contributed modeling results and/or detailed technical analyses and/or discussions for this study. All authors read and approved the final manuscript.

#### Funding

IW was funded at an initial state of this study by the ESA project 3D-Earth and by the Centre National des Etudes Spatiales (CNES) within the context of the project "Exploitation des mesures de la mission Swarm". DS was funded by CNES project "Exploitation des mesures de la mission Swarm". HA acknowledges financial support from "Programme National de Planétologie" of "Institut national des sciences de l'Univers". AC acknowledges financial support from CNES over the project "Indices d'Activité Magnétique : Météorologie de l'Espace et Géomagnétisme Spatial". This work was supported by CNES.

#### Data availability statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup> Institut de Physique du Globe de Strasbourg, Université de Strasbourg/EOST, CNRS, UMR 7516, Strasbourg, France. <sup>2</sup> Laboratoire de Planétologie et Géodynamique, Université de Nantes, Université d'Angers, CNRS, UMR 6112, Nantes, France. <sup>3</sup> Centre National d'Etudes Spatiales (CNES), Paris, France. <sup>4</sup> Laboratoire Magmas et Volcans, Université de Clermont Auvergne, CNRS UMR 6524, Clermont Ferrand, France.

## Appendix A: Theory of singular spectrum analysis

We briefly recall some aspects of the singular spectrum analysis (SSA). The SSA is a non-parametric spectral estimation method that incorporates aspects of classical time-series analysis, signal processing, multi-variate statistics and geometry (Vautard et al. 1992; Golyandina et al. 2001). We first formulate the theoretical concepts for uni-variate time series, and then extend them to multi-variate time series. For a deeper review and also its application to a variety of phenomena in meteorology, oceanography and climatology we refer to Ghil et al. (2002) and references therein.

#### Uni-variate analysis

The analysis is based on the embedding of a time series  $y_t$  in a vector space of dimension  $M$  that determines the longest periodicity captured by the analysis. The spectral information on a time series are obtained by diagonalizing the  $M \times M$  covariance matrix  $C_y$  of  $y_t$ . The covariance matrix  $C_y$  can be estimated directly from the data, i.e., its entries  $c_{ij}$  depend only on the lag  $\delta = |i - j|$ :

$$c_{ij} = \frac{1}{N - \delta} \sum_{t=1}^{N-\delta} y_t \cdot y_{t+\delta}. \quad (\text{A.1})$$

Usually, the decomposition of the covariance matrix  $C_y$  in  $q$  orthogonal eigenvectors  $E_q$  that are called temporal empirical orthogonal functions (EOF) is done by singular value decomposition. The eigenvalues  $\epsilon_q$  of  $C_y$  account for the partial variance in the direction  $E_q$  and the sum of the eigenvalues, i.e., the trace of  $C_y$ , gives the total variance of the original time series  $y_t$ . By a projection of the time series onto the EOF the temporal principal component  $A_t^k$  can be obtained:

$$A_t^k = \sum_{j=1}^M X(t + j - 1) E_k(j). \quad (\text{A.2})$$

In fact, this method decomposes the time series in parts that correspond to trends, oscillatory modes or noise. Therefore, it allows the time series to be reconstructed and to be forecasted by using linear combinations of the temporal principal components and EOFs. The reconstructed components  $R_t^k$  are given by

$$R_t^k = \frac{1}{N_t} \sum_{k \in K} \sum_{j=L_t}^{U_t} A^k(t-j+1)E^k(j), \quad (\text{A.3})$$

where  $K$  is the set of  $k$  EOFs and temporal principal components on which the reconstruction is based. We refer to  $k$  as the truncation level of the reconstruction and forecast. This truncation level is  $k \leq q$ . The values of the normalization factor  $N_t$ , as well as of the lower and upper bound of summation  $L_t$  and  $U_t$ , differ between the central part of the time series and in the vicinity of its endpoints (Ghil et al. 2002).

An oscillatory mode is characterized by a pair of nearly equal eigenvalues and by associated principal components that are in approximate phase quadrature (Ghil et al. 2002). Such a pair can represent a non-linear, non-harmonic oscillation, because a single pair of eigenmodes are more sensitive to the basic periodicity of an oscillatory mode than methods with fixed basis functions, such as the sines and cosines used in the Fourier transform.

### Multi-variate analysis

The multi-variate (or multi-channel) singular spectrum analysis (MSSA) is a generalization of the SSA from univariate to multi-variate time series, such as time series of individual Gauss coefficients. Its use was proposed theoretically in the context of non-linear dynamics (Broomhead and King 1986) and numerous examples of successful application of this methods can be found, e.g., Plaut and Vautard 1994.

