TU-Berlin candidate model for IGRF-14

Baerenzung Julien[1], Kada Martin[1], Saynisch-Wagner Jan[2], and Holschneider Matthias[3]

[1]Institute of Geodesy and Geoinformation Science, Technische Universität Berlin
[2]Section 1.3: Earth System Modelling, Geoforschungszentrum Potsdam
[3]Institute for Mathematics, Universität Potsdam

We present the candidate models from the Technische Universität Berlin for the 14^{th} edition of the international geomagnetic reference field. These models are derived using the Kalmag geomagnetic field model, as described by [Baerenzung et al., 2022], which has been extended in time for the purposes of this study. Kalmag combines a Kalman filter with a smoothing algorithm, enabling the generation of both mean field estimates and associated uncertainties. From these we propose the three following candidate products:

- A main field solution for the 2020.0 epoch.
- A main field solution for the 2025.0 epoch.
- A secular variation solution for the $2025.0 - 2030.0$ period.

Data

The Kalmag model is based on geomagnetic field measurements, either in the form of vector field or intensity data, collected from 1900 to the present day. These data come from a combination of satellite missions, ground-based observatories, and land, airborne, and marine surveys. Five satellite missions were considered, specifically: POGO (1965-1971) [Cain and Sweeney, 1973], MagSat (1979-1980) [Langel and Estes, 1985], Oersted (since 1999) [Neubert et al., 2001], CHAMP (2000-2010) [Rother et al., 2000], and SWARM (since 2013) [Olsen et al., 2013] missions. For ground observatories, hourly mean vector field data from the World Data Center for Geomagnetism (WDC) spanning 1900 to 2000 were used, supplemented by definitive, if available, or quasi definitive, if not, minute-resolution data from INTERMAGNET since 2000.
In this study, we focus on the selection criteria applied to SWARM satellite data and INTERMAGNET observatory data from 2013 to September 2024, as these datasets are most relevant to the period of interest. Detailed descriptions of the preprocessing applied to other datasets can be found in [Baerenzung et al., 2022].
For the SWARM constellation, only data from the Alpha and Bravo satellites were used, with the latest baseline corrections applied. The data were sampled simultaneously at a rate of one datum every 20 seconds. The selection of these data was based on the following criteria:

- At low geomagnetic latitudes (between $\pm 60°$) only night time data were retained.
- The z-component of the Interplanetary magnetic field(IMF) was required to be positive.
- The geomagnetic activity index Hp30, as defined by [Yamazaki et al., 2022], had to be ≤ 2.0 at the time of observation.

For the INTERMAGNET ground-based observatory data, minute-resolution measurements were first averaged over 1-hour periods. These hourly means were then subjected to the same selection criteria as the SWARM data, with the exception of the IMF condition. Two distinct datasets were constructed from the observatory data: one for the magnetic field (Observatories MF) and one for secular variation (Observatories SV). The magnetic field dataset consists of the hourly mean values sampled every 6 hours. For the secular variation dataset, hourly means were averaged over 0.1-year intervals, and annual differences were calculated.
The number of data points used for each 0.1-year period between 2013 and September 2024 is shown in Figure 1.

