IGRF-14 DGRF2020/IGRF2025/PGRF2025 CANDIDATE MODELS SUBMITTED BY Wuhan University

Chao Xiong[1], Bohao Qian[1], Xinyi Rang[1], Fengjue Wang[1], Yuyang Huang[1]
1. Department of Space Physics, Wuhan University, 430072 Wuhan, Hubei, China.

Information about the candidate models

This document explains the processes by Wuhan University to construct the candidate models for IGRF-14, which include the geomagnetic field models for 2020 (DGRF2020), 2025 (IGRF2025), and a prediction of how the Earth's magnetic field will change from 2025 to 2030 (SV2025-2030).

Our team developed a temporal variation model of the Earth's main magnetic field, called the WuHan University Earth's Magnetic field Model (WHUEMM-01), which serves as the parent model for providing candidate models to IGRF-14. This model was constructed using magnetic data from several low Earth orbit satellites, including Swarm A (from January 2019 to July 2024), Swarm B (from January 2019 to July 2024), GRACE Follow-On 01 (from January 2019 to July 2024), and CryoSat-2 (from January 2019 to July 2024). The internal spherical harmonic coefficients, up to degree 13, are time-dependent. This temporal dependence is described by 6th-order B-splines (De Boor, 2001), with a 6-month knot separation and five-fold knots at the endpoints. All IGRF-14 model parameters are derived from WHUEMM-01. The spherical harmonic coefficients for the DGRF2020 model correspond to those of the WHUEMM-01 model as of January 1, 2020. The coefficients for IGRF2025 are obtained by linearly extrapolating the values of WHUEMM-01 from June 2024 to January 1, 2025. The SV2025-2030 model is based on the first derivative of the WHUEMM-01 coefficients as of January 1, 2024.

(1) Which satellite, observatory and/or repeat station data sets were used?

The scalar and vector magnetic field data from multiple satellite missions, including Swarm A, Swarm B, GRACE-FO-1, and CryoSat-2, were used to construct the WHUEMM-01 model. All satellite data covered the period from January 1, 2019, to July 31, 2024, and were obtained from ESA's Swarm data platform. The Swarm data were used with time resolution of 1 Hz products (versions 0605/0606) and included both scalar and vector components. The GRACE-FO-1 vector data are with also time resolution of 1 Hz, following the calibration procedures outlined by Stolle et al. (2021). The CryoSat-2 vector data products, following the calibration procedures described by Olsen et al. (2020), were also utilized in the construction of the core field model. No ground-based observatory data or repeat station measurements were included in this analysis, as our focus was solely on satellite data.

(2) What were the data selection and rejection criteria?

The following criteria have been applied for selecting the data:
- Only data with kp ≤ 2⁻ nT were used, and the variation rate of the storm-time disturbance index (dDst/dt) was limited to ≤ 3 nT/hr to minimize contamination from geomagnetic storms.

- Magnetic field observations were selected only for the time when the sun was at least 10° below the horizon to reduce the influence of ionospheric currents.
- Vector data from satellites are used for regions equator-ward of 55° quasi-dipole latitude (QDlat), where core magnetic fields dominate. In addition, scalar data were taken from regions pole-ward of 55° QDlat, where the scalar field is more sensitive to external fields.
- Solar wind merging electric field: data were excluded if the solar wind merging electric field (E_m) at the magnetopause, averaged over the preceding 2 hours, exceeded 0.8 mV/m. This threshold was chosen to eliminate data contaminated by strong solar wind-magnetosphere interactions. The E_m is defined as Equation (1) (Newell et al., 2007; Xiong et al., 2015):

$$E_m = \frac{1}{3000} V_{SW^3} (\sqrt{B_y^2 + B_z^2})^{\frac{2}{3}} \sin^{\frac{8}{3}}(\frac{\theta}{2}) \tag{1}$$

- Observations were rejected if the scalar field or vector components differed by more than 100 nT from the CHAOS-7x8 model, which provided a reliable baseline for identifying outliers.
- The CHAOS-7x8 model was used to subtract the external magnetic field contributions, including magnetospheric and lithospheric sources, leaving only the core field contributions.

(3) What weights were allocated to the different kinds of data?