The MSSA allows the identification of dynamically relevant temporal patterns that are coherent in series that form a multi-variate time series. These individual series are often called channels. By analogy to the SSA, each of  $L$ -channel data vectors  $y_{l,t} : l = 1, \dots, L, t = 1, \dots, N$  is expanded to a state vector

$$Y_t = \begin{pmatrix} y_{l,1} & y_{l,2} & \dots & y_{l,M} \\ y_{l,2} & y_{l,3} & \dots & y_{l,M+1} \\ \dots & \dots & \dots & \vdots \\ y_{l,N-M} & \dots & \dots & y_{l,N-1} \\ y_{l,N-M+1} & \dots & \dots & y_{l,N} \end{pmatrix} \quad (\text{A.4})$$

for each channel  $l = 1, \dots, L$ , and the window length  $M$ . Following the approach of Broomhead and King (1986); Allen and Robertson (1996), then the MSSA relies on the construction a grand covariance matrix  $C_x$  like

$$C_x = \frac{1}{N-M+1} Y^T Y = \begin{pmatrix} C_{11} & C_{12} & C_{13} & \dots & C_{1L} \\ & C_{22} & C_{23} & \dots & C_{2L} \\ & & C_{33} & \dots & C_{3L} \\ & & & \ddots & \vdots \\ & & & & C_{LL} \end{pmatrix}, \quad (\text{A.5})$$

where each block  $C_{ll'}$  is a covariance matrix between channels  $l$  and  $l'$ . The blocks  $C_{ll'}$  have the entries

$$(C_{ll'})_{jj'} = \frac{1}{N-M+1} \sum_{t=1}^{N-M+1} Y_{t+j-1}^l Y_{t+j'-1}^{l'}. \quad (\text{A.6})$$

The  $LM \times LM$  matrix  $C_x$  is symmetric and by diagonalizing,  $LM$  eigenvectors  $E^k : k = 1, \dots, LM$  can be obtained. The eigenvectors are composed of  $L$  consecutive segments with length  $M$ . Similarly to (A.2) the associated principal components  $A^k$  can be computed by projecting  $Y$  in the directions of the eigenvectors (i.e. onto the EOFs):

$$A_t^k = \sum_{j=1}^M \sum_{l=1}^L Y_{t+j-1}^l E^{l,k}(j). \quad (\text{A.7})$$

$E^{l,k}(j)$  are the elements of the eigenvectors. The  $A^k(t)$  are single-channel time series and likewise to (A.3) the  $k$ th reconstruction of the signal of channel  $l$  can be obtained by:

