IGRF-14 PGRF2030 candidate model submitted by NASA/GSFC

Kyle Gwirtz, Terence J. Sabaka, Weijia Kuang

Overview

The GSFC candidate model for the average secular variation (SV) from 2025.0 to 2030.0 is produced via data assimilation with a frozen flux model (see, e.g., Roberts and Scott, 1965). Assimilations are performed via an Ensemble Kalman Filter (EnKF, see, e.g., Evensen, 2006) accompanied by a smoother (see, e.g., Rauch et al., 1965) . For an initial ensemble of the core flow and magnetic field, we use snapshots of a long 3-D dynamo simulation. The Kalmag field model (Baerenzung et al. (2022), https://ionocovar.agnld.uni-potsdam.de/Kalmag/) is assimilated, through degree 13, at one year intervals, with Kalman smoothing in between. Below we provide a description of the dynamic model (frozen flux) and the assimilation procedure.

Dynamic model

The frozen flux model consists of a stream function ψ defining a horizontal fluid flow $\mathbf{v}$, just below the CMB, according to

$$v_\theta = \frac{1}{\sin\theta}\frac{\partial\psi}{\partial\phi}, \quad v_\phi = -\frac{\partial\psi}{\partial\theta}, \tag{1}$$

where v_θ and v_ϕ are the polar and azimuthal velocity components, respectively. The flow advects the radial component of the magnetic field B_r according to

$$\frac{dB_r}{dt} = -(\mathbf{v}\cdot\nabla)B_r. \tag{2}$$

Both the stream function ψ and the radial magnetic field B_r are defined by spherical harmonic expansions, truncated at degree and order 20,

$$B_r = \sum_{\ell=1}^{20}\sum_{m=0}^{\ell} b_\ell^m Y_\ell^m(\theta,\phi) + C.C., \quad \psi = \sum_{\ell=1}^{20}\sum_{m=0}^{\ell} \psi_\ell^m Y_\ell^m(\theta,\phi) + C.C., \tag{3}$$

where b_ℓ^m and ψ_ℓ^m are the spherical harmonic coefficients of B_r and ψ, respectively, $Y_\ell^m(\theta,\phi)$ are the degree ℓ and order m orthonormal spherical harmonic functions of colatitude θ and longitude ϕ, and $C.C.$ indicates the complex conjugate. Therefore, the state of the dynamic model can be completely described by a single vector

$$\mathbf{x} = (...b_\ell^m...\psi_\ell^m....)^T \tag{4}$$

containing the spherical harmonic expansion coefficients b_ℓ^m and ψ_ℓ^m. The time integration of Equation 2 is performed via a pseudo-spectral method with second order Runge-Kutta.

Data assimilation procedure

Data

We assimilate Gauss coefficients of the core field, through degree 13, from the Kalmag field model. The coefficients from the midpoint of ten successive years are assimilated (2015.5, 2016.5, 2017.5,..., 2024.5) with EnKF and smoothing iteration performed between each pair of assimilation times (see below and Figure 1). The Gauss coefficients of the field model are related to the spherical harmonics of the dynamic model according to

$$b_\ell^m = \mathcal{B}_\ell^m (-1)^m (\ell+1) r_s^2 \sqrt{\frac{2\pi(\delta_0^m+1)}{2\ell+1}} \left(\frac{r_s}{r_m}\right)^\ell (g_\ell^m - i h_\ell^m) \tag{5}$$

where $r_s = 6371.2$km, is the mean surface of the Earth, $r_m = 3467$ km is the dimensional radius of the model, $\delta_0^m = 1$ for $m = 0$ and $\delta_0^m = 0$ otherwise, g_ℓ^m and h_ℓ^m are the Gauss coefficients of the field model, and $\mathcal{B}_\ell^m$ are fixed scaling factors relating the physical field to the non-dimensional model.