As shown in Table 1, initial weights were determined based on the σ values of the vector and scalar components. For each satellite (Swarm A, Swarm B, GRACE-FO-1, and CryoSat-2), the data were assigned weights corresponding to their respective RMS values for the B_r, B_θ, B_φ, and the scalar total field (B_F) components. These weights were further updated during the inversion process by adjusting for the geographic colatitude of each data point. The final weighting function was $\sin\theta$, where θ is the geographic colatitude, to reduce the polar regions high density satellite data.

Table 1. Swarm A, Swarm B, GRACE-FO-1, CryoSat-2 residual statistics (unweighted) for CHAOS-7x8

Dataset	Region	Component	N	σ/nT
Swarm A	Polar	F_{polar}	63928	15.09
	Non-polar	B_r	316576	3.42
		B_θ	316576	5.11
		B_φ	316576	5.663
Swarm B	Polar	F_{polar}	64320	13.53
	Non-polar	B_r	321360	3.28
		B_θ	321360	4.71
		B_φ	321360	5.2
GRACE-FO-1	Non-polar	B_r	324307	5.23
		B_θ	324307	7.55
		B_φ	324307	5.27
CryoSat-2	Non-polar	B_r	322834	3.95
		B_θ	322834	5.54
		B_φ	322834	4.80

Robust estimation techniques were applied using Iteratively Reweighted Least Squares (IRLS) with

Huber weighting. For each iteration, residuals within 1.5σ of their respective distributions were treated as Gaussian noise, and residuals outside this range were treated as Laplacian noise as Equation (2) (Sabaka et al., 2004):

$$w_{i,k} = \frac{1}{\sigma_i^2} \, min\left(\frac{c\sigma_i}{|e_{i,k}|}, 1\right) \tag{2}$$

(4) Was data weighted for equal spatial or temporal coverage?

All data were weighted by $\mathbf{sin\theta}$, where $\boldsymbol{\theta}$ is their respective geographic colatitude, in order to mitigate high-density data in the polar regions.

(5) How was the forward extrapolation from October 2024 to 2025.0 carried out?

For constructing our model, satellite data were selected only until July 31, 2024. The forward extrapolation for the IGRF2025 model was achieved through a simple linear extrapolation of the Gauss coefficients from the WHUEMM-01 model, based on the final available data point at epoch 2024.0 as Equation (3):

$$\begin{cases} g_n^m(t = 2025.0) = g_n^m(t = 2024.0) + 1.0 \cdot \dot{g}_n^m(t = 2024.0) \\ h_n^m(t = 2025.0) = h_n^m(t = 2024.0) + 1.0 \cdot \dot{h}_n^m(t = 2024.0) \end{cases} \tag{3}$$

Here g_n^m, h_n^m represent each of the Gauss coefficients, while $\dot{g}_n^m$, $\dot{h}_n^m$ represent the Secular Variation (SV) in nT/yr. The resulting spherical harmonic coefficients for the internal field in epoch 2025.0 up to degree and order 13 were output to 0.01 nT.

(6) How is the average secular variation from 2025.0 to 2030.0 predicted?

SV 2025-2030 were derived from the first derivative of the spline-based WHUEMM-01 model. Due to the localized nature of the spline basis functions, the secular variation represents a weighted time average rather than an instantaneous rate of change. Regularization was applied to smooth the gaussian spherical harmonic coefficient, especially at the temporal endpoints of the model. As the endpoint effects are a common issue in spline models, additional constraints were imposed to minimize the impact of edge discontinuities on the SV coefficients. However, it is important to note that despite these measures, endpoint effects cannot be completely eliminated, and the resulting SV estimates may exhibit some sensitivity to the regularization parameters. Therefore, when evaluating the SV model for 2025-2030, we relied on the SV coefficients from the parent model at epoch 2024.0, rather than those at 2024.6, and did not attempt any extrapolation. The SV spherical harmonic coefficients (the first derivative of the spline model) for the internal field at epoch 2014.0, up to degree and order 8, were then output with a precision of 0.01 nT/year.

Moreover, we applied machine learning techniques to predict the SV coefficients for 2025–2030, using these predictions to validate the SV model we developed. This machine learning model was trained on historical IGRF Gauss coefficient data and provided an independent prediction of SV based solely on the past behavior of the magnetic field, without direct reliance on real-time magnetic field observations.

