Earth, Planets and Space

FULL PAPER

Open Access

Kalmag: a high spatio-temporal model of the geomagnetic field

Baerenzung Julien[1]* , Holschneider Matthias[2], Jan Saynisch-Wagner[1] and Maik Thomas[1]

Abstract: We present the extension of the Kalmag model, proposed as a candidate for IGRF-13, to the twentieth century. The dataset serving its derivation has been complemented by new measurements coming from satellites, ground-based observatories and land, marine and airborne surveys. As its predecessor, this version is derived from a combination of a Kalman filter and a smoothing algorithm, providing mean models and associated uncertainties. These quantities permit a precise estimation of locations where mean solutions can be considered as reliable or not. The temporal resolution of the core field and the secular variation was set to 0.1 year over the 122 years the model is spanning. Nevertheless, it can be shown through ensembles a posteriori sampled, that this resolution can be effectively achieved only by a limited amount of spatial scales and during certain time periods. Unsurprisingly, highest accuracy in both space and time of the core field and the secular variation is achieved during the CHAMP and Swarm era. In this version of Kalmag, a particular effort was made for resolving the small-scale lithospheric field. Under specific statistical assumptions, the latter was modeled up to spherical harmonic degree and order 1000, and signal from both satellite and survey measurements contributed to its development. External and induced fields were jointly estimated with the rest of the model. We show that their large scales could be accurately extracted from direct measurements whenever the latter exhibit a sufficiently high temporal coverage. Temporally resolving these fields down to 3 hours during the CHAMP and Swarm missions, gave us access to the link between induced and magnetospheric fields. In particular, the period dependence of the driving signal on the induced one could be directly observed. The model is available through various physical and statistical quantities on a dedicated website at https://ionocovar.agnld.uni-potsdam.de/Kalmag/.

Keywords: Geomagnetic field, Lithospheric field, Secular variation, Magnetospheric field, Induced field, Assimilation, Kalman filter, Machine learning

---

*Correspondence: baerenzung@gmx.de
[1] Section 1.3: Earth System Modelling, Geoforschungszentrum Potsdam, Potsdam, Germany
Full list of author information is available at the end of the article

Graphical abstract
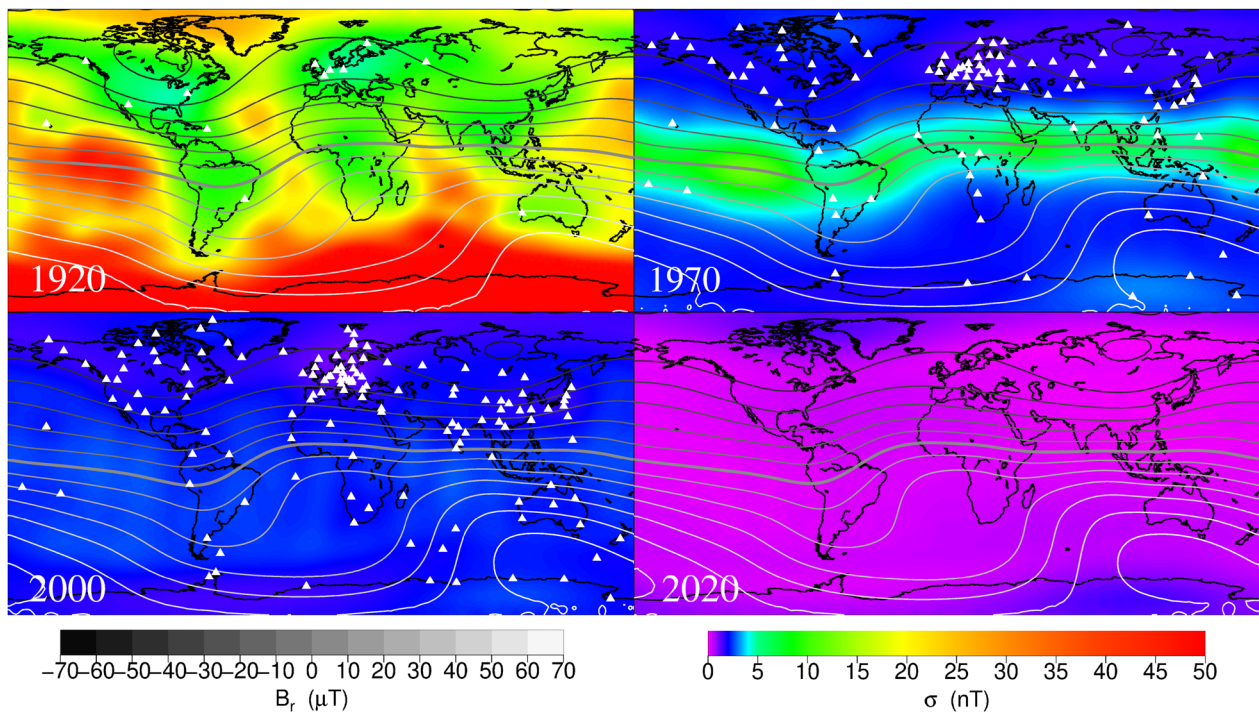
Introduction

Separating the different contributions to the Earth's magnetic field from direct measurements of it is a difficult task. The main reason making this problem complex is the wide range of spatial and temporal scales overlapping one another. The core field, which is sustained by dynamo action in the Earth's outer core, is at the Earth's surface the dominant large-scale field, and it evolves on timescale ranging from months to millennia. On the opposite, the lithospheric field is dominant at small scales. Emanating from the remnant magnetization of the rocks lying within the crust, it follows the motions of the latter and therefore varies very slowly with time. External sources, such as the magnetospheric fields or the ionospheric field are driven by thermospheric winds and solar radiations. Their direct link to solar activity make them subject to intense variations from very short up to decadal timescales. These fluctuations induce currents within the electrically conducting parts of the crust and the mantle which in return generate a secondary magnetic field. Induction processes also occurs within the oceans. The circulation or tidal motions of the latter within the ambient magnetic field create electrical currents which also produce a secondary field.

From the seventeenth century to today, geomagnetic data have been continuously accumulated. First collected during marine and land surveys, measurements of the Earth's magnetic field were quickly complemented by instrumentation installed within ground-based observatories [see Jackson and Finlay (2007)]. The development of aviation in the 1950s offered another support to measure the field. But the biggest step in geomagnetic monitoring certainly comes from the rise of low-orbiting satellite missions. Starting in 1965 with the POGO mission, many spacecrafts dedicated to geomagnetic field modeling were later launched. These include non-exhaustively the Mag-Sat, the Oersted and the CHAMP spacecrafts and the Swarm constellation.

Technical constraints to build geomagnetic field models strongly depend on the type of data to be assimilated. Satellite missions provide measurements at a high frequency. The algorithms they are feeding therefore need to be adapted to treat a large amount of observations. Land, marine and airborne surveys operate at the level or slightly above the Earth's surface. As a consequence, the contribution of the small-scale lithospheric field to the

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 3 of 22

data they produce is important. Accounting for this field requires to model it at a very high resolution, a technical challenge.

Many models of the geomagnetic field have been proposed over the last decades [see Hulot et al. (2015)], and most of them were obtained with a regularized least square approach. This is the case for the CHAOS model series from Olsen et al. (2006) to Finlay et al. (2020), the comprehensive models by Sabaka et al. (2002, 2015, 2018, 2020), the GRIMM models by Lesur et al. (2008, 2010, 2015), the POMME models by Maus et al. (2005, 2010), or the gufm1 by Jackson et al. (2000). Least square methods are very efficient numerically, but the usually considered reweighed version can only provide unique solution. On the opposite, Bayesian inversions are computationally demanding but results are expressed in terms of posterior distributions, providing therefore predictions of mean solutions together with their associated uncertainties. Bayesian inversion in the context of geomagnetic field modeling was initiated by Gillet et al. (2013). Considering ground-based observatory, survey and satellite data, they could derive the COV-OBS model spanning the $1840 - 2010$ time window, a period which was recently increased to 2020 by Huder et al. (2020). Similar efforts have been followed by Holschneider et al. (2016) in a study where emphasis was put on better characterizing the spatial properties of the different magnetic sources through correlation kernels. Extending this work to the time domain, and sequentializing the problem, Baerenzung et al. (2020), Ropp et al. (2020) could derive geomagnetic field models from the combination of a Kalman filter and a smoothing algorithm. This approach conserves all the advantages of the Bayesian method proposed by Gillet et al. (2013) and alleviates most of its drawbacks. In particular, the dimension of the system, the amount of observations to be assimilated, or the non linear link between certain magnetic sources, are not anymore strong limiting factors.

In this paper, we present the extension of the Kalmag model by Baerenzung et al. (2020) to the twentieth century. Deriving only from CHAMP and Swarm data, Kalmag covered the $2000.5 - 2020$ time period, and was a candidate for the IGRF-13 model [see Alken et al. (2021)]. The present version resulted from the assimilation of extra measurements taken by ground-based observatories, POGO, MagSat and Oersted satellites and during land, airborne and marine (L.A.M.) surveys. To assimilate the latter type of data, we introduced a statistical approximation within the Kalman filter algorithm enabling us to resolve the lithospheric field up to spherical harmonics degree and order $\ell = 1000$. Therefore,

there is no need to subtract the lithospheric contribution to L.A.M. survey observations with high-resolution models such as the WDMAM by Lesur et al. (2016), the EMM model by Maus (2010) or the recent model of Thébault et al. (2021), to build the model. In addition, a small-scale lithospheric field model could be recovered without preprocessing of the data.

The article is organized as follows. In the first part, the dataset used to construct the model and the selection criteria applied are presented. In the second part, the different magnetic sources, their prior characterization and dynamical behavior are detailed. At the end of this section, the various formulations to assimilate data, update the model and sample it are provided. In "Results" section, the properties of our model for the core field, the secular variation, the lithospheric field and external and induced fields are discussed. The article ends with a discussion and some concluding remarks.

Data
The proposed model was derived from either vector field or intensity measurements of the geomagnetic field taken from 1900.0 to today by satellites, ground-based observatories and during land, airborne and marine surveys. Satellite observations from five different missions were considered. These are, the POGO (1965–1971) (e.g., Cain and Sweeney 1973), the MagSat (1979–1980) (e.g., Langel and Estes 1985a), the Oersted (since 1999) (e.g., Neubert et al. 2001), the CHAMP (2000-2010) (e.g., Rother et al. 2000), and the SWARM (since 2013) (e.g., Olsen et al. 2013) missions. For ground-based observatories, hourly mean vector fields provided by the World data center for geomagnetism from 1886 (e.g., Macmillan and Olsen 2013) and selected through the procedure which is detailed in the following, were used to derive secular variation data, the latter being used only to constrain the core field evolution. These types of observations, feeding also other models such as the CHAOS series by Olsen et al. (2006), Finlay et al. (2020), the C3FM by Wardinski and Holme (2011), Wardinski et al. (2020) or the COV-OBS model by Gillet et al. (2013), Huder et al. (2020) were here obtained by first averaging vector field measurements over 0.1-year time windows. The resulting mean values $\bar{\mathbf{b}}(t)$ were then used to derive secular variation data $\gamma(t)$ through the relation $\gamma(t) = \bar{\mathbf{b}}(t + 0.5yr) - \bar{\mathbf{b}}(t - 0.5yr)$. The location of each observatory taken into account is displayed with black triangles in Fig. 1. For aeromagnetic, land and marine survey data, three compilations served the model derivation (e.g., Quesnel et al. 2009). The first one is provided by British Geological Survey at www.wdc.bgs.ac.uk/, the second one by the National Oceanic and Atmospheric Administration at maps.ngdc.noaa.
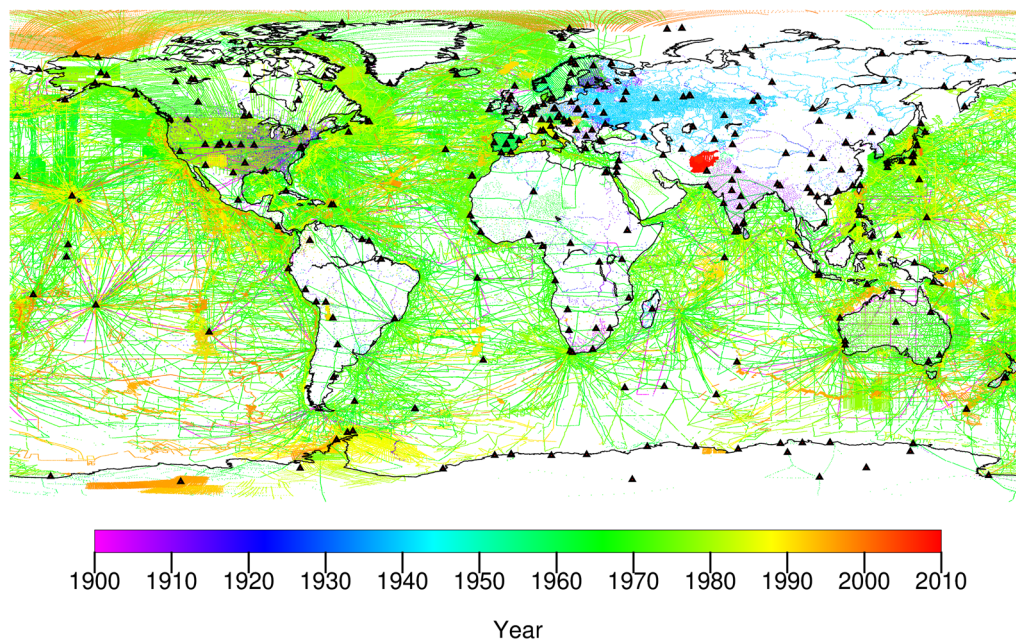
Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 4 of 22

Fig. 1 Locations (dots) and epoch (color) of each land, airborne and marine survey measurement. Black triangles correspond to every ground-based observatories feeding the model with data

gov and the third one is made accessible by the U.S. geophysical survey at www.mrdata.usgs.gov. The positions of L.A.M. data are typically given through the latitude, longitude and altitude location of the measuring vessel. For airborne measurements, whenever altitude was provided by radar altimeter it was corrected above land surfaces with the ETOPO1 global relief model of Amante and Eakins (2009).

Before being assimilated, each data containing vector information, such as North, East, Down or declination, inclination and intensity, was projected in geographic spherical coordinates. The resulting dataset was then subject to selection. The main purposes of this procedure are to avoid the contribution of the dayside ionospheric field which is not modeled, to operating during low geomagnetic activity and, for satellite observations, to be weakly perturbed by the substorm auroral electrojet. The latter two criteria were fulfilled through a selection based on the values of independently derived indices, respectively, a given threshold on the Kp geomagnetic index and the required positiveness of the z-component of the interplanetary magnetic field (IMF). The Kp threshold was set to 2^- for satellite data and to 4^- for all other observations. To limit the contribution of the dayside ionospheric field, only night-time measurements (when the sun is below the horizon)

were kept at magnetic latitude lying between $\pm 60°$. This constraint was nevertheless relaxed for MagSat data for which the satellite followed a dawn–dusk orbit and for some land survey data which were either not dated precisely enough to determine their local solar time, or only used to derive the lithospheric field model. Note also that for CHAMP and SWARM satellites, it was also required that measurements were taken when both the vector field magnetometer and the star tracker were functioning in nominal mode.