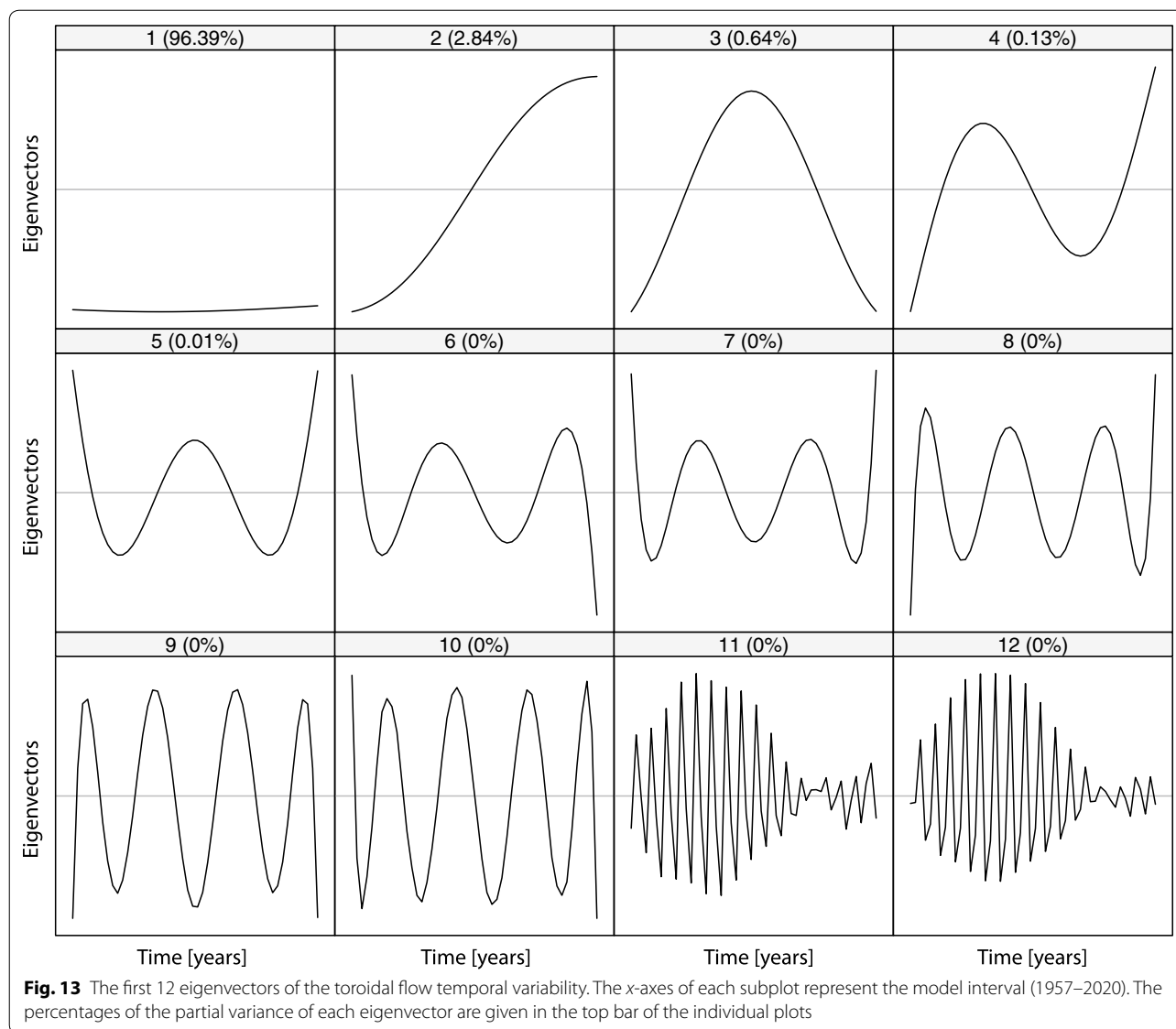
$$R_t^{l,k} = \frac{1}{M_t} \sum_{j=L_t}^{U_t} A^k(t-j+1)E^{l,k}(j), \quad (\text{A.8})$$

where  $L_T$  and  $U_t$  are the lower and upper bound of summation, respectively.

Figure 13 shows the first 12 eigenvectors of the toroidal flow decomposition. Clearly, the first two eigenvectors capture the slow variation of flow coefficients, whereas other eigenvectors represent the shorter temporal variations of the flow. The first few eigenvectors explain nearly the entire signal variance, which is given on top of each single plot. However, higher indexed eigenvectors also show non-zero partial variance, suggesting that these eigenvectors may carry relevant temporal information.

An important aspect of the analysis is how well components of the time series are separated from each other. The components generally group in two disjunctive parts, one corresponds to the signal, the other to the noise. The signal could be composed of slowly varying, periodic and/or quasi periodic components. A way to measure the separation between components, is to calculate the weighted correlation or w-correlations, as given by Golyandina et al. (2001).

In Fig. 14, the so-called, w-correlation matrix is displayed. It shows the weighted correlations for principal components,  $A^k$ , of the temporal flow variability. Accordingly, large values of w-correlation exist for diagonal elements of the matrix, where individual modes correlate with themselves (dark red color). Whereas, small values of the w-correlation (lighter



colors) indicate orthogonality between components, as the scalar product between perpendicular vectors vanishes. Therefore, small correlation values indicate a good separation of components or noise. In particular, noise components do not correlate with each other and other components, and therefore, show zero correlation.