Given the limited sample size, we employed the classical Autoregressive Integrated Moving Average (ARIMA) model (Anderson et al., 1978), which is well-suited for small-sample time series forecasting, to predict the spherical harmonic coefficients of the geomagnetic field for the period from 2024.7 to 2030. Each order of coefficients was predicted independently. After obtaining the coefficients, finite differences were applied to compute the SV2025-2030.
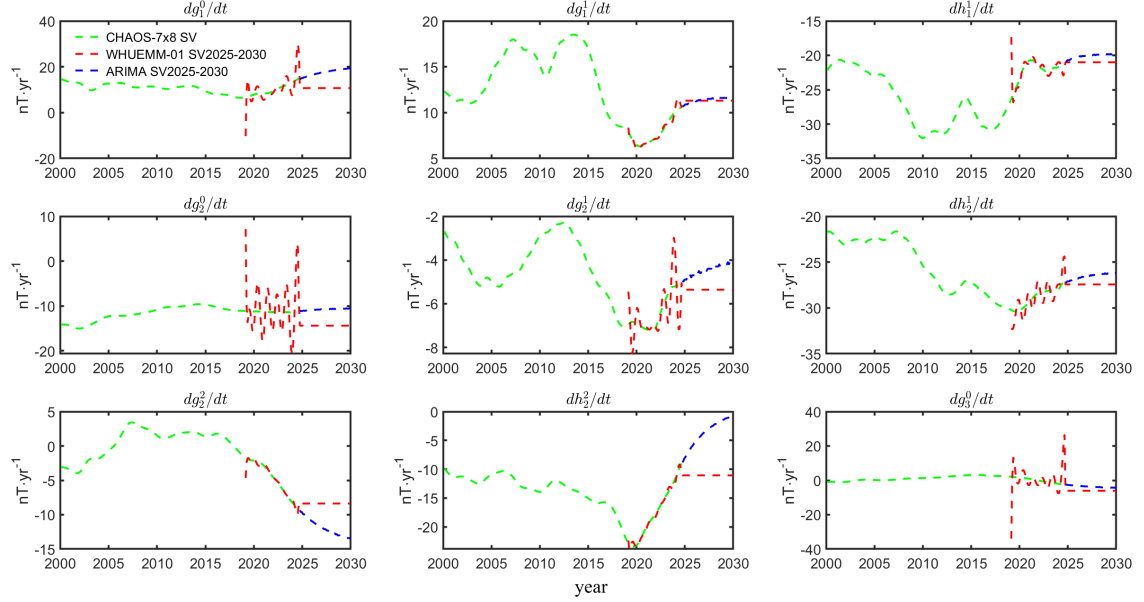
Figure 1. Comparison of SV model among models. The green line represents the SV from the CHAOS model, the red line represents the SV curve from the WHUEMM-01 model, and the blue line shows the ARIMA model prediction.

(7) If iterating using a least squares process, what starting model used and how many iterations were needed?

The CHAOS-7x8 core field model was used as the starting model for the least-square inversion processes, and this model provided the initial time-dependent polynomial coefficients for the core field, which were used to prior model parameter for WHUEMM-01. The inversion processes employed the Levenberg-Marquardt nonlinear least-square fitting algorithm (Aster et al., 2018), which iteratively updated the model parameters by minimizing the cost function. The step size for each iteration was determined by the convergence rate of the regularization term. After 10 iterations, the step size change became negligible, indicating that the model had reached a satisfactory level of convergence.

(8) If scalar data were used that required linearization of the inversion, what starting model used and how many iterations were needed?

When incorporating scalar data (total magnetic intensity, B_F), the non-linear relationship between B_F and the B_r, B_θ, and B_φ vector components required linearization.

$$B_F = \sqrt{B_r{}^2 + B_\theta{}^2 + B_\varphi{}^2} \tag{4}$$

This was accomplished using a Taylor series expansion around a reference point, allowing us to

express the scalar data in terms of small perturbations from the reference vector components.

$$\boldsymbol{B_F}(r,\theta,\lambda) = \boldsymbol{B_F}(r,\theta,\lambda)|_{g=g_0} + \frac{\partial \boldsymbol{B_F}(r,\theta,\lambda)}{\partial g}\Big|_{g=g_0} \bullet (g-g_0) \tag{5}$$

In nonlinear iteration, $\frac{\partial \boldsymbol{B_F}(r,\theta,\lambda)}{\partial g}\big|_{g=g_0}$ is the Jacobi matrix at g_0 in iterative calculation, as with the vector data, the CHAOS-7x8 model was used as the starting point for the inversion, and 10 iterations were performed using the Levenberg-Marquardt method.