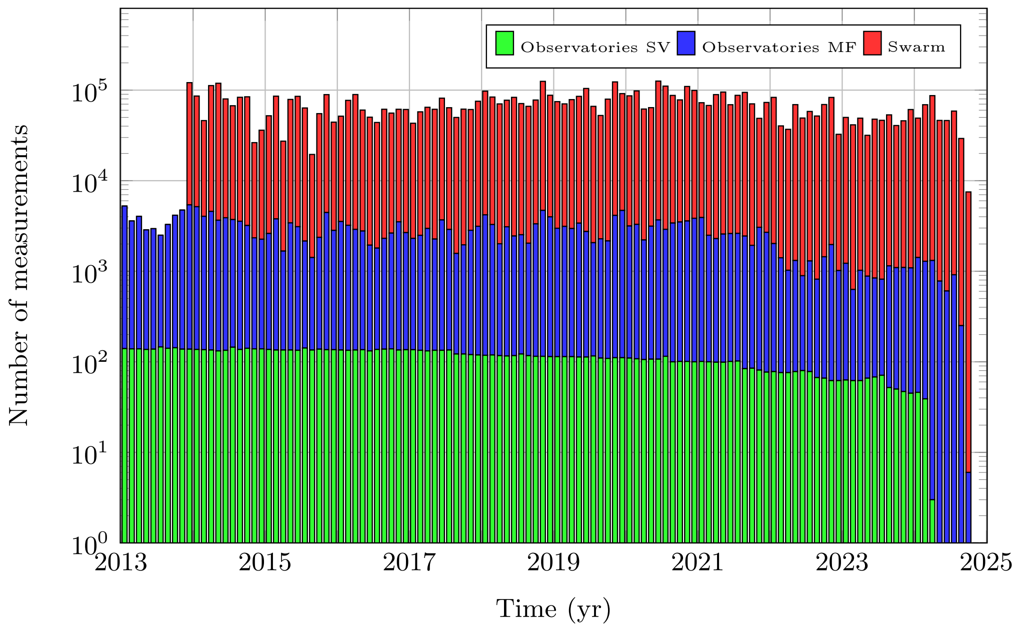

Figure 1: Number of vector field measurements assimilated every 0.1 year.

Model

Magnetic sources

The model consists of seven distinct sources: the core field (b_c), lithospheric field (b_l), induced/residual ionospheric field (b_{ii}), remote magnetospheric field (b_{rm}), near Earth magnetospheric field (b_m), fluctuating magnetospheric field (b_{fm}), and a source associated with field-aligned currents (b_{fac}).
All sources, except for b_{fac}, are assumed to be derived from scalar potentials V_s, where the magnetic field components are expressed as $b_s = -\nabla V_s$. For the field-aligned currents source (b_{fac}), the associated currents are assumed to arise from a vector potential V_{fac}, with the magnetic field defined as $b_{fac} = -\mathbf{r} \times \nabla V_{fac}$.
Each potential V_s is expanded in spherical harmonics (SH), with the contributions from internal and external sources represented respectively by the following formulations:

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

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

where $Y_{\ell,m}$ are Schmidt semi-normalized spherical harmonics of degree ℓ and order m, considered up to a maximum degree $\ell_{\max}$ and order $m_{\max}$. The reference radius is denoted by a_s, and the spherical harmonic coefficients $g_{s,\ell,m}(t)$ (hereafter referred to as g_s) are evaluated at this reference radius. Each field component is projected onto a specific spherical coordinate system r, θ_s, ϕ_s, as outlined in Table 1. These coordinate systems may correspond to geographic (GEO), magnetic (MAG), solar magnetic (SM), or geocentric solar magnetospheric (GSM) frames (see [Laundal, 2017] for further details).

Modeling technique

The Kalmag model is constructed using a sequential Kalman filter approach (see [Kalman, 1960]), which operates through two alternating steps: forecast and analysis. During the forecast phase, the model is propagated in space and time until new measurements become available. The analysis phase then updates the model based on these observations. Since the Kalman filter provides the posterior distribution conditioned only on the previously assimilated data, it is complemented by a smoothing algorithm that works backward in time, allowing for corrections to be applied retrospectively using the full dataset.
To predict the spatiotemporal evolution of the different sources we employ stochastic autoregressive processes (ARPs) which can be described by the following equation:

$$z_s(t + \Delta t) = F_s(\Delta t) z_s(t) + \xi_i(t, \Delta t) \tag{3}$$

Table 1: Magnetic sources considered in the model (first column) together with the coordinate systems they are expressed in (second column). GEO stands for geographic, SM for solar magnetic, MAG for magnetic and GSM for geocentric solar magnetospheric. ℓ_{max} and m_{max} are respectively the maximum degree and order of the SH expansion.

Source	Coordinate	ℓ_{max}	m_{max}
Core g_c	GEO	20	ℓ_{max}
Lithospheric g_l	GEO	150	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

where z_s represents the state vector of the s^{th} source, $F_s(\Delta t)$ is the autoregressive transition matrix, and $\xi_i(t, \Delta t)$ is a Gaussian white noise characterized by the spatial distribution $\mathcal{N}\left(0, \Sigma^{\infty}_{z_s} - F_s \Sigma^{\infty}_{z_s} F_s^T\right)$ with $\Sigma^{\infty}_{z_s}$ denoting the stationary state covariance matrix. For most sources, a first-order ARP is used, where $z_s(t)$ corresponds to the vector of spherical harmonic coefficients $g_s(t)$, and the transition matrix $F_s(\ell, \Delta t)$ is defined as:

$$F_s(\ell, \Delta t) = \exp\left[-|\tau|/\tau_s(\ell)\right] . \tag{4}$$

Here, $\tau_s(\ell)$ is a scale-dependent characteristic time, specific to each source.
For the core field, a second-order ARP is employed to capture the coupling between the field (g_c) and its time derivative ($\partial_t g_c$), with the state vector $z_c = (g_c, \partial_t g_c)^T$. The transition matrix for the core field process is given by:

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

The corresponding stationary covariance matrix for the core field is:

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

A description of the parameters $\Sigma^{\infty}_{z_s}$ and $\tau_c(\ell)$, which fully define the ARP dynamics, is given in appendix A for each source.
To forecast the evolution of all sources simultaneously, the transition matrices F_s from the autoregressive processes (ARPs) are combined into a global matrix $\mathbf{F}$. Likewise, the stationary covariance matrices Σ^{∞} for each source are combined into a global covariance matrix $\boldsymbol{\Sigma}^{\infty}$. The covariance matrix of the forecast error (Gaussian white noise) is given by $\tilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{\infty} - \mathbf{F}\boldsymbol{\Sigma}^{\infty}\mathbf{F}^T$. The evolution of the mean and covariance of the model from time step $k-1$ to step k is then expressed as:

$$\begin{aligned} E[\mathbf{z}_{k|k-1}] &= \mathbf{F}_{k-1} E[\mathbf{z}_{k-1}] & (7) \\ \boldsymbol{\Sigma}_{\mathbf{z}_{k|k-1}} &= \mathbf{F}_{k-1} \boldsymbol{\Sigma}_{\mathbf{z}_{k-1}} \mathbf{F}^T_{k-1} + \tilde{\boldsymbol{\Sigma}} \ . & (8) \end{aligned}$$

After the forecast step, when new measurements become available, the model is updated through a Bayesian inversion, following the Kalman filter approach. The Kalman gain matrix, $\mathbf{K}_k$, is computed as:

$$\mathbf{K}_k = \boldsymbol{\Sigma}_{\mathbf{z}_{k|k-1}} \mathbf{H}_k^T \left(\mathbf{H}_k \boldsymbol{\Sigma}_{\mathbf{z}_{k|k-1}} \mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \tag{9}$$

where $\mathbf{R}_k$ is the measurement error covariance matrix, and $\mathbf{H}_k$ is the matrix projecting the model onto the observations d_k at time step k. Using $\mathbf{K}_k$, the updated state and covariance are obtained as:

$$\begin{aligned} E[\mathbf{z}_{k|\mathbf{d}_k}] &= E[\mathbf{z}_{k|k-1}] + \mathbf{K}_k \left(d_k - \mathbf{H}_k E[\mathbf{z}_{k|k-1}]\right) & (10) \\ \boldsymbol{\Sigma}_{\mathbf{z}_{k|\mathbf{d}_k}} &= \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right) \boldsymbol{\Sigma}_{\mathbf{z}_{k|k-1}} \ . & (11) \end{aligned}$$

The measurement errors, $\mathbf{R}_k$, are modeled as diagonal matrices, with standard deviations of 0.1 nT for magnetic field components and 4.85 nT/yr for secular variation components.
Once all available data have been assimilated, a smoothing algorithm (see [Rauch et al., 1965]) is applied, starting from the final Kalman filter step and iterating backward in time. This process refines the model by utilizing the entire dataset, according to the following equations:

$$\begin{aligned} \mathbf{G}_{k-1} &= \boldsymbol{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}_{k-1}}} \mathbf{F}_k^T \boldsymbol{\Sigma}^{-1}_{\mathbf{z}_{k|k-1}} & (12) \\ E[\mathbf{z}_{k-1|\mathbf{d}}] &= E[\mathbf{z}_{k-1|\mathbf{d}_{k-1}}] + \mathbf{G}_{k-1} \left(E[\mathbf{z}_{k|\mathbf{d}}] - E[\mathbf{z}_{k|k-1}]\right) & (13) \\ \boldsymbol{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}}} &= \boldsymbol{\Sigma}_{\mathbf{z}_{k-1|\mathbf{d}_{k-1}}} + \mathbf{G}_{k-1} \left(\boldsymbol{\Sigma}_{\mathbf{z}_{k|\mathbf{d}}} - \boldsymbol{\Sigma}_{\mathbf{z}_{k|k-1}}\right) \mathbf{G}^T_{k-1} \ . & (14) \end{aligned}$$

The timestep of the Kalman filter and the smoothing algorithms were respectively set to $\Delta t = 30$ minutes and $\Delta t = 0.1$ year.