Ensemble Kalman Filter

At each time an assimilation is performed, the "observations" of Equation 5 are collected into the vector $\mathbf{y}$ and are taken to be related to the "true" state of the dynamic model $\mathbf{x}^{\text{true}}$ according to

$$\mathbf{y} = \mathbf{H}\mathbf{x}^{\text{true}} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \tag{6}$$

where $\mathbf{H}$ is the observation operator relating the model state to the observations and $\boldsymbol{\eta}$ is a Gaussian random variable with mean zero and covariance $\mathbf{R}$. The observation covariance is taken to be diagonal, with the core-field uncertainties coming from the Kalmag model. A *forecast* ensemble $\{\mathbf{x}_i^f\}_{i=1}^{N_e}$ of N_e prior state estimates is adjusted by the observations to form an *analysis* ensemble $\{\mathbf{x}_i^a\}_{i=1}^{N_e}$ according to

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}(\mathbf{y} + \boldsymbol{\epsilon}_i - \mathbf{H}\mathbf{x}^f), \quad \mathbf{K} = \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1} \tag{7}$$

where $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, and $\mathbf{P}^f$ is the covariance of the forecast ensemble, i.e.,

$$\mathbf{P}^f = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (\mathbf{x}_i^f - \bar{\mathbf{x}}^f)(\mathbf{x}_i^f - \bar{\mathbf{x}}^f)^T, \tag{8}$$

with $\bar{\mathbf{x}}^f$ being the forecast ensemble mean. For the validation experiments and SV forecast we use an ensemble of $N_e = 512$ and apply Noise Informed Covariance Estimation (Vishny et al., 2024) to modify $\mathbf{P}^f$ and account for sampling error. The initial ensemble is a subset of the coefficients defining the coreflow and poloidal magnetic field from snapshots of a long run of a self-consistent 3-D numerical dynamo. Specifically, we use the initial ensemble of the **M1** dynamo model outlined in Gwirtz et al. (2024).

Ensemble Kalman Smoother

Between each pair of assimilation times, we apply an ensemble-based approach to the fixed-lag Rauch-Tung-Striebel smoother (Rauch et al., 1965) Let $\mathbf{X}_{k|K}$ be an array of N_e ensemble members (columns) $\mathbf{x}$, at time k, conditioned on data up to time K and let $\boldsymbol{\mu}_{k|K}$ be an array of N_e identical columns equal to the ensemble mean $(\bar{\mathbf{x}}_{k|K})$. Then the smoothing is performed according to

$$\mathbf{X}_{k|K} = \mathbf{X}_{k|k} + (\mathbf{X}_{k|k} - \boldsymbol{\mu}_{k|k})(\mathbf{X}_{k+1|k} - \boldsymbol{\mu}_{k+1|k})^{+}(\mathbf{X}_{k+1|K} - \mathbf{X}_{k+1|k}), \tag{9}$$

where the superscript $^{+}$ indicates the pseudoinverse (Penrose, 1955). To iterate the filter and smoother, we start a new EnKF run from time k, with an updated ensemble $\tilde{\mathbf{X}}_{k|k-1}$ from the smoother ensemble. To produce the updated ensemble, the mean of the original, prior ensemble at time k, is shifted to that of the smoother, i.e.,

$$\tilde{\mathbf{X}}_{k|k-1} = \mathbf{X}_{k|k-1} - \boldsymbol{\mu}_{k|k-1} + \boldsymbol{\mu}_{k|K}. \tag{10}$$

Between each year span (two observations) we iterate this process five times (see Figure 1).