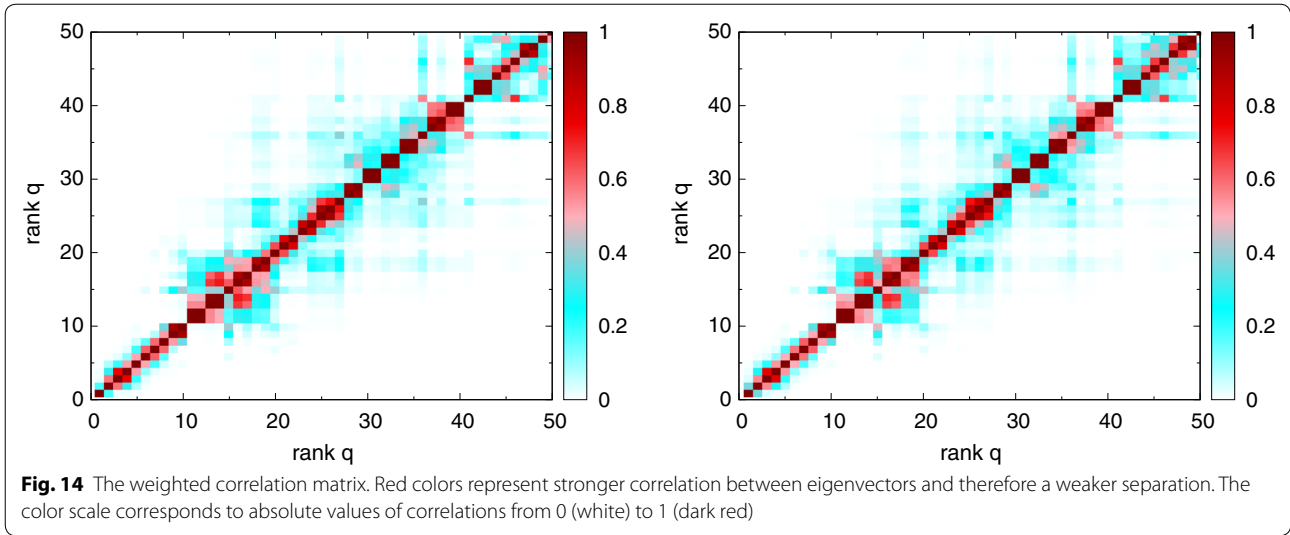
The w-correlation matrix (Fig. 14) is found to be diagonal symmetric, and we mainly identify two different regions. The first region relates to the first 11 eigenvalues and their modes, where correlations quickly diminish away from the central diagonal. In the second region, from the 11th eigenvector onward, off-diagonal correlations emerge. This suggests that principal

components do not fully separate. However, these correlations are minor and those principal components may still carry important information. This is almost identical for toroidal and poloidal flow coefficients.

The forecast of the flow variability requires to identify significant and non-noisy components. There are several possibilities to quantify the statistical significance of the temporal principal components. In the next section we develop a rule that provides a criterion for the selection of a sufficient correct time-series model of the flow variability.

**Model selection criterion**

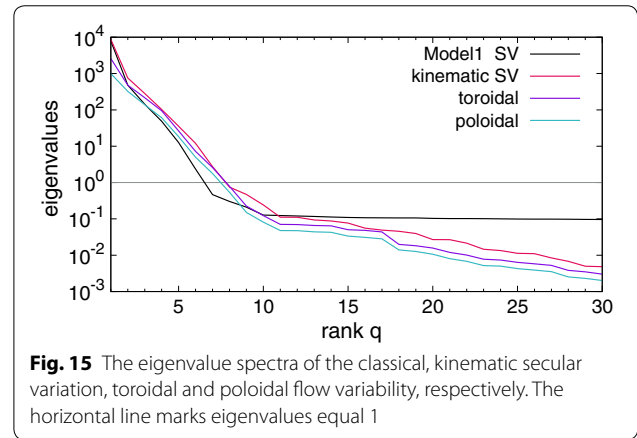
Generally, to what extent the variability of the time series can be explained strongly depends on  $k$ , the



number of model parameter. Therefore, a selection criterion is applied which ascertains the statistical significance of the eigenvalues and their related EOF detected by the MSSA. The recovery of the temporal variability of the time series can be considered to be most complete, when the maximum number  $k$  of EOF's (truncation level) are used for the reconstruction of the time series. However, this is for several reasons not desirable, for instance some modes may represent noise that is part of the time series (and we want a noise-free representation of the flow temporal variability).

Figure 15 shows the eigenvalue spectra of different analyses of the secular variation and the toroidal and poloidal flow variability. Broadly, the first 7 eigenvalues are  $\geq 1$ . A first break in the spectral slopes occurs at  $k \sim 10$ , which is related to features of the  $w$ -correlation matrices shown in Fig. 14. A second break in slope occurs at  $k \sim 15$ . Thereafter, spectra of the kinematic secular variation and the flow variability continuously decrease. Based on these results a MSSA-model truncated at a level  $k > 15$  should provide a sufficient reconstruction of the temporal flow variability. However, in the kinematic forecast, we use  $k = 22$ , at this value past and predicted flow variability become continuous. This may indicate significant portion of the signal to be carried by eigenmodes for  $15 < k < 23$ .