Candidate products

The models proposed as candidates for the IGRF-14 for epoch 2025.0 are derived from the Kalman filter solution at the latest available data point, corresponding to epoch 2024.74 (September 26, 2024). To extend this solution to 2025.0, we apply the forecast step of the Kalman filter, as follows:

$$E[\mathbf{z}](t = 2025.0) = \mathbf{F}_{\tau=0.26yr} E[\mathbf{z}](t = 2024.74) \tag{15}$$

$$\mathbf{\Sigma_z}(t = 2025.0) = \mathbf{F}_{\tau=0.26yr} \mathbf{\Sigma_z}(t = 2024.74) \mathbf{F}^T_{\tau=0.26yr} + \tilde{\mathbf{\Sigma}}_{\tau=0.26yr} . \tag{16}$$

Here, the forecast operator $\mathbf{F}_{\tau=0.26yr}$ is applied to both the model's mean estimate and its covariance matrix. The resulting secular variation model for 2025.0 is the mean estimated secular variation, denoted by $E[\partial_t g_c](t = 2025.0)$. The associated uncertainties are obtained by extracting the square root of the diagonal elements of the secular variation covariance matrix $\Sigma_{\partial_t g_c}(t = 2025)$, providing the standard deviation for each spherical harmonic coefficient of $\partial_t g_c$.

Our internal field model for epoch 2025.0 is computed as the sum of the mean core field and the mean lithospheric field at this epoch, i.e., $E[g_c] + E[g_l]$. The uncertainty associated with this model is derived from the covariance matrix:

$$\Sigma_{g_c+g_l} = \Sigma_{g_c} + \Sigma_{g_l} + \Sigma_{g_{cl}} + \Sigma^T_{g_{cl}} \tag{17}$$

where Σ_{g_c} and Σ_{g_l} represent the covariances of the core and lithospheric fields, respectively, and $\Sigma_{g_{cl}}$ is the cross covariance between the core and lithospheric fields. The square root of the diagonal elements of $\Sigma_{g_c+g_l}$ provides the standard deviations associated with the combined field model $E[g_c] + E[g_l]$.

For the DGRF 2020.0 candidate model, a similar approach is employed. However, instead of using the forecast solution, the core and lithospheric fields are taken from the smoothing solution, providing a more refined estimate for that epoch. The same covariance structure is used to compute the associated uncertainties.

A Autoregressive parameters

Stationary state covariance matrices $\Sigma^{\infty}_{z_s}$ and characteristic timescales $\tau_s(\ell)$ are completely characterizing the autoregressive process associated with each source g_s.
$\Sigma^{\infty}_{z_s}$ are derived from energy spectra $E_s(\ell, a_s)$ evaluated at radii a_s and follow the form:

$$\Sigma^{\infty}_{g_s}(\ell, m, \ell', m', r = a_s) = \frac{E_s(\ell, a_s)}{N_m R(\ell)} \delta(\ell - \ell')\delta(m - m') \tag{18}$$

where N_m is the number of SH coefficients per degree ℓ. $R(\ell)$ is defined as $\ell + 1$ for internal sources and ℓ for external sources. The energy spectra may either be flat, with $E_s(\ell) = A_s^2$ or of the C-based type, $E_s = A_s^2(2\ell + 1)R(\ell)$ as proposed by [Holschneider et al., 2016]. The parameters a_s, A_s and the dipole magnitudes D_s define the free parameters of the stationary state covariance matrices.
Characteristic timescales are parameterized as power laws of the form $\tau_s(\ell) = M_s \ell^{-\alpha_s}$ where M_s is a magnitude and α_s is a slope, which may vary across different harmonic degrees. The parameters of the ARPs were estimated using a machine learning algorithm on a subset of CHAMP and Swarm data, as detailed in [Baerenzung et al., 2020], and are summarized in Table 2.
For the field-aligned currents and the fluctuating magnetospheric field (at $\ell > 1$), the timescales are extremely short, $\tau_{fac}(\ell) = 1$ minute and $\tau_{fm}(\ell > 1) = 18$ minutes. Given that these timescales are smaller than the Kalman filter time step (30 minutes), these fields are modeled as temporally white noise, though spatial correlations are maintained during the analysis phase. Their covariance is expressed as:

$$E[g_s(\ell, t) g_s(\ell', t + \Delta t)] = \Sigma^{\infty}_s(\ell) \exp\left[-|\tau|/\tau_s(\ell)\right] \delta(\ell - \ell') . \tag{19}$$