Finally, each L.A.M. surveys and satellite dataset were subsampled. For POGO, MagSat, and Oersted satellites, a rate of 1 datum every 10s (0.1Hz) was chosen. For CHAMP satellite, the sampling rate was increased to 0.2Hz. For SWARM, only satellites Alpha and Bravo are considered with a simultaneous sampling rate of 0.1Hz. Distance criteria were applied to subsample L.A.M. surveys data. In a first selection, a minimum distance of 5 km between any data point within 1-h time windows was imposed. Every measure lying too close to the previously selected ones were removed. The resulting dataset was then split in 8 subsets in which the minimum distance was set to 40 km. Therefore, at a given epoch within the Kalman filter algorithm, data from each of these subsets were sequentially assimilated whenever they were available.

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 5 of 22

Table 1 Dataset used to derive the model. Missions (first column) and their time span (second column)

Mission	Period	Selection type			# data	
		Night time	K_p	IMF $B_z > 0$	Vector components	Intensity
Observatories	1900.0 − 2019.5	×	$K_p < 4^o$		$3 \times 74,410$	
L.A.M. surveys	1900.0 − 2009.0	×	$K_p < 4^o$		$3 \times 132,502$	$2,106,816$
Pogo	1965.7 − 1971.4	×	$K_p < 2^o$	×		$1,405,585$
MagSat	1979.8 − 1980.4		$K_p < 2^o$	×	$3 \times 102,459$	
Oersted	1999.2 − 2000.5	×	$K_p < 2^o$	×	$3 \times 152,705$	
CHAMP	2000.5 − 2010.7	×	$K_p < 2^o$	×	$3 \times 6,103,759$	
Swarm	2013.8 − 2022.2	×	$K_p < 2^o$	×	$3 \times 5,843,961$	
Total	1900 − 2022.2				$3 \times 12,409,799$	$3,512,401$

Selection criteria applied the data: night-time (third column), K_p threshold (fourth column) and positiveness of the z-component of the IMF (fifth column). Number of vector field components (sixth column) and intensity measurements (seventh column)

Percentage of rejected data through the (eighth column)

In Table 1, the time period, the selection criteria and the type and total number of measurements associated with each dataset are summarized.

Magnetic sources

Seven sources compose the Kalmag model. These are a core field (b_c), a lithospheric field (b_l), an induced/residual ionospheric field (b_{ii}), a remote (b_{rm}), a close (b_m) and a fluctuating (b_{fm}) magnetospheric fields and a source associated with field-aligned currents (b_{fac}). Except for b_{fac}, each of these sources b_s is assumed to derive from a potential V_s such as $b_s = -\nabla V_s$. For b_{fac}, as in Sabaka et al. (2004) the currents themselves are assumed to derive from a potential V_{fac}. Waters et al. (2001) has shown that under this assumption the resulting magnetic field could be expressed as $b_s = -\mathbf{r} \times \nabla V_s$.

The potentials V_s are then expanded in spherical harmonics (SH) such as potentials of internal and external origin are, respectively, given by:

$$V_s^I(r, \theta_s, \phi_s, t)$$
$$= a_s \sum_{\ell \leq \ell_{max}} \sum_{m=-m_{max}}^{m=m_{max}} \left(\frac{a_s}{r}\right)^{l+1} g_{s,\ell,m}^I(t) Y_{\ell,m}(\theta_s, \phi_s),$$
(1)

$$V_s^E(r, \theta_s, \phi_s, t) = a_s \sum_{\ell \leq \ell_{max}} \sum_{m=-m_{max}}^{m=m_{max}} \left(\frac{r}{a_s}\right)^{l} g_{s,\ell,m}^E(t) Y_{\ell,m}(\theta_s, \phi_s),$$
(2)

Where $Y_{\ell,m}$ are Schmidt semi-normalized spherical harmonics of degree ℓ and order m considered, respectively, up to ℓ_{max} and m_{max}, a_s is a reference radius, and $g_{s,\ell,m}(t)$ (later referred as g_s) are the spherical harmonics coefficients expressed at a_s. Each field is projected in a given spherical coordinate system $\{r, \theta_s, \phi_s\}$ as indicated in

Table 2 Magnetic sources considered in the model (first column) together with the coordinate systems they are expressed in (second column)

Source	Coordinate	ℓ_{max}	m_{max}
Core g_c	GEO	20	ℓ_{max}
Lithospheric g_l	GEO	$100^a, 150^b, 1000^c$	ℓ_{max}
Remote magnetospheric g_{rm}	GSM	1	0
Close magnetospheric g_m	SM	15	1
Fluctuating magnetospheric g_{fm}	SM	15	0
Residual ionospheric/ induced g_{ii}	MAG	50	1
Field-aligned currents g_{fac}	SM	15	1

GEO geographic, *SM* solar magnetic, *MAG* magnetic, *GSM* geocentric solar magnetospheric

ℓ_{max} and m_{max} are, respectively, the maximum degree and order of the SH expansion

[a] POGO, MagSat, Oersted data.

[b] CHAMP, Swarm data.

[c] L.A.M. surveys data

Table 2. These systems can either be geographic (GEO), magnetic (MAG), solar magnetic (SM), or geocentric solar magnetospheric (GSM) (see Laundal (2017)).

Depending on the observations which are being assimilated, the spatial resolution of the lithospheric field is varied. Whereas for CHAMP and Swarm data, the latter is expanded up to $\ell = 150$, it is only modeled up to $\ell = 100$ for other satellite measurements. Since L.A.M. survey data are taken close to the Earth's surface, they contain a strong contribution of the small scale lithospheric field. To assimilate such measurements, the lithospheric field is therefore parameterized up to spherical harmonics degree $\ell = 1000$ with an approximation of the associated covariance matrix between $100 < \ell \leq 1000$ as detailed in the following.

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 6 of 22

Sequential modeling

The Kalmag model is constructed sequentially through a Kalman filter approach [see Kalman (1960)]. This technique proceeds in two alternating steps, namely a forecast and an analysis. In the forecast, the model is propagated in space and time until some measurements become available. Then the analysis takes place and the model is updated accordingly to them. Because this method provides the posterior distribution of the model only given the previously assimilated data, it is complemented by a smoothing algorithm. Performing backward in time, this algorithm enables us to correct the model at any time according to the complete dataset.

Dynamical model

The spatio-temporal evolution of the various sources composing the geomagnetic field is of complex nature. Involving nonlinear couplings, a large range of spatial and temporal scales, some regimes which are not yet numerically achievable or simply not sufficiently well characterized, the dynamics of the Earth's magnetic field cannot be directly simulated. This is why, as initiated by Gillet et al. (2013) in the context of geomagnetic modeling, we chose simplified stochastic equations, namely auto-regressive processes (or ARPs), to predict the evolution of the different fields. Mimicking dispersion and memory effects occurring within dynamical systems, such processes are computationally cheap to simulate and are formulated within a Gaussian framework as required by the Kalman filter approach. A priori, each source is characterized by its own process which is independent from the others. As shown in Appendix A, ARPs in their sequential form, can be described by the following general relation:

$$z_s(t + \Delta t) = F_s(\Delta t)z_s(t) + \xi_i(t, \Delta t), \quad (3)$$

where z_s is a quantity characterizing the sth source to be propagated, $F_s(\Delta t)$ is the parameter of the ARP and $\xi_i(t, \Delta t)$ is a temporal Gaussian white noise spatially characterized by the distribution $\mathcal{N}\left(0, \Sigma_{z_s}^\infty - F_s\Sigma_{z_s}^\infty F_s^T\right)$, where $\Sigma_{z_s}^\infty$ is the stationary state covariance matrix associated with z_s. Except for the lithospheric field which is assumed to be static, and for the core field which evolution is prescribed by a second-order process, the dynamics of each source is controlled by a first-order ARP. In this case, $z_s(t)$ simply corresponds to the vector of SH coefficients $g_s(t)$ associated with the sth field and the parameter of the process is given by:

$$F_s(\ell, \Delta t) = \exp\left[-|\Delta t|/\tau_s(\ell)\right], \quad (4)$$

where $\tau_s(\ell)$ is a parameterized scale-dependent characteristic time which is specified for each source in the following. For the core field, the use of a second-order ARP induces a coupling between the field itself (g_c) and its first time derivative ($\partial_t g_c$). Therefore, $z_c = (g_c, \partial_t g_c)^T$ and the parameter of the process is given by:

$$F_c(\ell, \Delta t) = \begin{pmatrix} 1 + |\Delta t|/\tau_c(\ell) & \Delta t \\ -\Delta t/\tau_c^2(\ell) & 1 - |\Delta t|/\tau_c(\ell) \end{pmatrix} \\ \exp\left[-|\Delta t|/\tau_c(\ell)\right], \quad (5)$$

where $\tau_c(\ell)$ is also chosen to be scale dependent. Contrary to first-order ARPs where the stationary state covariance matrices are given by $\Sigma_{z_s}^\infty = \Sigma_{g_s}^\infty$, for the core field it reads:

$$\Sigma_{z_c}^\infty = \Sigma_{g_c, \partial_t g_c}^\infty = \begin{pmatrix} \Sigma_{g_c}^\infty & 0 \\ 0 & \Sigma_{g_c}^\infty/\tau_c^2(\ell) \end{pmatrix}, \quad (6)$$

as shown by Hulot and Le Mouël (1994). With the proposed setup, $\Sigma_{z_s}^\infty$ and $\tau_s(\ell)$ completely defines the dynamical behavior of the ARPs. The covariance matrices characterizing the stationary state of each source are assumed to derive from energy spectra $E_s(\ell, a_s)$ expressed at given radii a_s such as:

$$\Sigma_{g_s}^\infty(\ell, m, \ell', m', r = a_s) = \frac{E_s(\ell, a_s)}{N_m R(\ell)}\delta(\ell - \ell')\delta(m - m'), \quad (7)$$

where N_m is the number of modeled spherical harmonics coefficients per degree ℓ, and R is given by $R(\ell) = \ell + 1$ and $R(\ell) = \ell$ for internal and external sources, respectively. The shape of each energy spectrum is imposed. It can either be flat, such as $E_s(\ell) = A_s^2$ or identical to the correlation kernels proposed by Holschneider et al. (2016) that we refer as of C-based type with $E_s = A_s^2(2\ell + 1)R(\ell)$, where A_s is the magnitude of the spectrum. For most sources, the dipole part is assumed to be independent from the rest of the spectrum such as $E_s(\ell = 1) = D_s^2$. Under these assumptions, the radii a_s, the amplitudes A_s and the dipole magnitudes D_s form the free parameters of the stationary state covariance matrices $\Sigma_{z_s}^\infty$. Characteristic timescales are parameterized by power laws such as $\tau_s(\ell) = M_s\ell^{-\alpha_s}$ with given magnitudes (M_s) and slopes (α_s) which are for some sources allowed to continuously vary from one range of spherical harmonics to the other. The ARP's parameters were estimated through a machine learning algorithm with a subsample of CHAMP and Swarm data as detailed in Baerenzung et al. (2020). The same values are used in this study. They are reported in Table 3.

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 7 of 22

Table 3 Magnetic sources parameters as described in "Dynamical model" section and evaluated by Baerenzung et al. (2020)

Field	Spectrum	Radius a (km)	A (nT)	M	α
Core	Flat	3456	D: 1.12×10^5 9.74×10^4	$\tau_c(1)$: 935 yrs $M(\ell \geq 2) = 514$ yrs	1.06
Lithospheric $1 \leq \ell \leq 74$	C-based	6287	0.16	∞	0
$75 \leq \ell \leq 1000$	Flat	6367.9	6.5	∞	0
Close magnetospheric	C-based	12524	D: 9.16 1.88	$\tau_m(1)$: 1.54 days $M(\ell \geq 2) = 18$ min	0
Remote magnetospheric	C-based	235570	7.3	10.31 yrs	0
Fluctuating magnetospheric	C-based	13028	D: 3 4.56	$\tau_{fm}(1)$: 0.36 day $\tau_{fm}(2)$: 0.55 days $M(\ell \geq 3) = 4$ days	1.15
Residual ionospheric/ induced	Flat	6324	D: 5.48 4.39	$\tau_s(1)$: 0.71 day $M(\ell \geq 2) = 1.76$ day	0.93
Field-aligned currents	C-based	7917	D: 0 1.22	$\tau_{fac}(1)$: 0 $M(\ell \geq 2) = 1$ min	0

The prior spatial covariance matrices are derived from energy spectra expressed at some radii a_s which are either flat with $E_s^\infty(\ell) = A_s^2$ or of the C-based type [see Holschneider et al. (2016)] with the form $E_s^\infty(\ell) = A_s^2(2\ell + 1)R(\ell)$, where $R(\ell) = \ell + 1$ and $R(\ell) = \ell$ for, respectively, internal and external sources

The characteristic timescales are parameterized by $\tau_s(\ell) = M_s\ell^{-\alpha_s}$

Note that here, the energy spectrum of the lithospheric field is split into two ranges. In the first one, between $\ell = 1$ and $\ell = 74$, the spectrum is of the C-based type and exhibits a characteristic radius of $a_l = 6287$ km and a magnitude of $A_l = 0.16$ nT. These are values obtained by Baerenzung et al. (2020). In the second range, between $\ell = 75$ and $\ell = 1000$, the spectrum is flat with $a_l = 6367.9$ km and $A_l = 6.5$ nT. In this case, the parameters were estimated through a least square fit between $\ell = 75$ and $\ell = 400$ of the energy spectrum associated with the WDMAM model of Lesur et al. (2016).

The source associated with field-aligned currents, as well as the components at SH degree larger than $\ell = 1$ of the fluctuating magnetospheric field, exhibit very small characteristic timescales of, respectively, $\tau_{fac}(\ell) = 1$ min and $\tau_{fm}(\ell > 1) = 18$ min. These timescales being smaller than the time step of the Kalman filter algorithm (here set to 30 min), the associated fields are assumed to temporally evolve as a white noise but are correlated in space and time during the analysis. Setting a priori a zero mean for both fields their covariance can be expressed as:

$$E[g_s(\ell, t)g_s(\ell', t + \Delta t)] = \Sigma_s^\infty(\ell) \exp\left[-|\Delta t|/\tau_s(\ell)\right]\delta(\ell - \ell') . \tag{8}$$

Filtering, smoothing, sampling

The prior statistical properties as well as the dynamics of the different magnetic sources being characterized, assimilation can be initiated. As a first step, a vector $\mathbf{z}$ containing the spherical harmonics coefficients of each field is constructed. For the full model, $\mathbf{z}$ is composed of $N_M = 1002696$ entries. The lithospheric field which is expanded up to $\ell = 1000$ is filling more than 99.9% of the $\mathbf{z}$ vector. With such a model dimension, the size of the covariance matrix associated with $\mathbf{z}$, namely $\Sigma_\mathbf{z}$, should be of $N_M \times N_M \sim 10^{12}$. Yet computations with such a matrix would be numerically impossible. This is why we approximate the predicted uncertainties of the small scale lithospheric field (for $101 \leq \ell \leq 1000$) by only keeping its variance information (the one associated with each of its spherical harmonics coefficients). Under such an assumption the dimension of $\Sigma_\mathbf{z}$ reduces to $N_M \times N_M \sim 10^8$, a computationally conceivable size. This strong approximation, which induces a complete loss of the predicted spatial correlations of the lithospheric field beyond $\ell = 100$ is evaluated in "Lithospheric field" section.