The eigenvalue spectrum of the secular variation of Model 1 breaks at eigenvalue 11. Thereafter, it is flat, but eigenvalues remain an order of magnitude larger, than those of the kinematic secular variation and the flow coefficients.



### Appendix B: Kinematic field models

Characteristics of kinematic field models for a range of damping parameters are given in Table 3. We explored purely toroidal (x0a) and tangential geostrophic (x3\*) flow assumptions, respectively. The parameter controlling the frozen-flux constraint is set to  $\lambda_1 = 1 \times 10^{-9}$  and  $\lambda_1 = 1 \times 10^{-7}$ . Different setups of the other constraints are also tested, as well as different temporal derivatives of the temporal norm. We tested

$$\begin{aligned}
 T = 0 &: \int_T \int_{S(c)} (\mathbf{u})^2 d\Omega dt \\
 T = 1 &: \int_T \int_{S(c)} (\partial_t \mathbf{u})^2 d\Omega dt.
 \end{aligned}
 \tag{B.1}$$

**Table 3 Specification of damping parameters of the frozen-flux field model based on the tangential geostrophic flow assumption, except model x0a, which is based on a purely toroidal flow assumption**

	$\lambda_s$	$\lambda_T$	$\lambda_{TG}$	$\lambda_E$	S/T	R	M
x0a	$1 \times 10^{-3}$	$1 \times 10^{-4}$	–	$1 \times 10^2$	0/1	$0.957 \times 10^{-3}$	2.1184
x3a	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^3$	$1 \times 10^2$	0/1	$0.791 \times 10^{-3}$	2.1196
x3b	$1 \times 10^{-3}$	$1 \times 10^{-5}$	$1 \times 10^3$	$1 \times 10^2$	0/1	$0.941 \times 10^{-3}$	2.1185
x3c	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^4$	$1 \times 10^2$	0/1	$0.123 \times 10^{-2}$	2.1233
x3d	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^5$	$1 \times 10^2$	0/1	$0.102 \times 10^{-2}$	2.1237
x3e	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^6$	$1 \times 10^2$	0/1	$0.166 \times 10^{-3}$	2.1431
x3f	$1 \times 10^{-3}$	$1 \times 10^{-3}$	$1 \times 10^4$	$1 \times 10^2$	0/1	$0.774 \times 10^{-3}$	2.1322
x3g	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^{10}$	$1 \times 10^2$	0/1	$0.136 \times 10^{-2}$	2.1627
x3g1 <sup>†</sup>	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^{10}$	$1 \times 10^2$	0/1	$0.431 \times 10^{-3}$	2.1856
x3h	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^{12}$	$1 \times 10^2$	0/1	$0.106 \times 10^{-2}$	2.1632
x3j	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^8$	$1 \times 10^2$	0/0	$0.311 \times 10^{-2}$	2.1695
x3k	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^7$	$1 \times 10^2$	0/0	$0.696 \times 10^{-2}$	2.1634
x3k1 <sup>†</sup>	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^7$	$1 \times 10^2$	0/0	$0.332 \times 10^{-2}$	2.2221
x3l	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^6$	$1 \times 10^2$	0/0	$0.129 \times 10^{-1}$	2.1525
x3m	$1 \times 10^{-3}$	$1 \times 10^{-4}$	$1 \times 10^4$	$1 \times 10^2$	0/0	$0.524 \times 10^{-3}$	2.1396

The parameter controlling the frozen-flux constraint is kept fixed, i.e.  $\lambda_1 = 1 \times 10^{-9}$  for most of the inversions, except those marked with<sup>†</sup>, there we set  $\lambda_1 = 1 \times 10^{-7}$ . In all inversions, the parameter controlling the flow acceleration at model end points is set to  $\lambda_4 = 1 \times 10^2$ . S/T identifies which temporal derivative of the spatial and temporal norm is applied, respectively (see text). R and M are defined by (26) and (25), respectively. M is given in nT/year. The model x0a represents a field model with purely toroidal flow