References

[Baerenzung et al., 2022] Baerenzung, J., Holschneider, M., Saynisch-Wagner, J., and Thomas, M. (2022). Kalmag: a high spatio-temporal model of the geomagnetic field. *Earth, Planets and Space*, 74(1):139.

Table 2: Parameters for the stationary state covariance matrices and characteristic timescales associated with each source's autoregressive process.

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 $1 \leq \ell \leq 74$ $75 \leq \ell \leq 150$	C-Based FLAT	6287 6367.9	0.16 6.5	∞ ∞	0 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

[Baerenzung et al., 2020] Baerenzung, J., Holschneider, M., Wicht, J., Lesur, V., and Sanchez, S. (2020). The Kalmag model as a candidate for IGRF-13. *Earth, Planets, and Space*, 72(1):163.

[Cain and Sweeney, 1973] Cain, J. C. and Sweeney, R. E. (1973). The POGO data. *Journal of Atmospheric and Terrestrial Physics*, 35:1231.

[Holschneider et al., 2016] Holschneider, M., Lesur, V., Mauerberger, S., and Baerenzung, J. (2016). Correlation-based modeling and separation of geomagnetic field components. *Journal of Geophysical Research (Solid Earth)*, 121:3142–3160.

[Kalman, 1960] Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82:35–45.

[Langel and Estes, 1985] Langel, R. A. and Estes, R. H. (1985). The near-earth magnetic field at 1980 determined from magsat data. *Journal of Geophysical Research*, 90(B3):2495–2510.

[Laundal, 2017] Laundal, R. (2017). Magnetic Coordinate Systems. *Space Science Reviews*, 206:27–59.

[Neubert et al., 2001] Neubert, T., Mandea, M., Hulot, G., von Frese, R., Primdahl, F., Jørgensen, J. L., Friis-Christensen, E., Stauning, P., Olsen, N., and Risbo, T. (2001). Ørsted satellite captures high-precision geomagnetic field data. *EOS Transactions*, 82(7):81–88.

[Olsen et al., 2013] Olsen, N., Friis-Christensen, E., Floberghagen, R., Alken, P., Beggan, C. D., Chulliat, A., Doornbos, E., da Encarnação, J. T., Hamilton, B., Hulot, G., van den IJssel, J., Kuvshinov, A., Lesur, V., Lühr, H., Macmillan, S., Maus, S., Noja, M., Olsen, P. E. H., Park, J., Plank, G., Püthe, C., Rauberg, J., Ritter, P., Rother, M., Sabaka, T. J., Schachtschneider, R., Sirol, O., Stolle, C., Thébault, E., Thomson, A. W. P., Tøffner-Clausen, L., Velímský, J., Vigneron, P., and Visser, P. N. (2013). The Swarm Satellite Constellation Application and Research Facility (SCARF) and Swarm data products. *Earth, Planets and Space*, 65(11):1189–1200.

[Rauch et al., 1965] Rauch, H. E., Striebel, C. T., and Tung, F. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8):1445–1450.

[Rother et al., 2000] Rother, M., Michaelis, K., and Olsen, N. (2000). Resolution studies of fluid flow models near the core-mantle boundary using bayesian inversion. In Hansen, P., Jacobsen, B., and Mosegaard, K., editors, *Methods and Applications of Inversion*, volume 92 of *Lecture Notes in Earth Sciences*, pages 255–275. Springer Berlin Heidelberg.

[Yamazaki et al., 2022] Yamazaki, Y., Matzka, J., Stolle, C., Kervalishvili, G., Rauberg, J., Bronkalla, O., Morschhauser, A., Bruinsma, S., Shprits, Y. Y., and Jackson, D. R. (2022). Geomagnetic Activity Index Hpo. *Geophysical Research Letters*, 49(10):e98860.