To forecast $\mathbf{z}$, each parameter matrix F_s of equation 3 are incorporated in a global matrix $\mathbf{F}$. The same operation is performed for the stationary state covariance matrices Σ^∞ which are assembled into the covariance matrix $\mathbf{\Sigma}^\infty$. Given $\mathbf{F}$ and $\mathbf{\Sigma}^\infty$, the covariance matrix associated with the Gaussian white noise of the full model forecast step reads $\tilde{\mathbf{\Sigma}} = \mathbf{\Sigma}^\infty - \mathbf{F}\mathbf{\Sigma}^\infty\mathbf{F}^T$. Therefore, the evolution of the mean model and its covariance from time step $k - 1$ to step k is then given by:

$$E[\mathbf{z}_{k|k-1}] = \mathbf{F}_{k-1}E[\mathbf{z}_{k-1}] \tag{9}$$

$$\mathbf{\Sigma}_{\mathbf{z}_{k|k-1}} = \mathbf{F}_{k-1}\mathbf{\Sigma}_{\mathbf{z}_{k-1}}\mathbf{F}_{k-1}^T + \tilde{\mathbf{\Sigma}} . \tag{10}$$

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 8 of 22

After the forecast, whenever measurements are available, the model is updated accordingly. This operation is performed through a Bayesian inversion which reads:

$$\mathbf{K}_k = \mathbf{\Sigma}_{\mathbf{z}_{k|k-1}} \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{\Sigma}_{\mathbf{z}_{k|k-1}} \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1}, \tag{11}$$

$$E[\mathbf{z}_{k|\mathbf{d}_k}] = E[\mathbf{z}_{k|k-1}] + \mathbf{K}_k \left(d_k - \mathbf{H}_k E[\mathbf{z}_{k|k-1}] \right), \tag{12}$$

$$\mathbf{\Sigma}_{\mathbf{z}_{k|\mathbf{d}_k}} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{\Sigma}_{\mathbf{z}_{k|k-1}}, \tag{13}$$

where $\mathbf{R}_k$ is the covariance matrix associated with measurement errors, $\mathbf{K}_k$ is the Kalman gain matrix and $\mathbf{H}_k$ is the operator projecting the model to the observations d_k at iteration k. $\mathbf{R}_k$ is chosen to be diagonal with constant standard deviations of 0.1 nT for intensity data [see Quesnel et al. (2009)] and vector field measurements, and of 4.85 nT/yr for each component of secular variation data as we estimated it with a similar algorithm used to calibrate Kalmag (see Baerenzung et al. (2020)). When d_k corresponds to intensity measurements, the linearization approach proposed by Mauerberger et al. (2020); Schanner et al. (2022) is applied. In their developments, they showed that at first order, the predicted intensity I_k could be related to the predicted magnetic field $\mathbf{B}_k$ through the relation $I_k \sim E[\mathbf{B}_k]^T \mathbf{B}_k / \tilde{I}_k$, where $\tilde{I}_k$ is the intensity derived from the mean magnetic field $E[\mathbf{B}_k]$. Note that this projection is realized with the mean magnetic field prediction. Therefore, no iteration over the updated solutions is required. When d_k correspond to secular variation data, $\mathbf{H}_k = \mathbf{0}$ for each source except for the core field where $\mathbf{H}_k$ projects its associated secular variation on the data. Once all data have been assimilated, the different modeled epochs are corrected through a smoothing algorithm (see Rauch et al. (1965)). Starting at the final step of the Kalman filter it performs iteratively backward in time through the following relations:

$$\mathbf{G}_{k-1} = \mathbf{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}_{k-1}}} \mathbf{F}_k^T \mathbf{\Sigma}_{\mathbf{z}_{k|k-1}}^{-1}, \tag{14}$$

$$E[\mathbf{z}_{k-1|\mathbf{d}}] = E[\mathbf{z}_{k-1|\mathbf{d}_{k-1}}] + \mathbf{G}_{k-1} \left(E[\mathbf{z}_{k|\mathbf{d}}] - E[\mathbf{z}_{k|k-1}] \right), \tag{15}$$

$$\mathbf{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}}} = \mathbf{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}_{k-1}}} + \mathbf{G}_{k-1} \left(\mathbf{\Sigma}_{\mathbf{z}_{k|\mathbf{d}}} - \mathbf{\Sigma}_{\mathbf{z}_{k|k-1}} \right) \mathbf{G}_{k-1}^T, \tag{16}$$

where $\mathbf{d}$ corresponds to the full dataset. The smoothing algorithm only provides snapshots of the posterior model, therefore the resulting solution does not contain information about temporal correlations. Although the posterior covariance between the model at different epochs can be analytically derived, obvious storage limitations makes this option numerically inapplicable. This

is why we introduced a formulation to sample ensembles from the posterior model which are correlated both in space and time. Starting with an ensemble $\mathbf{z}^e$ randomly drawn from the last state of the Kalman filter solution, the algorithm proceeds similarly to the smoothing algorithm, backward in time with:

$$\mathbf{z}_{k-1|\mathbf{d}}^e = E[\mathbf{z}_{k-1|\mathbf{d}_{k-1}}] + \mathbf{G}_{k-1} \left(\mathbf{z}_{k|\mathbf{d}}^e - E[\mathbf{z}_{k|k-1}] \right) + \zeta^e, \tag{17}$$

where ζ^e is a random realization from the Gaussian distribution characterized by a $\mathbf{0}$ mean and a covariance matrix given by:

$$\mathbf{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}}}^E = \mathbf{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}_{k-1}}} - \mathbf{G}_{k-1} \left(\mathbf{\Sigma}_{\mathbf{z}_{k|k-1}} \right) \mathbf{G}_{k-1}^T. \tag{18}$$

Note that to correct deviations, due to sampling errors, between the ensembles and the true posterior means, the ensembles were recentered at each epochs accordingly to the mean smoothing solutions. For this study, we used an ensemble of 1024 members.

Model construction

To construct the model, the time step of the Kalman filter algorithm was set to $\Delta t = 30$ min. Nevertheless, whenever the distance between two analysis windows exceeded this value, Δt was increased accordingly. With a dataset covering the 20th century and the last 22 years, the direct approach would have been to start assimilating measurements in 1900.0 and to progress forward in time until today. However, we did not proceed this way. Instead, the Kalman filter simulation was initiated in 2000.5 to first assimilate ground-based observatories, CHAMP and then Swarm data until 2022.18, the last epoch at which measurements were currently available. The smoothing algorithm was then applied to update the model within this time window. In a third part, the smoothing solution in 2000.5 was used as a restart file to assimilate, backward in time, the measurements taken prior to this date. Finally, the smoothing algorithm was applied from 1900 to today with a slight modification beyond 2000.5 which is detailed in Appendix B.

Two reasons motivated this choice of splitting the assimilation process. The first one is to possess a well-resolved large-scale lithospheric field before assimilating survey data. This, in order to be able to distinguish the gain of assimilating such observations on this part of the field. To this end, the lithospheric field was fully modeled up to SH degree $\ell = 150$ during the CHAMP and Swarm eras. The full solution (mean and covariance) was then truncated at $\ell = 100$ in 2000.5 to restart the Kalman filter between 2000.5 and 1900.0. Beyond $\ell = 100$ only

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 9 of 22

the mean and variance were kept from the CHAMP and Swarm solution. This part was finally extrapolated with a zero mean and the prior variance of equation 7 between $\ell = 150$ and $\ell = 1000$. The second reason for splitting the assimilation process was motivated by the fact that the older the measurements, the lower their accuracy and spatial coverage. Yet, before assimilating any measurement an outlier detection is performed. The latter process consists in checking that the measurements do not excessively deviate from their predicted values, in particular that each vector field or intensity measure lies within the 95.6% confidence interval of the model prediction. On top of this selection, the misfit of the sequentially assimilated tracks was evaluated. Whenever the misfit value exceeded the imposed threshold of 3, the corresponding track was dismissed. The algorithm to detect outliers performs better when the model accuracy is high, which occurs when the data quality and coverage are good. This

is why starting with a very well constrained solution in 2000.5 and assimilating data backward in time enabled us to optimize the detection process.

Over the entire time span of the model, each source, except the one associated with field-aligned currents, is stored every 0.1 year, setting up the temporal resolution of the model to this time step. However, to better track the evolution of rapidly evolving sources, such as the close and fluctuating magnetospheric fields or the residual ionospheric/induced fields, the latter were stored every 3 hours during the CHAMP and Swarm eras and every 5 days between 1900 and 2000.5.

Results

Main field and secular variation

The Kalman filter and smoothing algorithms provide a model in terms of mean solution and associated covariance matrix. Combining these two quantities gives a precise
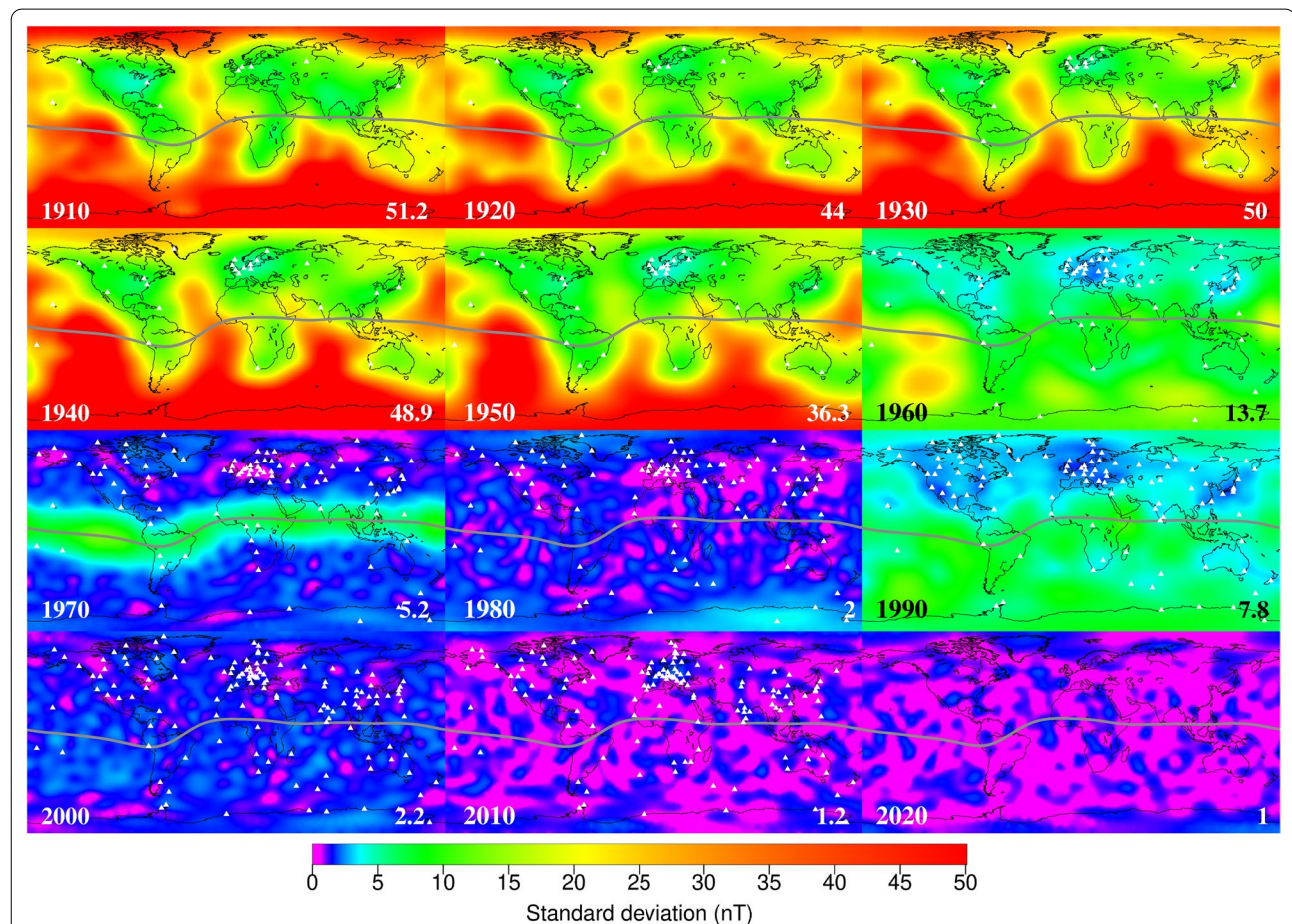

Fig. 2 Standard deviation associated with the radial component at the Earth's surface of the sum of the core field and the lithospheric field expanded up to spherical harmonics degree $\ell = 20$. Each panel corresponds to a different epoch which is displayed on their bottom left. Isocontours show the mean $B_r = 0$ solution. White triangles represent the locations of ground-based observatories available at the presented epochs. On the bottom right of each map is indicated the r.m.s standard deviation $\bar{\sigma}$ in nT