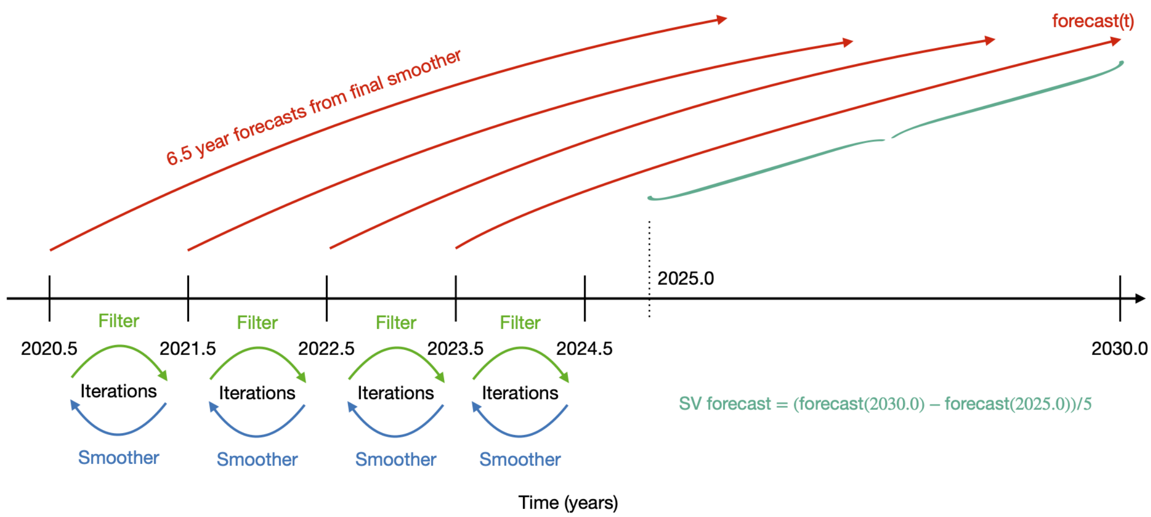

Figure 1: Illustration of the assimilation scheme used in producing the SV forecast. The filtering/smoothing procedure outlined in the text is iterated five times for each pair of observations (separated by one year) before moving to the next pairing of observations, where the process is repeated. At the end of each set of iterations, the smoother ensemble is run forward by the model 6.5 years to produce a forecast. The last such forecast reaches the year 2030.0 and is used to compute the forecasted average SV between 2025.0 and 2030.0

To test this configuration, it was used to generate SV forecasts for past periods where the PGRF was produced and DGRFs are available as a measure of the "truth". For example, we take the average annual change in coefficients between DGRF-2015 and DGRF-2010 to give $\dot{g}_\ell^{m,\mathrm{true}}$, $\dot{h}_\ell^{m,\mathrm{true}}$ for the 2010.0-2015.0 period. We then compare our forecasts for average SV for this period, $\dot{g}_\ell^{m,f}$, $\dot{h}_\ell^{m,f}$ according to

$$\mathrm{Error} = \sqrt{\sum_{\ell=1}^{8}\sum_{m=0}^{\ell}(\ell+1)[(\dot{g}_\ell^{m,f} - \dot{g}_\ell^{m,\mathrm{true}})^2 + (\dot{h}_\ell^{m,f} - \dot{h}_\ell^{m,\mathrm{true}})^2]}. \tag{11}$$

We find an average error of 19.40nT for our forecasts compared to 24.08nT for the PGRF releases for the same periods.

Final SV forecast

To produce the candidate SV forecast, the smoother ensemble of the last iteration is run forward by the dynamic model to 2030.0 (see Figure 1). The average annual change for each coefficient, between 2025.0 and 2030.0 is computed for each ensemble member, and using Equation 5, is converted to a predicted SV in the Gauss coefficients at $r_s = 6371.2$km. Let $\{\mathbf{SV}_i\}_{i=1}^{N_e}$ be the corresponding vectors of forecasted SV for Gauss coefficients from the $N_e = 512$ ensemble members. We report, for our forecasted SV and SV uncertainty, the ensemble mean and five times the standard deviation, as computed below

$$\bar{\mathbf{SV}} = \frac{1}{N_e}\sum_{i=1}^{N_e}\mathbf{SV}_i, \quad \boldsymbol{\sigma}_{\mathrm{SV}}^{\circ 2} = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}(\mathbf{SV}_i - \bar{\mathbf{SV}})^{\circ 2}, \tag{12}$$

where $^{\circ 2}$ indicates the element-wise square of the vector. The choice of reporting the value of five standard deviations ($5\boldsymbol{\sigma}$) for the uncertainties is based the testing of the forecast system against past PGRF periods. We found that 90% of SV coefficient forecasts during testing, fell within five ensemble standard deviations of the "true" SV.

References

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.

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

Gwirtz, K., Kuang, W., Yi, C., and Tangborn, A. (2024). Impact of localization and inflation on geomagnetic data assimilation. *Physics of the Earth and Planetary Interiors*, 355.

Penrose, R. (1955). A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(3):406–413.

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

Roberts, P. H. and Scott, S. (1965). On analysis of the secular variation 1. a hydromagnetic constraint: theory. *J. Geomag. Geoelect.*, 17:137–151.

Vishny, D., Morzfeld, M., Gwirtz, K., Bach, E., Dunbar, O. R. A., and Hodyss, D. (2024). High-dimensional covariance estimation from a small number of samples. *Journal of Advances in Modeling Earth Systems*, 16(9):e2024MS004417.