Received: 30 April 2020 Accepted: 19 August 2020

Published online: 22 October 2020

## References

- Abramowitz M, Stegun IA (1973) Handbook of mathematical functions. Dover, New York
- Alken P, Thébaud E, Finlay CC, Beggan CD, Alken P, Aubert J, Barrois O, Bertrand F, Bondar T, Boness A, Brocco L, Canet E, Chambodut A, Chulliat A, Coisson P, Civet F, Du A, Fournier A, Fratter I, Gillet N, Hamilton B, Hamoudi M, Hulot G, Jager T, Korte M, Kuang W, Lalanne X, Langlais B, Léger J-M, Lesur V, Lowes FJ, Macmillan S, Manda M, Manoj C, Maus S, Olsen N, Petrov V, Ridley V, Rother M, Sabaka TJ, Saturnino D, Schachtschneider R, Sirol O, Tangborn A, Thomson A, Toffner-Clausen L, Vigneron P, Wardinski I, Zvereva T (2020) International geomagnetic reference field: the thirteenth generation. *Earth Planets Space* 67:81. <https://doi.org/10.1186/s40623-020-01288-x>
- Allen MR, Robertson AW (1996) Distinguishing modulated oscillations from coloured noise in multivariate datasets. *Clim Dyn* 12:775–784
- Amit H (2014) Can downwelling at the top of the Earth's core be detected in the geomagnetic secular variation? *Phys Earth Planet Inter* 229:110–121
- Armstrong JS (1985) From crystal ball to computer, 2nd edn. Wiley, New York
- Aubert J (2015) Geomagnetic forecasts driven by thermal wind dynamics in the Earth's core. *Geophys J Int* 203(3):1738–1751
- Aubert J, Finlay CC (2019) Geomagnetic jerks and rapid hydromagnetic waves focusing at Earth's core surface. *Nat Geosci* 12(5):393–398
- Barrois O, Hammer MD, Finlay CC, Martin Y, Gillet N (2018) Assimilation of ground and satellite magnetic measurements: inference of core surface magnetic and velocity field changes. *Geophys J Int* 215(1):695–712
- Beggan C, Whaler K (2010) Forecasting secular variation using core flows. *Earth Planets Space* 62(10):821–828. <https://doi.org/10.5047/eps.2010.07.004>
- Bloxham J (1988) The dynamical regime of fluid flow at the core surface. *Geophys Res Lett* 15:585–588
- Bloxham J, Jackson A (1992) Time-dependent mapping of the magnetic field at the core-mantle boundary. *J Geophys Res* 97:19537–19563
- Bloxham J, Zatman S, Dumberry M (2002) The origin of geomagnetic jerks. *Nature* 420(6911):65–68
- Box GEP, Jenkins GM (1976) Time series analysis, forecasting and control. Holden Day, San Francisco
- British Geological Survey - Edinburgh (2020) Worldwide observatory annual means. [http://www.geomag.bgs.ac.uk/data\\_service/data/annual\\_means.shtml](http://www.geomag.bgs.ac.uk/data_service/data/annual_means.shtml)
- Brockwell PJ, Davis RA (2002) Introduction to time series and forecasting, 2nd edn. Springer, Berlin
- Broomhead DS, King GP (1986) Extracting qualitative dynamics from experimental data. *Phys D Nonlinear Phenomena* 20:217–236
- Brown WJ, Mound JE, Livermore PW (2013) Jerks abound: an analysis of geomagnetic observatory data from 1957 to 2008. *Phys Earth Planet Inter* 223:62–76
- Canciani A, Raquet J (2016) Absolute positioning using the earth's magnetic anomaly field. *Navigation* 63(2):111–126
- Chambodut A, Schwarte J, Langlais B, Lühr H, MM (2002) The selection of data in field modeling. In: Stauning P, Lühr H, Ulfré-Guérard P, LaBrecque J, Purucker M, Primdahl F, Jørgensen J, Christensen F, Høeg P, Lauritsen K (eds) In: OIST- Proceedings, 4th Oersted international science team conference. Danish Meteorological Institute Copenhagen Denmark, pp 201–213
- Chambodut A, Marchaudon A, Menvielle M, El-Lemdani Mazouz F, Lathuillere C (2013) The k-derived MLT sector geomagnetic indices. *Geophys Res Lett* 40(18):4808–4812
- Chen Z, Yang Y (2004) Assessing forecast accuracy measures. available on Researchgate
- Finlay CC, Maus S, Beggan CD, Bondar TN, Chambodut A, Chernova TA, Chulliat A, Golovkov VP, Hamilton B, Hamoudi M, Holme R, Hulot G, Kuang W, Langlais B, Lesur V, Lowes FJ, Lühr H, MacMillan S, Manda M, McLean S, Manoj C, Menvielle M, Michaelis I, Olsen N, Rauberg J, Rother M, Sabaka TJ, Tangborn A, Toffner-Clausen L, Thébaud E, Thomson AWP, Wardinski I, Wei Z, Zvereva TI (2010) International geomagnetic reference field: the eleventh generation. *Geophys J Int* 183:1216–1230
- Finlay CC, Aubert J, Gillet N (2016) Gyre-driven decay of the Earth's magnetic dipole. *Nat Commun* 7:10422