Julien *et al. Earth, Planets and Space* (2022) 74:139

Page 10 of 22

knowledge of locations where the solution is reliable and where it is not. As an illustration for the main field, i.e., the sum of the core field and the lithospheric expanded up to SH degree $\ell = 20$, Fig. 2 shows at different epochs the radial component of the mean field (isocontours) and its associated standard deviation (color maps). Locations where the maps are red correspond to locations where the mean solution is likely to deviate strongly from the true field. On the opposite, within blue and purple areas the model predicts that the true and the mean predicted field are close. These maps are complemented on their bottom right by a global measure of the predicted uncertainty. It corresponds to the r.m.s. standard deviation given in nT and expressed as:
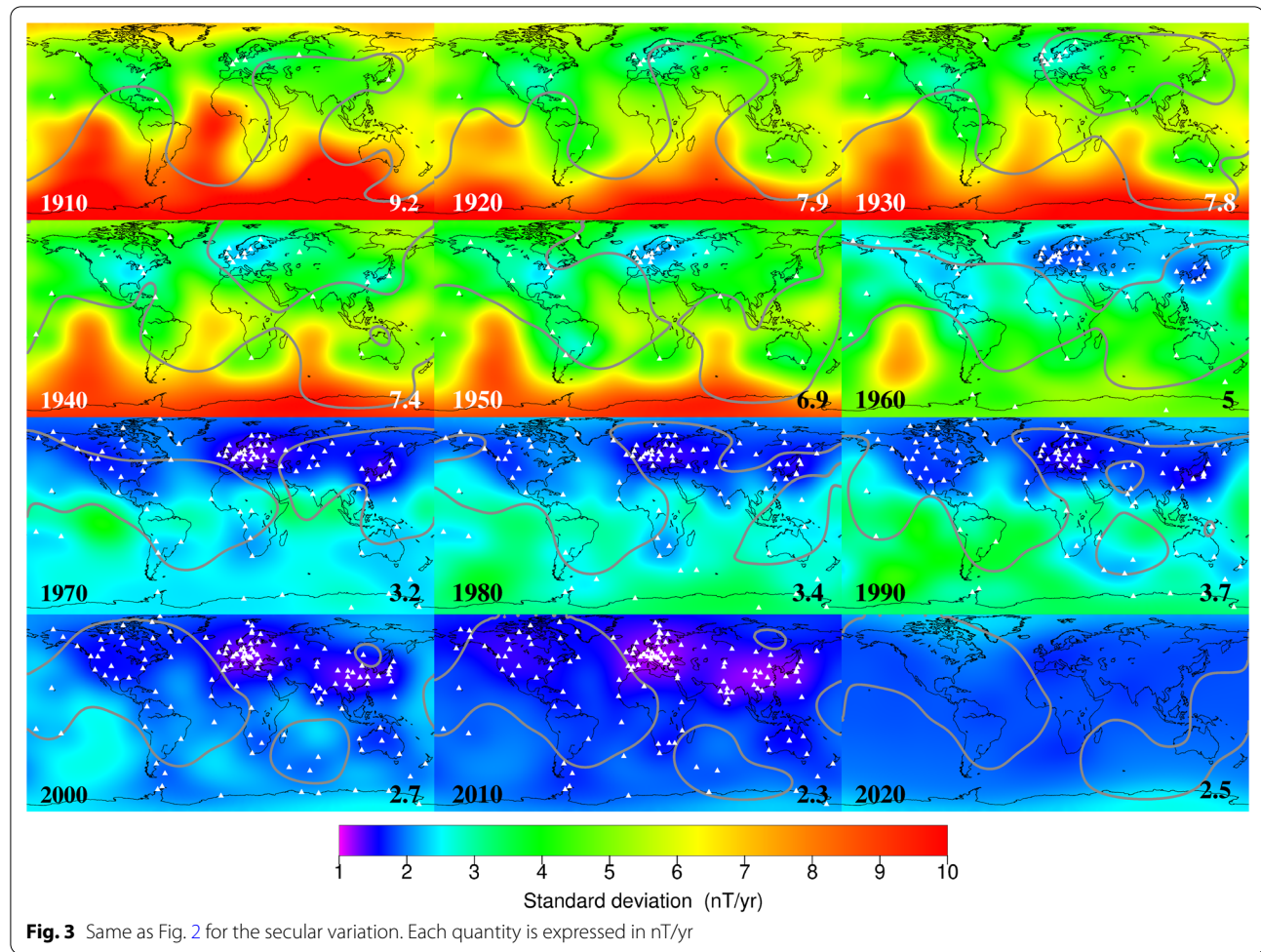
$$\bar{\sigma} = \sqrt{\int_\Omega \sigma^2 d\Omega / \int_\Omega d\Omega}, \qquad (19)$$

where σ is the standard deviation associated with the radial component of the field and Ω is the Earth's surface.

Until the 1960s, uncertainty maps exhibited a strong dichotomy between the Northern and Southern hemispheres. Whereas in the North, the standard deviation associated with the radial component of the field does not globally excess 25 nT, it reaches and even exceeds 50 nT in the South. The difference of predicted uncertainties is particularly important between land and oceanic surfaces reflecting the lack of measurements taken over the latter, the location where the field is best resolved is Europe. This is a benefit of the high density of ground-based observatories operating at this place and during this time period. When looking at the r.m.s. standard deviation, the year 1920 slightly stands out with $\bar{\sigma} = 44$ nT, whereas this value oscillates around $\bar{\sigma} \sim 50$ nT in 1910, 1930 and 1940. This phenomenon can be explained by the multiple land and marine surveys occurring at and around this epoch and which are offering a large data coverage of the globe (see Fig. 1). In 1960, the global resolution of the model is improved and the North–South dichotomy mostly disappears. Two reasons explain this gain of accuracy. The first one is the dense spatial coverage of survey data at this epoch (see Fig. 1). The second one is the time proximity of the POGO mission which started in 1965. One can also observe that observatories still play an important role to reduce the posterior variability as it is the case in and around Europe and Japan. In 1970, the jump in accuracy of the model is striking. At this period lying within the POGO era, the standard deviation associated with the Kalmag solution is strongly reduced. However, the model predicts a higher possible variability around the magnetic dip equator. This phenomenon is the transcription of the

Backus effect, or more generally the "perpendicular error" effects within the model. Indeed, as first recognized by Backus (1970), to be then generalized by Lowes (1975), when constructing a geomagnetic field model with intensity measurements alone, larger errors will contaminate the model near the equator. This effect is surely affecting our mean solution, but covariance information enables us to quantify it. With MagSat observations, which cover less than a year ($1979 - 1980$), the model precision is equivalent to the one obtained with POGO data except around the dip equator where vector field measurements eliminate the "perpendicular error" effects induced by the assimilation of intensity data. The map in 1990 highlights the importance of low-orbiting satellites to recover the Earth's magnetic field. Lying between MagSat and Oersted missions, in the middle of almost 20 years without satellite measurements, the solution obtained at this time is strongly degraded. It presents levels of uncertainties equivalent to the 1960 ones except in Northern America and Russia where the coverage with ground-based observatories has since been increased. The situation is ameliorated with Oersted measurements and becomes even better with CHAMP and Swarm observations. With the high-quality instrumentation of CHAMP and Swarm satellites, the model is extremely precise and this is almost everywhere at the Earth's surface. It is however worth noting that the constellation of Swarm satellites permits to obtain a slightly more accurate solution than the unique CHAMP spacecraft.

When looking at the mean secular variation (SV) and its associated standard deviation as displayed at similar epochs in Fig. 3, one can observe that the dichotomy in accuracy between the North and the South is also present for this quantity. The dichotomy persists until the year 2000, but with a lower contrast after 1960. Ground-based observatory data are of particular importance to constrain the secular variation, as locations where their density is high always coincide with areas of low posterior variability. Globally, uncertainties are decreasing with time except between 1970 and 2000, where the r.m.s. standard deviation fluctuates due to the lack of persistent low-orbiting satellite missions. In addition, the distribution of uncertainties over the different spatial scales is not homogeneous. Instead, small scales typically exhibit a higher posterior variability relatively to their mean signal than large scales. This effect can be observed in Fig. 4 where time series between 1900 and 2022 of the 68.2% confidence interval associated with some selected SH coefficients are displayed in red. In this figure, it is clearly visible that the larger the degree of the coefficient (from left to right and top to bottom), the larger its posterior standard deviation relatively to its mean

Julien *et al. Earth, Planets and Space* (2022) 74:139

Page 11 of 22

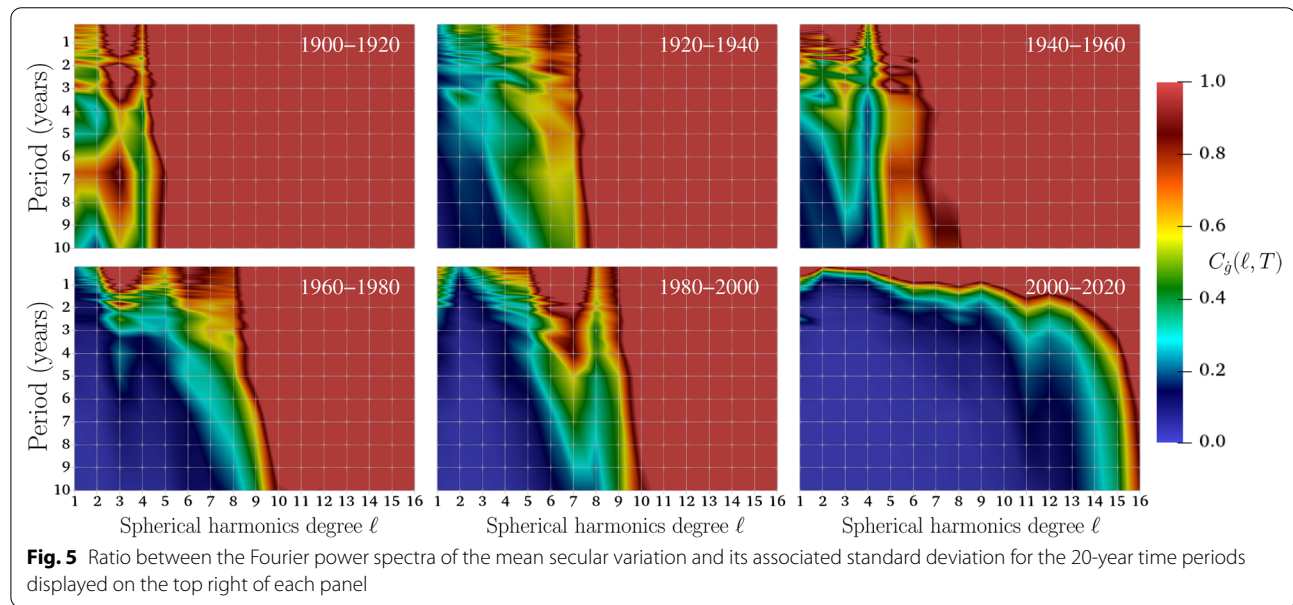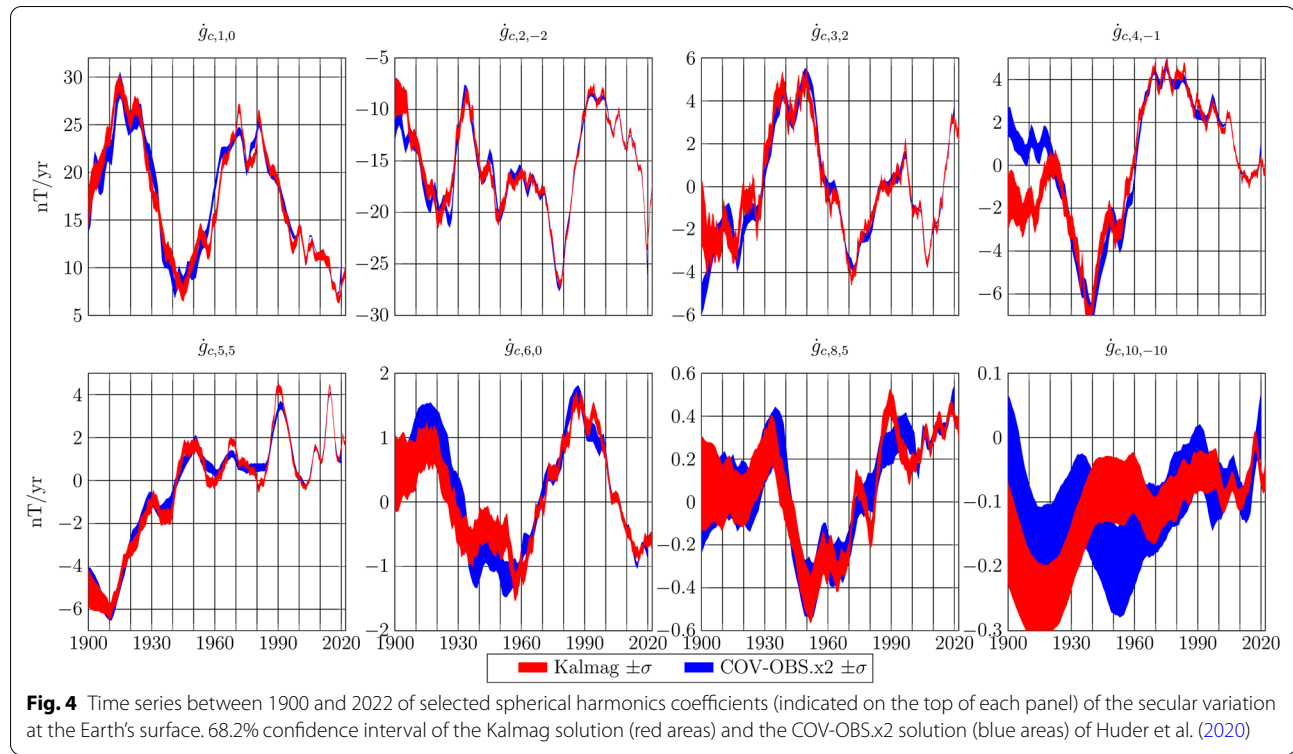**Fig. 3** Same as Fig. 2 for the secular variation. Each quantity is expressed in nT/yr

values. The COV-OBS.x2 model of Huder et al. (2020), exhibits a similar behavior as its predicted 68.2% confidence intervals (blue areas) show. Although the two models are mostly consistent with one another, small differences can nevertheless be distinguished, in particular in the predicted standard deviations. Until ∼ 1920 their level is lower for COV-OVS.x2, they become equivalent between COV-OVS.x2 and Kalmag until ∼ 1960 to be lower for Kalmag afterwards.

To precisely characterize the spatio-temporal resolution of the secular variation over the model time span, we computed the ratio $C_{\dot{g}}(\ell, k)$ between the Fourier power spectra of the mean secular variation and its associated standard deviation for 20 years time periods. This quantity, which was proposed by Gillet et al. (2015), can be expressed as:

$$C_{\dot{g}}(\ell, k) = \sum_{m=-\ell}^{m=l} \left\| E\left[\hat{\dot{g}}_{c,\ell,m}(k)\right] \right\|^2 \Big/ \sum_{m=-\ell}^{m=l} \left\| \sigma_{\hat{\dot{g}}_c,\ell,m}(k) \right\|^2,$$

(20)

where $\hat{\dot{g}}_{c,\ell,m}(k)$ is the Fourier transform of the secular variation, and $\sigma_{\hat{\dot{g}}_c,\ell,m}(k)$ is its associated standard deviation. To estimate the latter quantity, we used an ensemble of 1024 Fourier transform of secular variation time series. In Fig. 5, $C_{\dot{g}}(\ell, k)$ is displayed for 6 different time windows. The blue and red areas correspond to spatio-temporal scales which are, respectively, well resolved and not resolved. At early times, between 1900 and 1920, only some limited amount of temporal scales of the SV up to SH degree $\ell = 4$ are resolved. The situation slightly improves between 1920 and 1960 where some signal up to SH $\ell = 6$ can be accurately recovered, and this down to a few years for the largest spatial scales. The emergence of satellite missions and the increase of ground-based observatory and survey data helps improving the model resolution between 1960 and 2000. During this time interval some spherical harmonics coefficient up to degree $\ell = 5$ are either partially or fully resolved down to

Julien *et al. Earth, Planets and Space* (2022) 74:139

Page 12 of 22

Fig. 4 Time series between 1900 and 2022 of selected spherical harmonics coefficients (indicated on the top of each panel) of the secular variation at the Earth's surface. 68.2% confidence interval of the Kalmag solution (red areas) and the COV-OBS.x2 solution (blue areas) of Huder et al. (2020)