(9) What, if any, regularization was used, e.g., use of an a-priori model with specified (co-variance, or addition of some quadratic penalty function to the sum square deviation?

Three main regularization terms were applied during the model construction:

(a) Minimization of the third derivative of the non-axial dipole radial component across the core-mantle boundary (CMB), integrated over the entire surface of the CMB. This regularization term was time-averaged over the model's temporal window, ensuring smoothness in the time derivative of the core field.

$$\left|\frac{\partial^3 B_r^{\text{NAD}}}{\partial t^3}\right|^2 = \frac{1}{t_e - t_s}\int_{t_s}^{t_e}\int \left|\frac{\partial^3 B_r^{\text{NAD}}}{\partial t^3}\right|^2 \mathrm{d}\Omega_c dt = \boldsymbol{m}^T \boldsymbol{\Lambda_1} \boldsymbol{m} \tag{6}$$

(b) Axial dipole (g10) regularization: The g10 coefficient was treated separately due to its sensitivity to unmodeled external disturbances, such as magnetospheric currents. By regularizing this term, we aimed to reduce its susceptibility to external contamination.

$$\left|\frac{\partial^3 B_r^{\text{AD}}}{\partial t^3}\right|^2 = \frac{1}{t_e - t_s}\int_{t_s}^{t_e}\int \left|\frac{\partial^3 B_r^{\text{AD}}}{\partial t^3}\right|^2 \mathrm{d}\Omega_c dt = \boldsymbol{m}^T \boldsymbol{\Lambda_2} \boldsymbol{m} \tag{7}$$

(c) Endpoint regularization: To mitigate the endpoint effects associated with spline models, we imposed a penalty on the second derivative of the radial component (Br) at the temporal endpoints of the model. This regularization helped smooth the transition between the model window and the extrapolated periods.

$$\int \left|\frac{\partial^2 B_r}{\partial t^2}\right|^2 \mathrm{d}\Omega_c|_{t=t_s} + \int \left|\frac{\partial^2 B_r}{\partial t^2}\right|^2 \mathrm{d}\Omega_c|_{t=t_e} = \boldsymbol{m}^T \boldsymbol{\Lambda_3} \mathbf{m} \tag{8}$$

For the selection of regularization coefficients as shown in formula (8), the parameter values for regularization are λ_1=0.33 (nT·years^{-3})$^{-2}$, λ_2=10 (nT·years^{-3})$^{-2}$, λ_3=10 (nT·years^{-3})$^{-2}$ like CHAOS-4 (Olsen et al., 2014). We want to note that inappropriate regularization parameters can lead to significant oscillations in the spherical harmonic coefficients at the time endpoints. To ensure stability, we chose 2024.0 as the reference time point for the calculations of SV2025-2030 and for extrapolating IGRF-2025.

(10) What, if any, sources were co-estimated and removed? Were any a-priori models or information used in the co-estimation (for instance mantle conductivity models in the estimation of induced magnetospheric and/or ionospheric fields)?

No co-estimation of non-core fields was performed. We assume that only the core field remains after

data preprocessing.

(11) What was the method used to solve the least squares equations?

We solved the least squares equations using the Levenberg-Marquardt algorithm, which combines the advantages of gradient descent and the Gauss-Newton method. The iterative inversion process aims to minimize the cost function (9):

$$\Phi(m, d_{obs}) = [d_{mod}(m) - d_{obs}]^T W [d_{mod}(m) - d_{obs}] + \lambda_1{}^2 m^T \Lambda_1 m + \lambda_2{}^2 m^T \Lambda_2 + \lambda_3{}^2 m^T \Lambda_3 m \tag{9}$$

The algorithm iteratively minimized the cost function by updating the model parameters according to the Jacobian matrix, which represented the partial derivatives of the residuals with respect to the model parameters. Each iteration involved calculating the step size $\delta \boldsymbol{m}_k$, which was constrained by the regularization terms $\Lambda_1 \Lambda_2 \Lambda_3$ described earlier.