- Flores BE (1986) A pragmatic view of accuracy measurement in forecasting. *Omega* 14(2):93–98
- 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. <https://doi.org/10.1186/s40623-015-0245-8>
- Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME, Robertson AW, Saunders A, Tian Y, Varadi F, Yiou P (2002) Advanced spectral methods for climatic time series. *Rev Geophys* 40:1–41
- Golyandina N, Nekrutkin V, Zhigljavsky A (2001) Analysis of time series structure: SSA and related techniques. Chapman & Hall, Routledge
- Hamilton B, Ridley VA, Beggan CD, Macmillan S (2015) The BGS magnetic field candidate models for the 12th generation IGRF. *Earth Planets Space* 67:69. <https://doi.org/10.1186/s40623-015-0227-x>
- Hills RG (1979) Convection in the Earth's mantle due to viscous shear at the core-mantle interface and due to large-scale buoyancy, Ph.D. thesis, New Mexico State University, Las Cruces
- Holme R (2007) Chap 4 Large-scale flow in the core. In: Olson P (ed) *Core dynamics. Treatise on Geophysics*, vol 8. Elsevier, Amsterdam, pp 107–130
- Holme R, de Viron O (2005) Geomagnetic jerks and a high-resolution length-of-day profile for core studies. *Geophys J Int* 160:435–440
- Holme R, de Viron O (2013) Characterization and implications of intradecadal variations in length of day. *Nature* 499(7457):202–204
- Hulot G, Lhuillier F, Aubert J (2010) Earth's dynamo limit of predictability. *Geophys Res Lett* 37:6305
- International Service of Geomagnetic indices (2020) <http://isgi.unistra.fr>
- Jackson A (1997) Time-dependency of tangentially geostrophic core surface motions. *Phys Earth Planet Inter* 103:293–311
- Jackson A, Bloxham J, Gubbins D (1993) Time-dependent flow at the core surface and conservation of angular momentum in the coupled core-mantle system. In: LeMouél J-L, Smylie DE, Herring T (eds) *Dynamics of the Earth's deep interior and Earth rotation*. pp 97–107. AGU/IUGG
- Jault D, Gire C, LeMouél JL (1988) Westward drift, core motions and exchanges of angular momentum between core and mantle. *Nature* 333:353–356
- Jiménez AR, Seco F, Zampella F, Prieto JC, Guevara J (2012) Improved heuristic drift elimination with magnetically-aided dominant directions (MiHDE) for pedestrian navigation in complex buildings. *J Locat Based Serv* 6(3):186–210
- 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(10):775–785. <https://doi.org/10.5047/eps.2010.07.008>
- LeMouél J-L (1984) Outer core geostrophic flow and secular variation of Earth's magnetic field. *Nature* 311:734–735
- Lesur V, Wardinski I, Asari S, Minchev B, Manda M (2010) Modelling the Earth's core magnetic field under flow constraints. *Earth Planets Space* 62:503–516. <https://doi.org/10.5047/eps.2010.02.010>
- Lesur V, Wardinski I, Baerenzung J, Holschneider M (2018) On the frequency spectra of the core magnetic field Gauss coefficients. *Phys Earth Planet Inter* 276:145–158
- Livermore PW, Hollerbach R, Finlay CC (2017) An accelerating high-latitude jet in Earth's core. *Nat Geosci* 10(1):62–68
- Lloyd D, Gubbins D (1990) Toroidal fluid motion at the top of the Earth's core. *Geophys J Int* 100:455–467
- Manda M, Olsen N (2006) A new approach to directly determine the secular variation from magnetic satellite observations. *Geophys Res Lett* 33:15306
- Manda M, Holme R, Pais A, Pinheiro K, Jackson A, Verbanac G (2010) Geomagnetic jerks: rapid core field variations and core dynamics. *Space Sci Rev* 155(1–4):147–175
- Mathews BP, Diamantopoulos H (1994) Towards a taxonomy of forecast error measures: a factor-comparative investigation of forecast error dimensions. *J Forecast* 13:409–416