Fig. 5 Ratio between the Fourier power spectra of the mean secular variation and its associated standard deviation for the 20-year time periods displayed on the top right of each panel

time periods lower than a year. Reaching such a temporal resolution is impossible with the secular variation data derived from annual differences of observatory measurements. It can therefore only be achieved thanks to the high temporal coverage of satellite and survey data. In agreement with our previous results and with the study

of Gillet (2019), the secular variation is best resolved during the CHAMP and Swarm eras, where spatial scale up $\ell = 15$ can be partially resolved down to periods of approximately 5 years, and 2-year fluctuations can be very well captured up to $\ell = 10$.
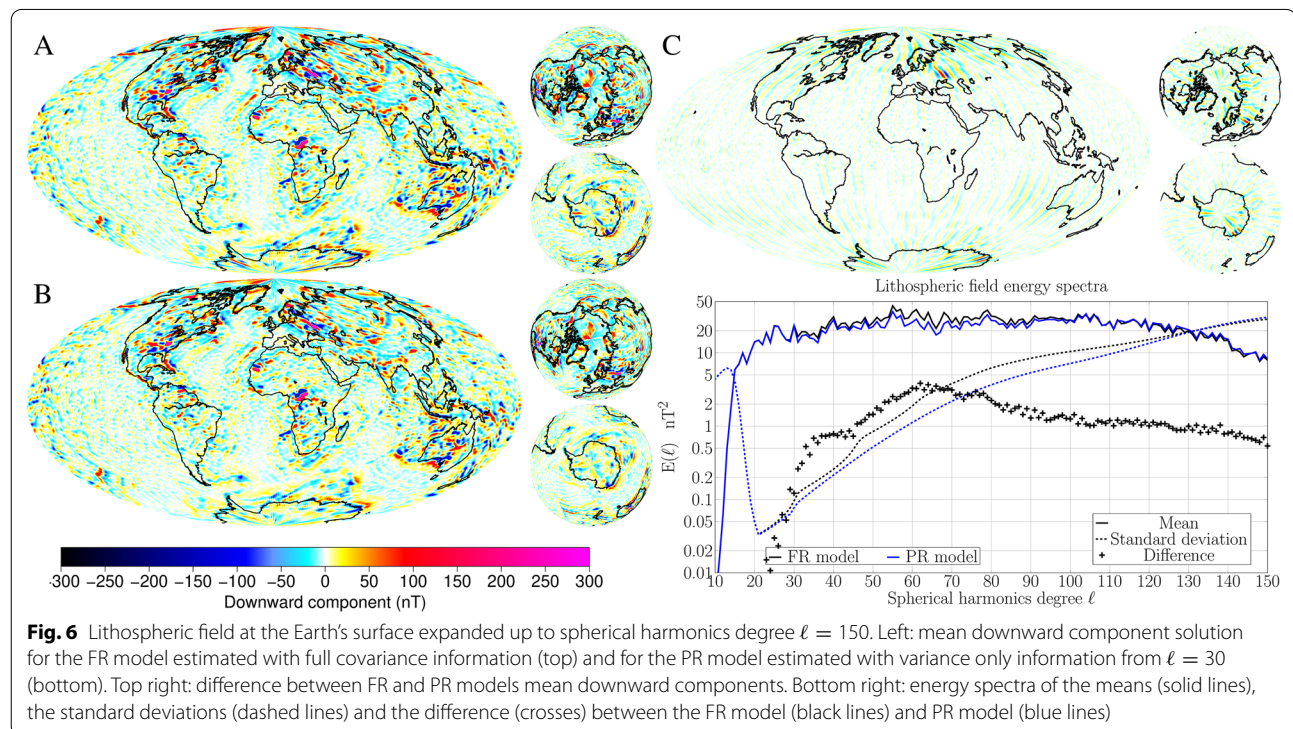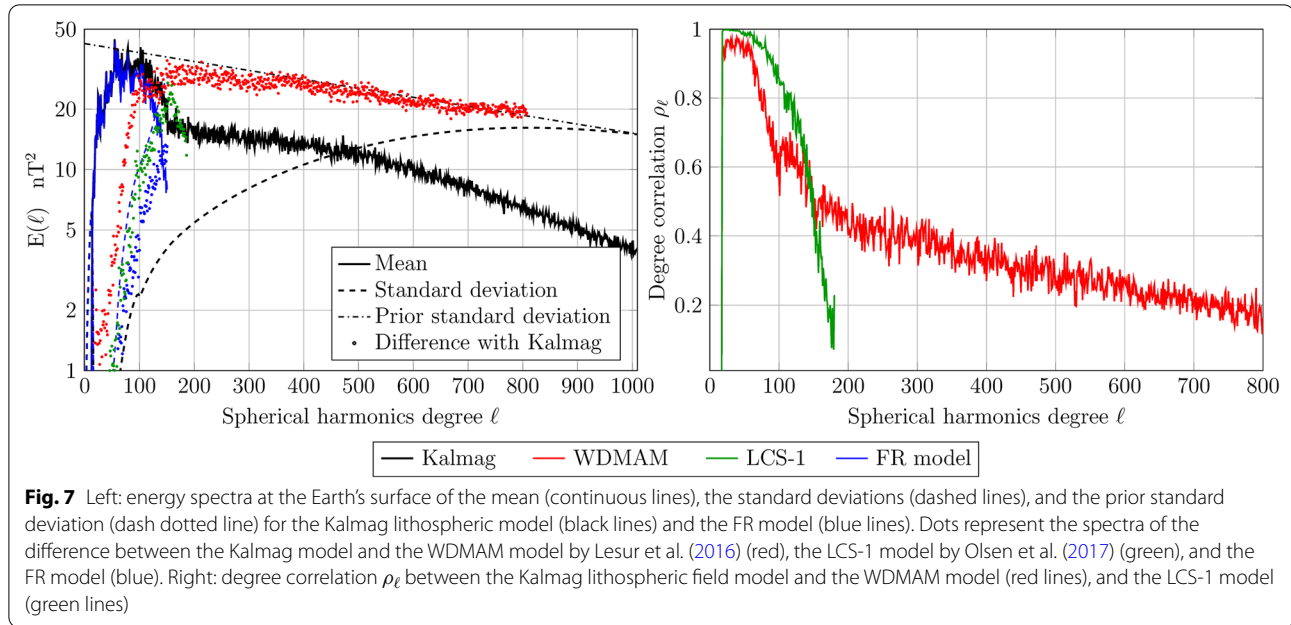
Lithospheric field

As previously mentioned, the lithospheric field model was built in multiple steps. During the CHAMP and Swarm eras, it was fully modeled up to SH degree $\ell = 150$. After applying the smoothing algorithm, the lithospheric in 2000.5 was divided in three parts. In the first one, between $\ell = 1$ and $\ell = 100$, the full smoothing solution (mean and associated covariance matrix) was kept. In the second part, between $\ell = 101$ and $\ell = 150$, only mean and variance information were considered. Finally, between $\ell = 151$ and $\ell = 1000$ a zero mean and the variance derived from equation 7 with parameters of Table 3 were a priori imposed. The Kalman filter algorithm was then launched backward in time with this prior lithospheric field between 2000.5 and 1900.

Keeping only variance information within the Kalman filter algorithm is a strong approximation. Before implementing it, this approximation was tested during the CHAMP and Swarm eras. For this evaluation phase, the lithospheric field was fully modeled up to $\ell = 30$ and partially modeled (keeping only variance information) between $\ell = 31$ and $\ell = 150$. The remaining part of the model was simulated normally and the dataset used is the one described in "Data" section. The resulting model is referred as the PR model. With this setup, comparisons with the solution obtained at full resolution (FR model) can be performed. In a first simulation, it was observed that the posterior variance associated with the

approximated solution had a tendency to be underestimated. In particular, the transition between the degree variance (the sum of the variances at a given degree) at SH degree $\ell = 30$ and $\ell = 31$ exhibited a pronounced discontinuity. To partially correct this effect, variances beyond the transition were increased by a multiplication factor. The latter was imposed to vary linearly with the degree of the SH expansion, and forced a smooth transition as well as a level of variance at the last modeled degree corresponding to stationary state variance of equation 7. Because of the latter operation, the lithospheric field resolution was increased to $\ell = 200$, a degree at which the signal at satellite altitude becomes very low as shown by Olsen et al. (2017).

The results of this evaluation phase are displayed in Fig. 6. On the left panels, the mean downward component of the lithospheric field at the Earth's surface is shown for both the solution obtained at full resolution (top) and the one obtained at partial resolution (bottom). These two maps look very similar and most features which can be recovered by the FR model are present in the PR model. This aspect is confirmed by the map which exhibits the difference between the two mean solutions (top right). Only at the level of Antarctica, Eastern Europe and Western Russia, discrepancies become quite intense. These discrepancies coincide with relatively large-scale errors (up to $\ell = 70$) as shown with crosses by the energy spectrum at the Earth's surface of the

Fig. 6 Lithospheric field at the Earth's surface expanded up to spherical harmonics degree $\ell = 150$. Left: mean downward component solution for the FR model estimated with full covariance information (top) and for the PR model estimated with variance only information from $\ell = 30$ (bottom). Top right: difference between FR and PR models mean downward components. Bottom right: energy spectra of the means (solid lines), the standard deviations (dashed lines) and the difference (crosses) between the FR model (black lines) and PR model (blue lines)

Fig. 7 Left: energy spectra at the Earth's surface of the mean (continuous lines), the standard deviations (dashed lines), and the prior standard deviation (dash dotted line) for the Kalmag lithospheric model (black lines) and the FR model (blue lines). Dots represent the spectra of the difference between the Kalmag model and the WDMAM model by Lesur et al. (2016) (red), the LCS-1 model by Olsen et al. (2017) (green), and the FR model (blue). Right: degree correlation ρ_ℓ between the Kalmag lithospheric field model and the WDMAM model (red lines), and the LCS-1 model (green lines)

difference between the two mean models (bottom right panel). Beyond $\ell = 70$, the level of error decreases. The computation of the degree correlation between the two models, as introduced by Langel and Hinze (1998) reads:

$$\rho_\ell = \frac{\sum_{\ell=-l}^{\ell=m} \left[g_{l,\ell,m} g_{s,\ell,m} \right]}{\sqrt{\sum_{\ell=-l}^{\ell=m} \left[g_{l,\ell,m}^2 + g_{m,\ell,m}^2 \right]}} \, , \tag{21}$$

also highlights their proximity. The latter reaches a minimum of 0.915 at $\ell = 66$ and stabilizes around the mean value of 0.979 beyond $\ell = 100$. The energy spectra associated with the standard deviations show that the model where only variance information was updated, had a tendency to underestimate the level of predicted uncertainties. Although the technique previously mentioned to rescale the variance was applied, it did not completely resolve this issue. Nevertheless, the fact that the small-scale lithospheric field was only marginally affected by the proposed modeling approximation comforted us to implement it for the complete model derivation.

The lithospheric field resulting from the assimilation of the entire dataset is first analyzed through energy spectra at the Earth's surface. In the left part of Fig. 7, the spectra of the mean, the standard deviation and the prior standard deviation of the lithospheric field are displayed with black lines. In this solution, energy populates the entire range of modeled scales. However, the mean field is predicted to be globally reliable only up to SH degree $\ell \sim 450$, where the spectrum of the mean and the spectrum of the standard deviation cross one another. In

addition, the discontinuity in the spectrum of the mean at SH degree $\ell = 150$ indicates that even up to $\ell \sim 450$ a non-negligible portion of the crustal signal remains unmodeled. Nevertheless, comparisons with the FR model previously discussed (blue lines and dots) demonstrate that the assimilation of survey data helps to better constrain the large-scale lithospheric field. Indeed, the mean signal of the final solution has gained in intensity, and its standard deviation has decreased. In the same figure, the spectra of the difference with two other lithospheric models, the WDMAM model by Lesur et al. (2016) (red dots), and the LCS-1 model by Olsen et al. (2017) (green dots), are also shown.