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k + \delta \boldsymbol{m}_k \tag{10}$$

$$\delta \boldsymbol{m}_k = [\boldsymbol{J}_k^T \boldsymbol{W}_k \boldsymbol{J}_k + (\lambda_1 \Lambda_1 + \lambda_2 \Lambda_2 + \lambda_3 \Lambda_3) + u_0 I]^{-1}$$
$$[\boldsymbol{J}_k^T \boldsymbol{W}_k (\boldsymbol{d}_{obs} - \boldsymbol{d}_{mod}(\boldsymbol{m}_k)) - (\lambda_1 \Lambda_1 - \lambda_2 \Lambda_2 - \lambda_3 \Lambda_3) \boldsymbol{m}_k] \tag{11}$$

J_k is the partial derivative matrix of the residual relative to the parameter vector of the $\boldsymbol{m}_k$, u_0 is the adaptive update step feature value. To reduce the influence of outliers, we used robust reweighting of the residuals ($\boldsymbol{W}_k$), employing the Huber function for residuals that exceeded a certain threshold. This approach ensured that large outliers had minimal impact on the final model. A total of 10 iterations were performed, after which the step size norm was sufficiently small to indicate convergence.

(12) What was the fit to the data?

Table 2. Swarm A, Swarm B, GRACE-FO-1, CryoSat-2 residual statistics for WHUEMM-01，Mean and RMS refer to Huber weighted mean and rms values.

Dataset	Region	Component	N	Mean/nT	RMS/nT
Swarm A	Polar	F_{polar}	63928	3.69	13.63
	Non-polar	B_r	316576	0.14	2.77
		B_θ	316576	0.46	4.26
		B_φ	316576	-0.03	4.5
Swarm B	Polar	F_{polar}	64320	3.49	12.29
	Non-polar	B_r	321360	-0.02	2.58
		B_θ	321360	0.32	4.03
		B_φ	321360	-0.08	4.24
GRACE-FO-1	Non-polar	B_r	324307	0.09	4.12
		B_θ	324307	0.27	6.6
		B_φ	324307	-0.01	4.51
CryoSat-2	Non-polar	B_r	322834	-0.07	3.51
		B_θ	322834	-1.36	4.91
		B_φ	322834	-0.28	4.18

(13) Please give some indication of the uncertainties of the resulting set of coefficients and how these were inferred (e.g., (co-)variances, possible known biases, etc). If possible, provide an estimate of the uncertainties in the spatial domain.
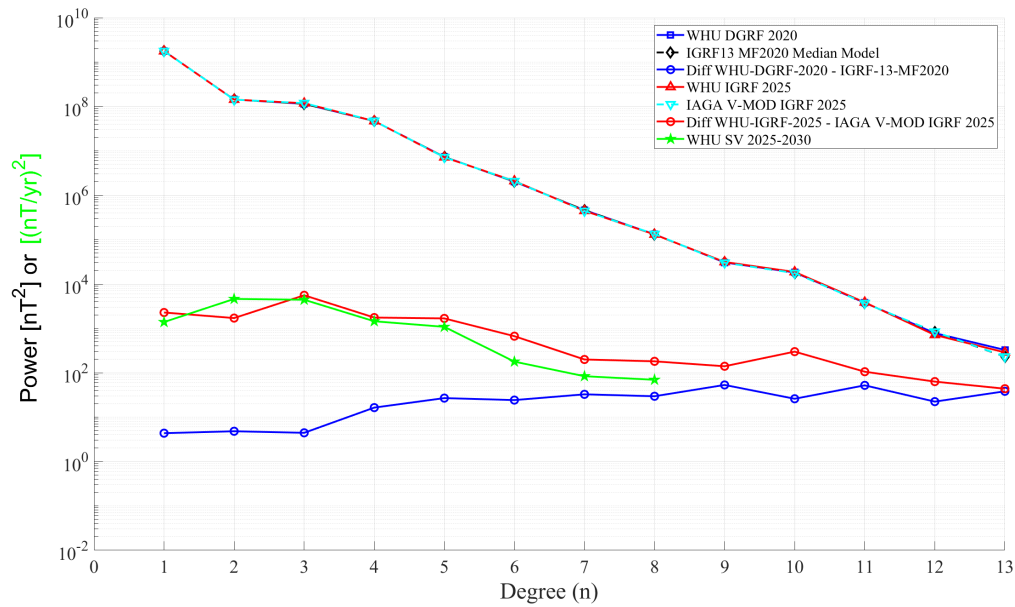