- Olson P, Amit H (2006) Changes in earth's dipole. *Naturwissenschaften* 93:519–542
- Pais MA, Jault D (2008) Quasi-geostrophic flows responsible for the secular variation of the Earth's magnetic field. *Geophys J Int* 173:421–443
- Pais MA, Oliveira O, Nogueira F (2004) Nonuniqueness of inverted core-mantle boundary flows and deviations from tangential geostrophy. *J Geophys Res* 109:8105
- Plaut G, Vautard R (1994) Spells of Low-Frequency Oscillations and Weather Regimes in the Northern Hemisphere. *J Atmos Sci* 51:210–236
- Purucker ME, Sabaka TJ, Langel RA (1996) Conjugate gradient analysis: a new tool for studying satellite magnetic data sets. *Geophys Res Lett* 23:507–510
- Roberts PH, Scott S (1965) On the analysis of secular variation, 1, A hydromagnetic constraint: theory. *J Geomag Geoelectr* 17:137–151
- Saturnino D, Langlais B, Amit H, Civet F, Manda M, Beucler E (2018) Combining virtual observatory and equivalent source dipole approaches to describe the geomagnetic field with swarm measurements. *Phys Earth Planet Inter* 276:118–133
- Slavinskis A, Kvell U, Kulu E, Sünter I, Kuuste H, Lätt S, Voormansik K, Noorma M (2014) High spin rate magnetic controller for nanosatellites. *Acta Astron* 95:218–226
- Terra-Nova F, Amit H, Hartmann GA, Trindade RIF, Pinheiro KJ (2017) Relating the South Atlantic Anomaly and geomagnetic flux patches. *Phys Earth Planet Inter* 266:39–53
- Thébault E, Finlay CC, Beggan CD, Alken P, Aubert J, Barrois O, Bertrand F, Bondar T, Boness A, Brocco L, Canet E, Chambodut A, Chulliat A, Coisson P, Civet F, Du A, Fournier A, Fratter I, Gillet N, Hamilton B, Hamoudi M, Hulot G, Jager T, Korte M, Kuang W, Lalanee X, Langlais B, Léger J-M, Lesur V, Lowes FJ, Macmillan S, Manda M, Manoj C, Maus S, Olsen N, Petrov V, Ridley V, Rother M, Sabaka TJ, Saturnino D, Schachtschneider R, Sirol O, Tangborn A, Thomson A, Toffner-Clausen L, Vigneron P, Wardinski I, Zvereva T (2015) International Geomagnetic Reference Field: the 12th generation. *Earth Planets Space* 67:79. <https://doi.org/10.1186/s40623-015-0228-9>
- Toffner-Clausen L, Nielsen JB (2018) Swarm Level 1b Product Definition, SW-RS-DSC-SY-0007, Issue 5.22. [https://earth.esa.int/documents/10174/1514862/Swarm\\_L1b\\_Product\\_Definition](https://earth.esa.int/documents/10174/1514862/Swarm_L1b_Product_Definition)
- Torta JM, Pavón-Carrasco FJ, Marsal S, Finlay CC (2015) Evidence for a new geomagnetic jerk in 2014. *Geophys Res Lett* 42(19):7933–7940
- Vautard R, Yiou P, Ghil M (1992) Singular-spectrum analysis: a toolkit for short, noisy chaotic signals. *Phys D Nonlinear Phenomena* 58:95–126
- Wardinski I, Holme R (2006) A time-dependent model of the earth's magnetic field and its secular variation for the period 1980 to 2000. *J Geophys Res* 111:12101
- Wardinski I, Holme R (2011) Signal from noise in geomagnetic field modelling: denoising data for secular variation studies. *Geophys J Int* 185:653–662
- Wardinski I, Lesur V (2012) An extended version of the C<sup>3</sup>FM geomagnetic field model: application of a continuous frozen-flux constraint. *Geophys J Int* 189:1409–1429
- Wardinski I, Holme R, Asari S, Manda M (2008) The 2003 geomagnetic jerk and its relation to the core surface flows. *Earth Planet Sci Lett* 267:468–481
- Whaler KA (1980) Does the whole of the Earth's core convect? *Nature* 287:528–530
- Whaler KA, Beggan CD (2015) Derivation and use of core surface flows for forecasting secular variation. *J Geophys Res* 120(3):1400–1414
- World Data Centre for Geomagnetism - Edinburgh (2019) Geomagnetic Data Master Catalogue. <http://www.wdc.bgs.ac.uk/catalog/master.html>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.