Although our solution is apparently closer at any degree to the LCS-1 model than to the WDMAM model, the examination of the degree correlation (right panel of Fig. 7) indicates that this aspect is only true up $\ell = 150$. Beyond this value, even if ρ_ℓ is relatively low, the correlation between Kalmag and WDMAM (red line) is higher. Contrary to the degree correlation between LCS-1 and Kalmag which decays smoothly, the one associated with Kalmag and WDMAM presents two transitions. One of them is at SH degree $\ell = 100$, the spatial scale delimiting the satellite data solution ($\ell \leq 100$) from the survey data solution ($\ell > 100$) of the WDMAM model. The other transition occurs at $\ell = 150$, the degree beyond which our model is only constrained by survey data. This second drop in ρ_ℓ may be explained by the lower spatial resolution that our solution exhibits in certain areas. This phenomenon can be observed in Fig. 8 where the downward components of WDMAM (top left) and Kalmag
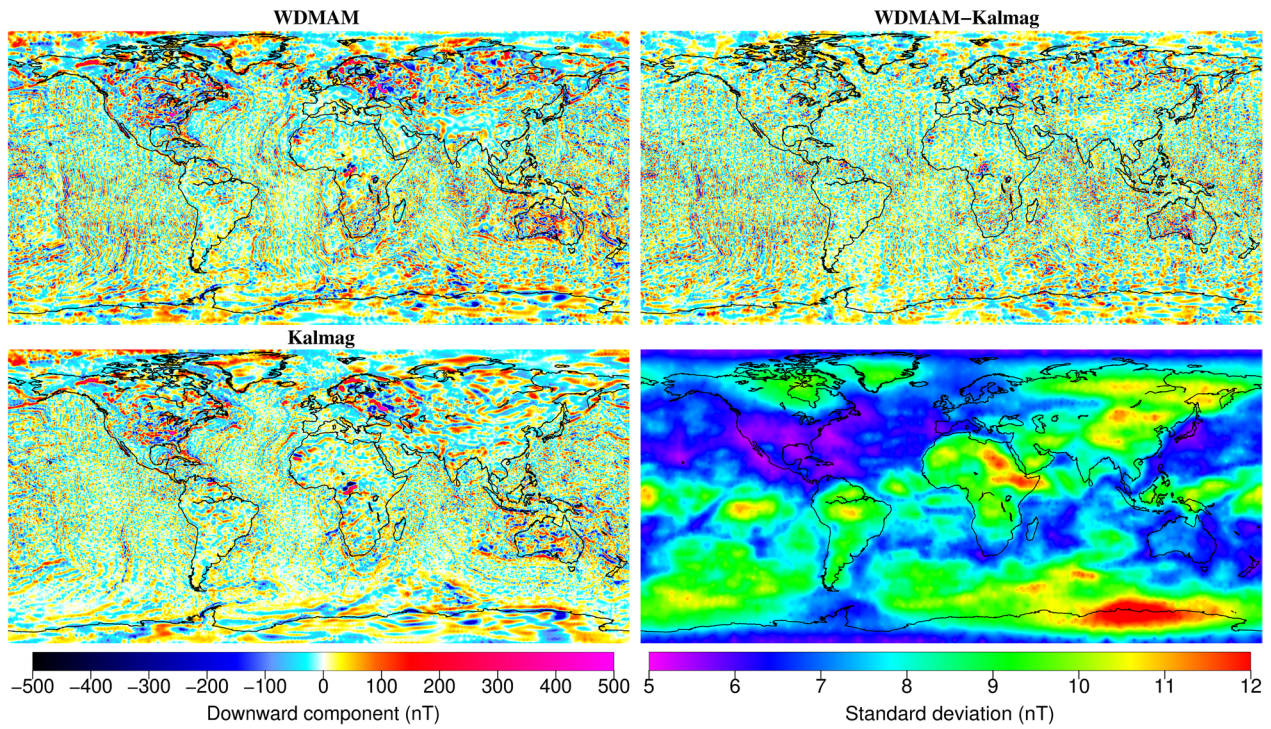
Fig. 8 Downward component of the lithospheric field at the Earth's surface expanded up to spherical harmonics degree $\ell = 450$ for the WDMAM model by Lesur et al. (2016) (top left), the Kalmag model (bottom left), and the difference between the WDMAM and the Kalmag model (top right). The color scale for these three maps is displayed below the mean Kalmag solution. Bottom right: standard deviation associated with the Kalmag mean downward component expanded up to $\ell = 100$
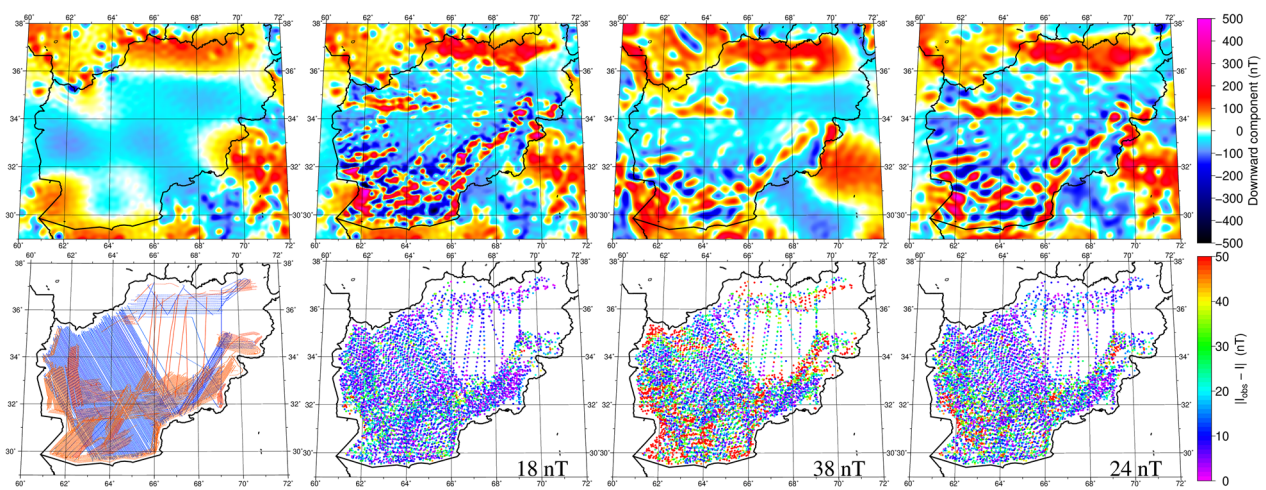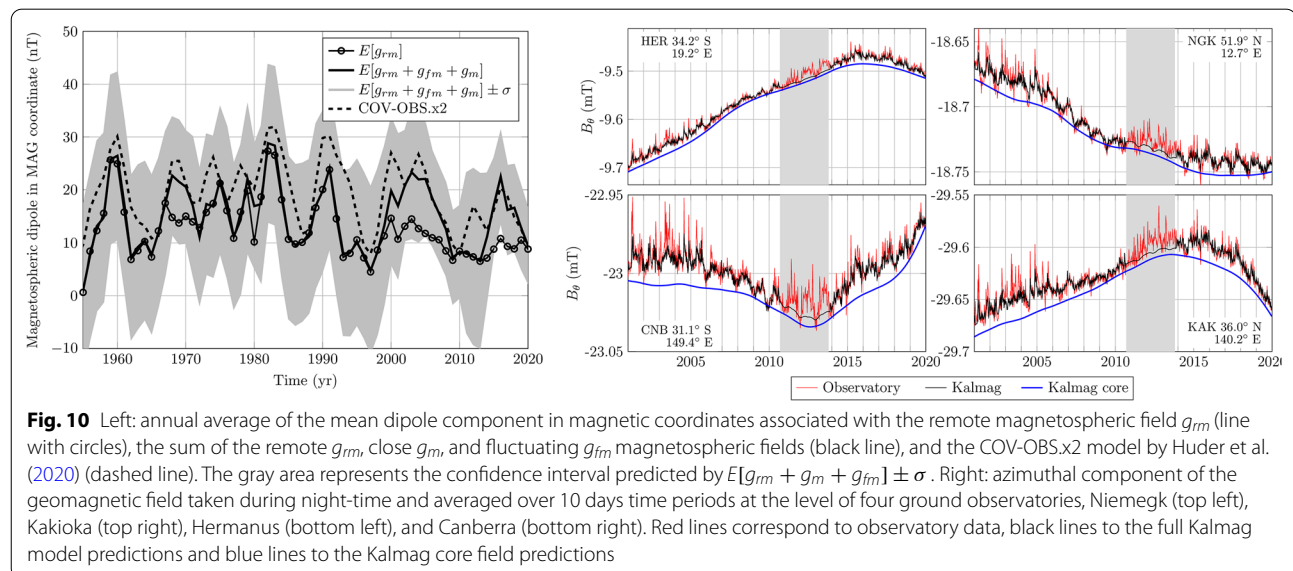

Fig. 9 Lithospheric field update at the level of Afghanistan. Top left: prior mean downward component in 2006.0 expanded up to $\ell = 1000$. Bottom left: locations of airborne intensity measurements taken in 2006 (blue dots) and 2008 (red dots). The second to the fourth map on the top show different models of the downward component of the lithospheric field. These are, respectively, from from left to right, the posterior mean expanded up to $\ell = 2000$, the EMM model by Maus (2010) taken up to $\ell = 790$ and posterior mean truncated at $\ell = 790$. Under each solution, the absolute value of the difference between measured intensities and predicted ones is shown. For predictions, the Kalmag mean core field was included. The values given on the bottom left of these maps correspond to the r.m.s. values of the differences between the model and the data

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 16 of 22

(bottom left) expanded up to $\ell = 450$ (the resolution up to which we predict a globally well-resolved solution) are displayed.

The intense signals predicted by WDMAM in the Southern parts of the Pacific, the Atlantic and the Indian oceans, or on large portions of continental areas are mostly absent in our solution. It is however worth noting that WDMAM does not only derive from direct measurements of the geomagnetic field, but also from the combination of ocean floor age map, relative plate motions and geomagnetic polarity time scale (see Dyment et al. (2015)). Logically, the difference between the downward component of both models (top right of Fig. 8) is larger at these oceanic and land locations than anywhere else. On the opposite, discrepancies are reduced in most areas where the standard deviation associated with the large scale part of the field (up to $\ell = 100$) is low (map on the bottom right). These uncertainty predictions which are tied to data coverage (see Fig. 1) therefore provide a good approximation of locations where the Kalmag model is likely to be well resolved.

The model being expressed in terms of posterior distributions, it can be used as a prior information to assimilate new data when some of them become available, and therefore be updated accordingly. To illustrate this aspect, airborne intensity measurements taken above Afghanistan in 2006 and 2008 were put aside from the dataset serving the model derivation. They are now used to update the lithospheric field following the method detailed in Appendix C. The locations at which each measure was taken during these surveys are shown with colored dots (blue for 2006 red for 2008) in the bottom left panel of Fig. 9. The downward component of

the mean prior lithospheric field, which comes from the smoothing solution taken up to $\ell = 1000$ in 2006.0, is shown on the top left panel. Its resolution was increased to $\ell = 2000$ before the Kalman filter simulation was launched. The result of the assimilation process is shown through the downward component of the mean posterior field in 2009.0 in the second panel of the top row of Fig. 9. On this map, it can be seen that structures which were completely invisible in the prior model appear in the posterior one. In particular, high-intensity anomalies could be detected along the Southern and Western border of Afghanistan. The field in the central part of the land is globally weaker. Such patterns are also predicted by the EMM 2017 model of Maus (2010) as shown on the third panel of the top row. They are nevertheless of lower magnitude, and less detailed due to the resolution of the model which is limited to $\ell = 790$. To make the comparison with the EMM solution possible, the posterior mean was truncated at SH degree $\ell = 790$. The resulting downward component is shown in the top right of the figure. Now the two models are looking more alike. Nevertheless, discrepancies in predicted intensity still remain. In order to assess the degree of compatibility of the different models with the observations, the absolute value of the difference between a subset of the measurements and the intensities predicted by the sum of the core and the different lithospheric field solutions was computed. The results are shown on the bottom panel below each corresponding downward components. The model exhibiting the higher degree of freedom, displayed on the second column, is without surprise the model which can better explain the data. As shown on the bottom of the map, the r.m.s. difference between the model and the

Fig. 10 Left: annual average of the mean dipole component in magnetic coordinates associated with the remote magnetospheric field g_{rm} (line with circles), the sum of the remote g_{rm}, close g_m, and fluctuating g_{fm} magnetospheric fields (black line), and the COV-OBS.x2 model by Huder et al. (2020) (dashed line). The gray area represents the confidence interval predicted by $E[g_{rm} + g_m + g_{fm}] \pm \sigma$. Right: azimuthal component of the geomagnetic field taken during night-time and averaged over 10 days time periods at the level of four ground observatories, Niemegk (top left), Kakioka (top right), Hermanus (bottom left), and Canberra (bottom right). Red lines correspond to observatory data, black lines to the full Kalmag model predictions and blue lines to the Kalmag core field predictions

Julien *et al. Earth, Planets and Space* (2022) 74:139

Page 17 of 22

measurements is of 18nT. Globally the predictions of the truncated model (right column) are closer to the data than the EMM predictions (third column). Of course Afghanistan is a particular location and no claim is made here that the Kalmag model would be globally more accurate than the EMM model since this is certainly not the case. However, this example shows that the method proposed in this study is well suited to construct regional high-resolution models of the lithospheric field and this even when data coverage is not optimal.

Magnetospheric and induced fields

With the proposed approach, magnetospheric and induced fields are jointly estimated with the rest of the model. A priori, the field generated by the currents flowing in the outer magnetosphere (g_{rm}) is predicted to evolve slowly with time ($\tau_{g_{rm}} = 10.3$ years) in comparison to other external sources. A posteriori, such a behavior is confirmed as illustrated by the evolution of the annual mean dipole component of $E[g_{rm}]$ projected in magnetic coordinates and shown in the left panel of Fig. 10 with circles. Note that prior to 1953, our model cannot correctly extract this field and the latter oscillates around 0 with a large posterior variance. However, g_{rm} alone cannot explain decadal variations of external sources as they can be detected at the Earth's surface or at the altitude of low-orbiting satellites. The rapidly evolving magnetospheric components also exhibit long-term trends whenever the latter can be captured. This effect can be observed when comparing the annual mean dipole component of $E[g_{rm}]$ to the one of $E[g_{rm} + g_m + g_{fm}]$ shown with a continuous line in Fig. 10. During satellite eras, the latter is always found to be more intense than the former, meaning that the ring current can generate some persistent annual signal as already documented by Lühr and Maus (2010). With our current method, this signal can only be recovered when temporal data coverage is high enough due to the fact that $E[g_m]$ and $E[g_{fm}]$ exhibit very low memory timescales. A possible way to improve the AR processes characterizing these sources would be to consider some extra timescales accounting for the slow varying part of the field generated by the ring current. The cycle of approximately 10.5 years highlighted by Huder et al. (2020) with the COV-OBS.x2 model (shown with dashed lines in Fig. 10) is also present in our solution. Although the mean solutions of both models slightly differ from one another, the COV-OBS.x2 dipole always lies within the 68.7 confidence interval predicted by our model (gray areas in Fig. 10).

To evaluate the model over short periods of time and when all sources are predicted to be well separated, we now compare predictions of the azimuthal component of the model 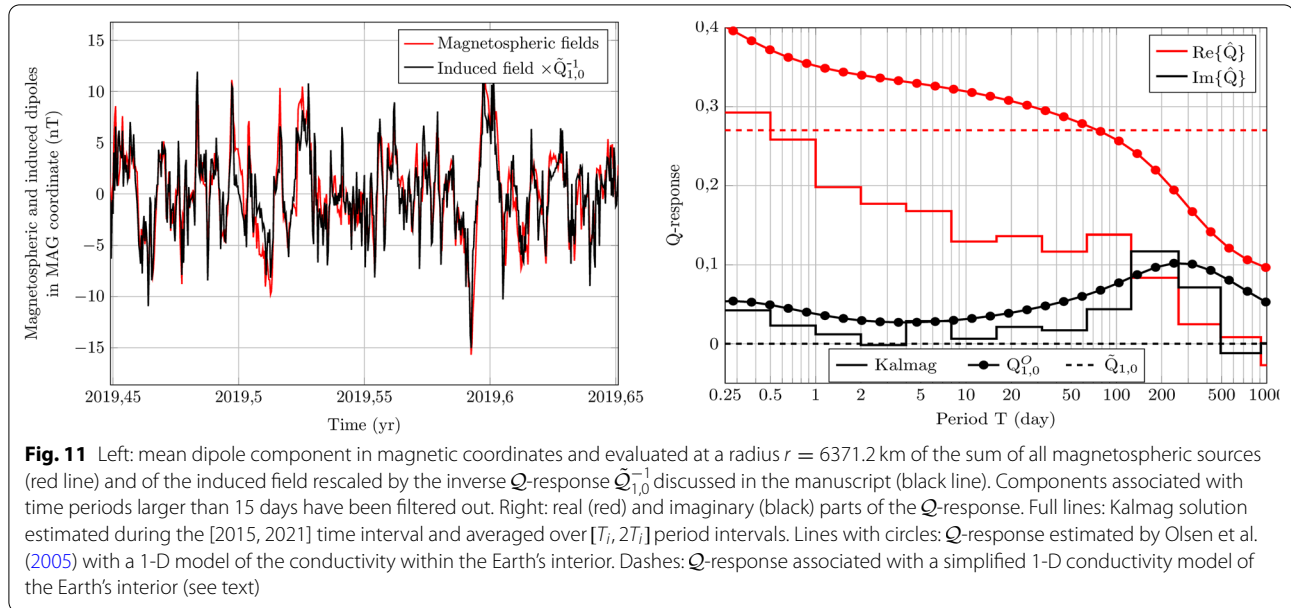with ground based observatory measurements taken at four different locations, Hermanus, Niemegk, Canberra and Kakioka. Observatory data being only assimilated to constrain the core field secular variation, they can be considered as independent measurements for external and induced fields. In order to make visual comparisons possible and to remain within the conditions the model was built in, only hourly night-time measurements and predictions were kept to be then averaged over 10 days time periods. The results are reported in the right panel of Fig. 10 with red lines for observatory data, black lines for the full model predictions and blue lines for the predictions of the core field alone. Globally, monthly and annual variations of B_θ are well captured by the model. Only during the time gap between the CHAMP and Swarm missions, when external sources are not updated anymore, predictions and observations differ strongly. One can also notice that the core field does not seem to be contaminated by external or induced fields, as its evolution does not reproduce the rapid variations observed in the data. The largest discrepancies between predictions and observations are in the magnitude of the signals. Intense excursions are not predicted by the model. The reason for this is that the model was trained on a dataset selected for very quiet magnetic conditions [see Baerenzung et al. (2020)]. Therefore, the selection algorithm of the Kalman filter prevents the assimilation of data containing a too strong signal from external sources. A recalibration of the model for more general conditions would certainly solve this issue.

Finally, our model contains a source for induced/residual ionospheric fields. The latter is a priori uncorrelated from magnetospheric fields. Yet rapid variations of external fields generate currents within the Earth's interior, which in return induce a secondary magnetic field (e.g., Schmucker 1985; Langel and Estes 1985b; Olsen et al. 2005; Finlay et al. 2020). The intensity and temporal evolution of the induced field depends on the conductivity of the crust, the mantle and the core. Under the assumption that conductivity only depends on depth, each spherical harmonics coefficient of the induced field will be linked the same coefficient of the external field through the relation:

$$\iota_{l,m}(t) = \int_{-\infty}^{\infty} \mathcal{Q}_{l,m}(t - t')\epsilon_{l,m}(t')\mathrm{d}t' , \qquad (22)$$

where ι is the induced field, ϵ the external fields, and $\mathcal{Q}$ is referred as the $\mathcal{Q}$-response. In our model, $\iota = g_{ii}$ and $\epsilon = g_{rm} + g_m + g_{fm}$, where g_{rm} is projected in magnetic coordinates.

In the particular case discussed by Olsen et al. (2005), where the mantle is assumed to be insulating until a given depth d followed by a superconductor, $\mathcal{Q}_l^m(t - t') = \tilde{\mathcal{Q}}_l^m\delta(t - t')$ and therefore

Fig. 11 Left: mean dipole component in magnetic coordinates and evaluated at a radius $r = 6371.2$ km of the sum of all magnetospheric sources (red line) and of the induced field rescaled by the inverse $\mathcal{Q}$-response $\tilde{\mathcal{Q}}_{1,0}^{-1}$ discussed in the manuscript (black line). Components associated with time periods larger than 15 days have been filtered out. Right: real (red) and imaginary (black) parts of the $\mathcal{Q}$-response. Full lines: Kalmag solution estimated during the [2015, 2021] time interval and averaged over $[T_i, 2T_i]$ period intervals. Lines with circles: $\mathcal{Q}$-response estimated by Olsen et al. (2005) with a 1-D model of the conductivity within the Earth's interior. Dashes: $\mathcal{Q}$-response associated with a simplified 1-D conductivity model of the Earth's interior (see text)

$\iota_{l,m}(t) = \tilde{\mathcal{Q}}_{l,m}\epsilon_{l,m}(t)$. Focusing on the dipole component of induced and external fields, and assuming a depth of $d = 1200$ km, leads to $\iota_{1,0}(t) = \tilde{\mathcal{Q}}_{1,0}\epsilon_{1,0}(t)$ with $\tilde{\mathcal{Q}}_{1,0} = 0.27$ as estimated by Langel and Estes (1985b) with POGO data. In the left panel of Fig. 11, the evolution of, respectively, $\epsilon_{1,0}$ and $\iota_{1,0}(t)/\tilde{\mathcal{Q}}_{1,0}$ is displayed between 2019.45 and 2019.65 with, respectively, red and black lines. In order to concentrate on rapid variations only (we recall that external and induced field were stored every 3 hours during the CHAMP and Swarm eras), temporal scales larger than 15 days have been filtered out from both time series. Furthermore, we chose the [2019.45, 2019.65] time period because temporal coverage of Swarm data is optimal during this interval. The two time series in Fig. 11 follow one another quite closely and $\tilde{\mathcal{Q}}_{1,0}^{-1}$ seems appropriate to rescale the induced field. Over the current Swarm time span, induced and external fields exhibit a Pearson correlation $\rho = \text{Cov}(\epsilon, \iota)/(\sigma_\epsilon \sigma_\iota)$, calculated here with the mean Kalmag solutions, of $\rho = 0.79$. It is of $\rho = 0.84$ over the time interval of Fig. 11 and of $\rho = 0.73$ over the CHAMP era. This lower correlation value is probably caused by the uncertainty level of external and induced fields which are higher during the CHAMP mission than during the Swarm one. However, the particular 1-D conductivity model leading to $\tilde{\mathcal{Q}}_{1,0}$ is known to be imperfect. More complex conductivity profiles are required to better model induction processes within the Earth's interior.

We now investigate the $\mathcal{Q}$-response predicted by our model when keeping the assumption that the conductivity within the Earth is only depth-dependent, but relaxing the constraint about its profile. For this evaluation, we operate in spectral space. Considering only dipole components of ι and ϵ and applying a Fourier transform to equation 22 the latter becomes:

$$\hat{\iota}_{1,0}(k) = \hat{\mathcal{Q}}_{1,0}(k)\hat{\epsilon}_{1,0}(k) . \tag{23}$$

From this equation, the real and imaginary parts of $\hat{\mathcal{Q}}(k)$ are, respectively, given by:

$$\text{Re}\{\hat{\mathcal{Q}}(k)\} = \frac{\text{Re}\{\hat{\epsilon}(k)\}\text{Re}\{\hat{\iota}(k)\} + \text{Im}\{\hat{\epsilon}(k)\}\text{Im}\{\hat{\iota}(k)\}}{\text{Re}\{\hat{\epsilon}(k)\}^2 + \text{Im}\{\hat{\epsilon}(k)\}^2}, \tag{24}$$

$$\text{Im}\{\hat{\mathcal{Q}}(k)\} = \frac{\text{Re}\{\hat{\epsilon}(k)\}\text{Im}\{\hat{\iota}(k)\} - \text{Im}\{\hat{\epsilon}(k)\}\text{Re}\{\hat{\iota}(k)\}}{\text{Re}\{\hat{\epsilon}(k)\}^2 + \text{Im}\{\hat{\epsilon}(k)\}^2} . \tag{25}$$

To evaluate these two quantities we considered induced and external fields during the [2015.0, 2021.0] time interval when the model reaches its peak accuracy. In the right panel of Fig. 11, $\text{Re}\{\hat{\mathcal{Q}}(2\pi/k)\}$ and $\text{Im}\{\hat{\mathcal{Q}}(2\pi/k)\}$ averaged at period $T_i = 2\pi/k_i$ over $[T_i, 2T_i]$ are, respectively, displayed with red and black continuous lines. For comparisons, the real and imaginary parts of $\tilde{\hat{\mathcal{Q}}}_{1,0}$ as well as the $\mathcal{Q}$-response (referred as $\mathcal{Q}^O$) estimated by Olsen et al. (2005) with a realistic conductivity model are, respectively, shown with dashed and dotted lines. The general behavior of the $\mathcal{Q}$-response we recover is coherent with our prior knowledge about it. Indeed, for short periods of time the real part of $\hat{\mathcal{Q}}$ is much more intense than its

imaginary part and its decay pattern is close to the one predicted by Olsen et al. (2005). However, in comparison to Re$\{\hat{\mathcal{Q}}^O\}$, Re$\{\hat{\mathcal{Q}}\}$ is globally underestimated. This effect might be due to the fact that induced fields vary rapidly with time, and when no data is feeding the model, its mean value tends quickly toward 0 contrary to the remote and close magnetospheric fields which evolves slower. The behavior of the imaginary part of $\hat{\mathcal{Q}}$, which reflects the temporal lag of the induced field response, is on the contrary very similar to the one predicted by the direct model of Olsen et al. (2005).

Conclusion

In this study, we proposed a method to assimilate different types of geomagnetic data in order to construct a high spatio-temporal model of the Earth's magnetic field. The model being expressed in terms of posterior distribution, it reflects the quality and spatial coverage of the measurements it is derived from. At the beginning of the twentieth century, the main field and the secular variation are quite uncertain in the Southern hemisphere and more particularly in oceanic areas and in Antarctica. With the first data collected by low-orbiting satellites, these two fields gain in precision and become very reliable during the CHAMP and Swarm eras. We demonstrated that the rapid dynamics of the core field could be captured by the model. However, the spatial resolution at which short timescale fluctuations are recovered is not constant over time and strongly depends on the spatial scale considered. Typically, rapid variations can only be accurately modeled at large spatial scale. On the opposite, fluctuations of the secular variation at high spherical harmonics degree can only be resolved for long periods of time. The model reaches its peak accuracy both spatially and temporally during the CHAMP and Swarm eras. It is therefore mandatory that such satellite missions are perpetrated in the future to better understand the nonlinear and wave dynamics occurring within the Earth's outer core (see Aubert and Gillet (2021), Gillet et al. (2021)).

To be able to consider land, airborne and marine survey observations, which contain an intense contribution of the small-scale lithospheric field, the latter was modeled up to spherical harmonic degree $\ell = 1000$. However, this operation could not be performed directly, since the dimension of the associated covariance matrix would have forbidden any numerical computation. We therefore introduced, and conclusively evaluated, a statistical approximation where only mean and variance information were updated beyond $\ell = 100$. The resulting mean solution exhibits highly detailed structures on every areas where data coverage

was dense enough. Furthermore, the part of the covariance which is still fully modeled (up to $\ell = 100$) provides a rough estimation of locations where the mean is likely to be well resolved.

An important aspect of the proposed approach is that whenever new observations become available, the model can be updated accordingly without restarting the entire assimilation process. The example presented with the dataset taken above Afghanistan demonstrates the flexibility of the method.

As for the core field, the accuracy of external and induced fields is not constant over the model time span. While signal of the remote magnetospheric field could be extracted from 1953 on, rapidly evolving sources such as the close and fluctuating magnetospheric fields or the induced field, could only be separated from the data when the latter exhibit a high temporal coverage. In general, optimal solution for such field was obtained during satellite eras and in particular during the CHAMP and Swarm ones. The global behavior of external fields is in agreement with previous studies of it (see Lühr and Maus (2010); Huder et al. (2020)). However, the training of the model under very quiet magnetic conditions forbids the reproduction of most intense external field variations. A recalibration of the model under more general conditions appears therefore as necessary. Although magnetospheric and induced fields were a priori assumed to be independent, their connection revealed itself a posteriori. Through the proposed approach we showed that external and induced fields could be jointly estimated from direct measurements of the geomagnetic field although the process characterizing their evolution remain quite simplistic. A refined parametrization of their dynamical behavior would certainly enhance the ability of the algorithm to extract such sources from the data.

The model will be frequently updated (at least once every 2 months), in particular with Swarm and observatory data. Furthermore, it can be accessed through different physical and statistical properties on a dedicated website at: https://ionocovar.agnld.uni-potsdam.de/Kalmag/.

Appendix A: Sequentialization

Describing the evolution of a given *g* quantity by continuous first and second-order auto-regressive processes can be preformed through the following relations:

$$\partial_t g + \frac{1}{\tau} g = \sigma \dot{\omega}, \tag{A.1}$$

$$\partial_t^2 g + \frac{2}{\tau}\partial_t g + \frac{1}{\tau^2}g = \sigma\dot\omega, \tag{A.2}$$

where $\dot\omega$ is a Gaussian white noise scaled by the factor σ. Introducing $z = g$ for first-order ARP and $z = (g, \partial_t g)^T$ for second-order ARP, equations (A.1) and (A.2) can be written as:

$$\partial_t z = -Az + \zeta, \tag{A.3}$$

with $A = 1/\tau$ and $\zeta = \sigma\dot\omega$ for first-order ARP and:

$$A = \begin{pmatrix} 0 & 1 \\ 1/\tau^2 & 2/\tau \end{pmatrix}, \quad \zeta = \begin{pmatrix} 0 \\ \sigma\dot\omega \end{pmatrix} \tag{A.4}$$

for second-order ARP.

The homogeneous solution of equation A.3 is given by:

$$z(t + \Delta t) = z(t)\exp(-A\Delta t) = Fz(t), \tag{A.5}$$

where $F = \exp(-A\Delta t)$ is the parameter of the ARP as expressed in equations 4 and 5. The general solution of equation A.3 is simply $z(t + \Delta t) = Fz(t) + \xi$, where the white noise ξ characterized by the distribution $\mathcal{N}(0, \tilde\Sigma)$ is chosen here to force the process to remain stationary. Under such a constraint, one can write that $\Sigma(t) = E\left[(z(t) - E[z(t)])(z(t) - E[z(t)])^T\right] = \Sigma(t = \infty) = \Sigma_\infty$. Therefore, calculating the spatial covariance of both sides of the solution $z(t + \Delta t) = Fz(t) + \xi$ and rearranging the result gives $\tilde\Sigma = \Sigma_\infty - F\Sigma_\infty F^T$.

Appendix B: Smoothing and merging

The model being constructed in multiple steps, the smoothing algorithm had to be adapted. In particular, information gained with the assimilation of data prior to 2000.5 had to be propagated to the model constructed after this date with CHAMP, Swarm and ground-based observatory data. To do so, the solution in 2000.4 of the smoothing algorithm running between 1900 and 2000.4 was taken as a reference model. Its is referred as $\mathbf{z}_r$, with mean $E[\mathbf{z}_r]$ and covariance $\Sigma_{\mathbf{z}_r}$. Information accumulated within this snapshot is then transferred to the first smoothing solution (the one running from 2022.2 to 2000.5). The algorithm to perform this task proceeds iteratively in time through the relations:

$$\mathbf{G}_k = \Sigma_{\mathbf{z}_k}\mathbf{F}_k^T\Sigma_{\mathbf{z}_r}^{-1}, \tag{B.1}$$

$$E[\mathbf{z}_k|\mathbf{z}_r] = E[\mathbf{z}_k] + \mathbf{G}_k(E[\mathbf{z}_k] - \mathbf{F}_k E[\mathbf{z}_r]), \tag{B.2}$$

$$\Sigma_{\mathbf{z}_k|\mathbf{z}_r} = \Sigma_{\mathbf{z}_k} + \mathbf{G}_k\left(\Sigma_{\mathbf{z}_k} - \mathbf{F}_k\Sigma_{\mathbf{z}_r}\mathbf{F}_k^T - \tilde\Sigma_{\mathbf{z}_r}\right)\mathbf{G}_k^T, \tag{B.3}$$

where $\mathbf{z}_k$ is a snapshot taken at iteration k of the $2000.5 - 2022.2$ smoothing model, $\mathbf{F}_k$ is the parameter of the ARP enabling to forecast $\mathbf{z}_r$ to the k iteration, and $\tilde\Sigma_{\mathbf{z}_r}$ is the covariance of the ARP white noise derived from $\mathbf{F}_k$.

Appendix C: Lithospheric field update

Whenever new survey data become available, one may wish to assimilate them to improve the lithospheric field model. A possibility would be to merge this new dataset with the global one and to relaunch the entire Kalman filter/smoothing algorithm. Yet this operation is extremely time consuming. A better option would be to use the posterior model resulting from the smoothing algorithm as a prior information to assimilate the new dataset and then to propagate the information gained on the lithospheric field to the entire model. To perform this task lets assume that new data become available within the time interval $[k - 1, k + 1]$, where k corresponds to a stored snapshot of the smoothing solution. Assimilating data between $k - 1$ and k is straightforward. One can simply simulate the Kalman filter algorithm with the smoothing solution $\mathbf{z}_{k-1}$ taken as a restart file. Arriving at k, if the new dataset offers only a limited coverage of the Earth's surface, the accuracy the core field and other sources exhibit will likely be lower than the accuracy of the smoothing solution at this epoch. It would therefore be beneficial to use this solution $\mathbf{z}_k$ as a prior model. At the same time nevertheless, information gained on the lithospheric field between $k - 1$ and k needs to be transferred to it. To perform this operation we proceeded as following:

$$\mathbf{G}_k = \Sigma_{\mathbf{z}_k g_k^s}\Sigma_{g_k^u}^{-1}, \tag{C.1}$$

$$E[\mathbf{z}_k|g_k^u] = E[\mathbf{z}_k] + \mathbf{G}_k\left(E[g_k^u] - E[g_k^s]\right), \tag{C.2}$$

$$\Sigma_{\mathbf{z}_k|g_k^u} = \Sigma_{\mathbf{z}_k} + \mathbf{G}_k\left(\Sigma_{g_k^u} - \Sigma_{g_k^s}\right)\mathbf{G}_k^T, \tag{C.3}$$

where g_k^u and g_k^s are the vectors of SH harmonics components associated with the lithospheric field for, respectively, the Kalman filter and the smoothing solution and $\Sigma_{\mathbf{z}_k g_k^s}$ is the smoothing cross-covariance matrix between $\mathbf{z}_k$ and g_k^s. These updated smoothing solutions at epoch k are then used as a prior information for the Kalman filter running between k and $k + 1$. Such operation is then repeated every 0.1 year, as constrained by the chosen temporal resolution of the model.

Once the entire new dataset has been assimilated, the Kalman filter solution in 1900 is updated with the new lithospheric field through equations C.1 to C.3 and the smoothing algorithm between 1900 and 2022 is simulated again.

Julien *et al. Earth, Planets and Space*   (2022) 74:139

Page 21 of 22

Author contributions
All the authors read and approved the final manuscript.

Declarations

Author details
[1]Section 1.3: Earth System Modelling, Geoforschungszentrum Potsdam, Potsdam, Germany. [2]Institute for Mathematics, University of Potsdam, Potsdam, Germany.

References
Alken P, Thébault E, Beggan CD, Amit H, Aubert J, Baerenzung J, Bondar TN, Brown WJ, Califf S, Chambodut A, Chulliat A, Cox GA, Finlay CC, Fournier A, Gillet N, Grayver A, Hammer MD, Holschneider M, Huder L, Hulot G, Jager T, Kloss C, Korte M, Kuang W, Kuvshinov A, Langlais B, Léger JM, Lesur V, Livermore PW, Lowes FJ, Macmillan S, Magnes W, Mandea M, Marsal S, Matzka J, Metman MC, Minami T, Morschhauser A, Mound JE, Nair M, Nakano S, Olsen N, Pavón-Carrasco FJ, Petrov VG, Ropp G, Rother M, Sabaka TJ, Sanchez S, Saturnino D, Schnepf NR, Shen X, Stolle C, Tangborn A, Tøffner-Clausen L, Toh H, Torta JM, Varner J, Vervelidou F, Vigneron P, Wardinski I, Wicht J, Woods A, Yang Y, Zeren Z, Zhou B (2021) International geomagnetic reference field: the thirteenth generation. Earth Planets Space 73(1):49. https://doi.org/10.1186/s40623-020-01288-x
Amante C, Eakins B (2009) ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis. NOAA technical memorandum NESDIS NGDC-24. Natl Geophys Data Center. https://doi.org/10.7289/V5C8276M
Aubert J, Gillet N (2021) The interplay of fast waves and slow convection in geodynamo simulations nearing Earth's core conditions. Geophys J Int 225(3):1854–1873. https://doi.org/10.1093/gji/ggab054

Backus GE (1970) Non-uniqueness of the external geomagnetic field determined by surface intensity measurements. J Geophys Res 75(31):6339. https://doi.org/10.1029/JA075i031p06339
Baerenzung J, Holschneider M, Wicht J, Lesur V, Sanchez S (2020) The Kalmag model as a candidate for IGRF-13. Earth Planets Space 72(1):163. https://doi.org/10.1186/s40623-020-01295-y
Cain JC, Sweeney RE (1973) The POGO data. J Atmos Terr Phys 35:1231. https://doi.org/10.1016/0021-9169(73)90021-4
Dyment J, Choi Y, Hamoudi M, Lesur V, Thebault E (2015) Global equivalent magnetization of the oceanic lithosphere. Earth Planet Sci Lett 430:54–65. https://doi.org/10.1016/j.epsl.2015.08.002
Finlay CC, Kloss C, Olsen N, Hammer MD, Tøffner-Clausen L, Grayver A, Kuvshinov A (2020) The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly. Earth Planets Space 72(1):156. https://doi.org/10.1186/s40623-020-01252-9
Gillet N (2019) Spatial and temporal changes of the geomagnetic field: insights from forward and inverse core field models. Geomagn Aeron Space weather. https://doi.org/10.48550/arXiv.1902.08098
Gillet N, Jault D, Finlay CC, Olsen N (2013) Stochastic modeling of the Earth's magnetic field: inversion for covariances over the observatory era. Geochem Geophys Geosyst 14:766–786. https://doi.org/10.1002/ggge.20041
Gillet N, Jault D, Finlay CC (2015) Planetary gyre, time-dependent eddies, torsional waves, and equatorial jets at the Earth's core surface. J Geophys Res 120:3991–4013. https://doi.org/10.1002/2014JB011786
Gillet N, Gerick F, Angappan R, Jault D (2021) A dynamical prospective on interannual geomagnetic field changes. Surv Geophys. https://doi.org/10.1007/s10712-021-09664-2
Holschneider M, Lesur V, Mauerberger S, Baerenzung J (2016) Correlation-based modeling and separation of geomagnetic field components. J Geophys Res 121:3142–3160. https://doi.org/10.1002/2015JB012629
Huder L, Gillet N, Finlay CC, Hammer MD, Tchoungui H (2020) COV-OBS.x2: 180 years of geomagnetic field evolution from ground-based and satellite observations. Earth Planets Space 72(1):160. https://doi.org/10.1186/s40623-020-01194-2
Hulot G, Le Mouël JL (1994) A statistical approach to the Earth's main magnetic field. Phys Earth Planet Inter 82(3–4):167–183. https://doi.org/10.1016/0031-9201(94)90070-1
Hulot G, Sabaka T, Olsen N, Fournier A (2015) 5.02 - the present and future geomagnetic field. In: Schubert G (ed) Treatise on geophysics, 2nd edn. Elsevier, Oxford, pp 33–78. https://doi.org/10.1016/B978-0-444-53802-4.00096-8
Jackson A, Finlay CC (2007) Geomagnetic secular variation and its applications to the core. Treatise Geophys 5:147–193. https://doi.org/10.1016/B978-044452748-6.00090-0
Jackson A, Jonkers ART, Walker MR (2000) Four centuries of geomagnetic secular variation from historical records, in Astronomy, physics and chemistry of H$^+_3$. Philos Trans R Soc Lond Ser A 358:957. https://doi.org/10.1098/rsta.2000.0569
Kalman RE (1960) A new approach to linear filtering and prediction problems. J Basic Eng 82:35–45. https://doi.org/10.1115/1.3662552
Langel RA, Estes RH (1985) The near-earth magnetic field at 1980 determined from magsat data. J Geophys Res 90(B3):2495–2510. https://doi.org/10.1029/JB090iB03p02495
Langel RA, Estes RH (1985) Large-scale, near-field magnetic fields from external sources and the corresponding induced internal field. J Geophys Res 90(B3):2487–2494. https://doi.org/10.1029/JB090iB03p02487
Langel RA, Hinze WJ (1998) The magnetic field of the earth's lithosphere. Cambridge University Press, Cambridge
Laundal R (2017) Magnetic coordinate systems. Space Sci Rev 206:27–59. https://doi.org/10.1007/s11214-016-0275-y
Lesur V, Wardinski I, Rother M, Mandea M (2008) GRIMM: the GFZ reference internal magnetic model based on vector satellite and observatory data. Geophys J Int 173:382–394. https://doi.org/10.1111/j.1365-246X.2008.03724.x
Lesur V, Wardinski I, Hamoudi M, Rother M (2010) The second generation of the GFZ reference internal magnetic model: GRIMM-2. Earth Planets Space 62:765–773. https://doi.org/10.5047/eps.2010.07.007
Lesur V, Whaler K, Wardinski I (2015) Are geomagnetic data consistent with stably stratified flow at the core-mantle boundary? Geophys J Int 201:929–946. https://doi.org/10.1093/gji/ggv031

Lesur V, Hamoudi M, Choi Y, Dyment J, Thébault E (2016) Building the second version of the World Digital Magnetic Anomaly Map (WDMAM), Earth. Planets Space 68:27. https://doi.org/10.1186/s40623-016-0404-6

Lowes FJ (1975) Vector errors in spherical harmonic analysis of scalar data. Geophys J Int 42(2):637–651. https://doi.org/10.1111/j.1365-246X.1975.tb05884.x

Lühr H, Maus S (2010) Solar cycle dependence of quiet-time magnetospheric currents and a model of their near-Earth magnetic fields. Earth Planets Space 62(10):843–848. https://doi.org/10.5047/eps.2010.07.012

Macmillan S, Olsen N (2013) Observatory data and the Swarm mission. Earth Planets Space 65(11):1355–1362. https://doi.org/10.5047/eps.2013.07.011

Mauerberger S, Schanner M, Korte M, Holschneider M (2020) Correlation based snapshot models of the archeomagnetic field. Geophys J Int 223(1):648–665. https://doi.org/10.1093/gji/ggaa336

Maus S (2010) An ellipsoidal harmonic representation of Earth's lithospheric magnetic field to degree and order 720. Geochem Geophy Geosyst 11(6):Q06015. https://doi.org/10.1029/2010GC003026

Maus S, Lühr H, Balasis G, Rother M, Mandea M (2005) Introducing POMME, the POtsdam magnetic model of the earth. Earth Obs Champ. https://doi.org/10.1007/3-540-26800-6_46

Maus S, Manoj C, Rauberg J, Michaelis I, Lühr H (2010) NOAA/NGDC candidate models for the 11th generation International geomagnetic reference field and the concurrent release of the 6th generation Pomme magnetic model. Earth Planets Space 62:729–735. https://doi.org/10.5047/eps.2010.07.006

Neubert T, Mandea M, Hulot G, von Frese R, Primdahl F, Jørgensen JL, Friis-Christensen E, Stauning P, Olsen N, Risbo T (2001) Ørsted satellite captures high-precision geomagnetic field data. EOS Trans 82(7):81–88. https://doi.org/10.1029/01EO00043

Olsen N, Sabaka TJ, Lowes F (2005) New parameterization of external and induced fields in geomagnetic field modeling, and a candidate model for IGRF 2005. Earth Planets Space 57:1141–1149. https://doi.org/10.1186/BF03351897

Olsen N, Lühr H, Sabaka TJ, Mandea M, Rother M, Toeffner-Clausen L, Choi S (2006) CHAOS-a model of the Earth's magnetic field derived from CHAMP, Oersted, and SAC-C magnetic satellite data. Geophy J Int 166:67–75. https://doi.org/10.1111/j.1365-246X.2006.02959.x

...Olsen N, Friis-Christensen E, Floberghagen R, Alken P, Beggan CD, Chulliat A, Doornbos E, da Encarnação JT, Hamilton B, Hulot G, van den IJssel J, Kuvshinov A, Lesur V, Lühr H, Macmillan S, Maus S, Noja M, Olsen PEH, Park J, Plank G, Püthe C, Rauberg J, Ritter P, Rother M, Sabaka TJ, Schachtschneider R, Sirol O, Stolle C, Thébault E, Thomson AWP, Tøffner-Clausen L, Velímský J, Vigneron P, Visser PN (2013) The Swarm Satellite Constellation Application and Research Facility (SCARF) and Swarm data products. Earth Planets Space 65(11):1189–1200. https://doi.org/10.5047/eps.2013.07.001

Olsen N, Ravat D, Finlay CC, Kother LK (2017) LCS-1: a high-resolution global model of the lithospheric magnetic field derived from CHAMP and Swarm satellite observations. Geophys J Int 211(3):1461–1477. https://doi.org/10.1093/gji/ggx381

Quesnel Y, CataláN M, Ishihara T (2009) A new global marine magnetic anomaly data set. J Geophys Res 114(B4):B04106. https://doi.org/10.1029/2008JB006144

Rauch HE, Striebel CT, Tung F (1965) Maximum likelihood estimates of linear dynamic systems. AIAA J 3(8):1445–1450. https://doi.org/10.2514/3.3166

Ropp G, Lesur V, Baerenzung J, Holschneider M (2020) Sequential modelling of the Earth's core magnetic field. Earth Planets Space 72(1):153. https://doi.org/10.1186/s40623-020-01230-1

Rother M, Michaelis K, Olsen N (2000) Resolution studies of fluid flow models near the core-mantle boundary using bayesian inversion. In: Hansen P, Jacobsen B, Mosegaard K (eds) Methods and applications of inversion, lecture notes in earth sciences, vol 92. Springer, Berlin, pp 255–275. https://doi.org/10.1007/BFb0010296

Sabaka TJ, Olsen N, Langel RA (2002) A comprehensive model of the quiet-time, near-Earth magnetic field: phase 3. Geophy J Int 151:32–68. https://doi.org/10.1046/j.1365-246X.2002.01774.x

Sabaka TJ, Olsen N, Purucker ME (2004) Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data. Geophys J Int 159(2):521–547. https://doi.org/10.1111/j.1365-246X.2004.02421.x

Sabaka TJ, Olsen N, Tyler RH, Kuvshinov A (2015) CM5, a pre-Swarm comprehensive geomagnetic field model derived from over 12 yr of CHAMP, Oersted, SAC-C and observatory data. Geophys J Int 200:1596–1626. https://doi.org/10.1093/gji/ggu493

Sabaka TJ, Tøffner-Clausen L, Olsen N, Finlay CC (2018) A comprehensive model of Earth's magnetic field determined from 4 years of Swarm satellite observations. Earth Planets Space 70:130. https://doi.org/10.1186/s40623-018-0896-3

Sabaka TJ, Tøffner-Clausen L, Olsen N, Finlay CC (2020) CM6: a comprehensive geomagnetic field model derived from both CHAMP and Swarm satellite observations. Earth Planets Space 72(1):80. https://doi.org/10.1186/s40623-020-01210-5

Schanner M, Korte M, Holschneider M (2022) ArchKalmag14k: A Kalman-filter based global geomagnetic model for the Holocene. J Geophys Res 127(2):e23166. https://doi.org/10.1029/2021JB023166

Schmucker U (1985) Magnetic and electric fields due to electromagnetic induction by external sources. In: Schmucker (eds.) Landolt-Börnstein, New-Series, 5/2b. Springer, Berlin, p 100-125

Thébault E, Hulot G, Langlais B, Vigneron P (2021) A spherical harmonic model of Earth's lithospheric magnetic field up to degree 1050. Geophys Res Lett 48(21):e95147. https://doi.org/10.1029/2021GL095147

Wardinski I, Holme R (2011) Signal from noise in geomagnetic field modelling: denoising data for secular variation studies. Geophys J Int 185(2):653–662. https://doi.org/10.1111/j.1365-246X.2011.04988.x

Wardinski I, Saturnino D, Amit H, Chambodut A, Langlais B, Mandea M, Thébault E (2020) Geomagnetic core field models and secular variation forecasts for the 13th International geomagnetic reference field (IGRF-13). Earth Planets Space 72(1):155. https://doi.org/10.1186/s40623-020-01254-7

Waters CL, Anderson BJ, Liou K (2001) Estimation of global field aligned currents using the iridium® System magnetometer data. Geophys Res Lett 28(11):2165–2168. https://doi.org/10.1029/2000GL012725

Publisher's Note