Figure 2. Spherical harmonic spectrum at the Earth's surface. Power spectrum of IGRF-14 related Gaussian coefficients submitted by Wuhan University.
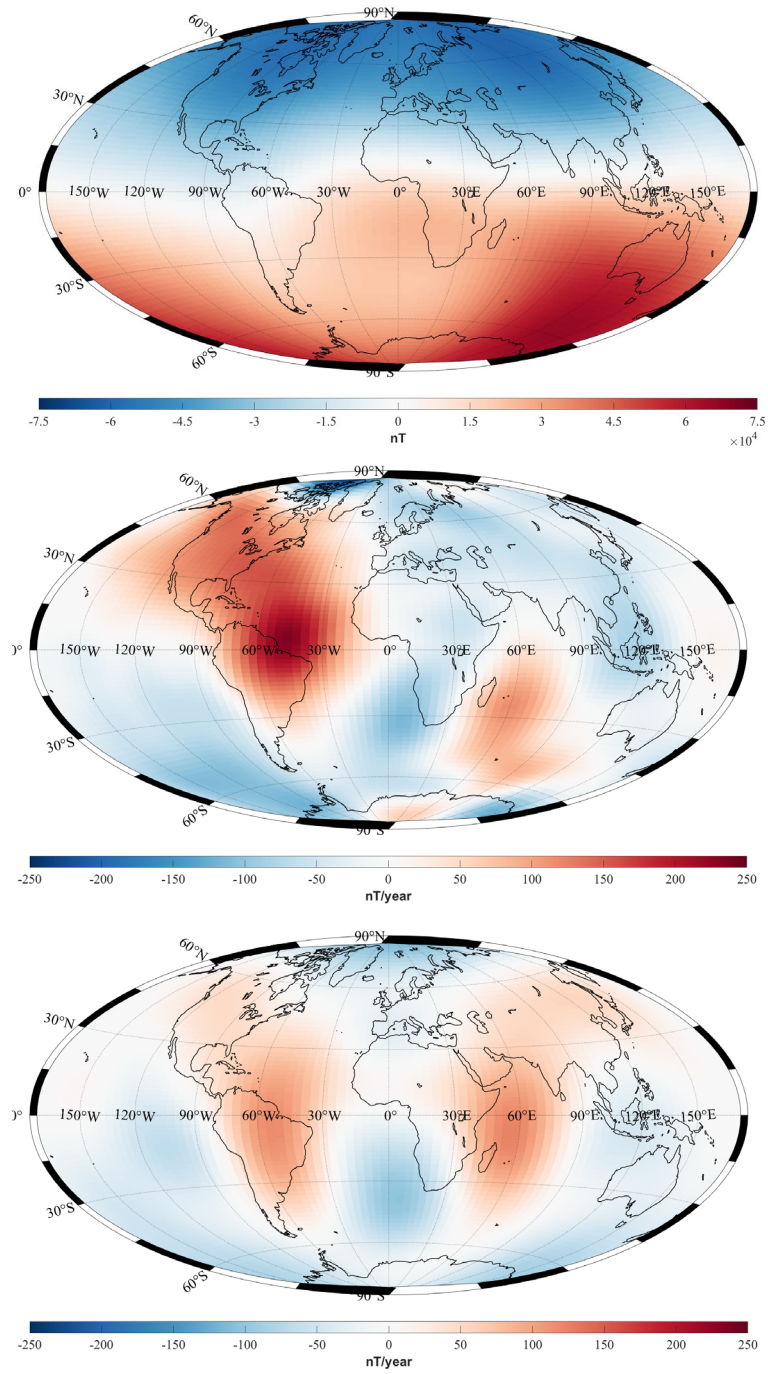
Figure 3. Global distribution of the radial magnetic field (top row), the secular variation (spherical harmonic degrees 1-13) (middle row), and the secular variation (spherical harmonic degrees 1-8) (bottom row) based on the WHUEMM-01 model.
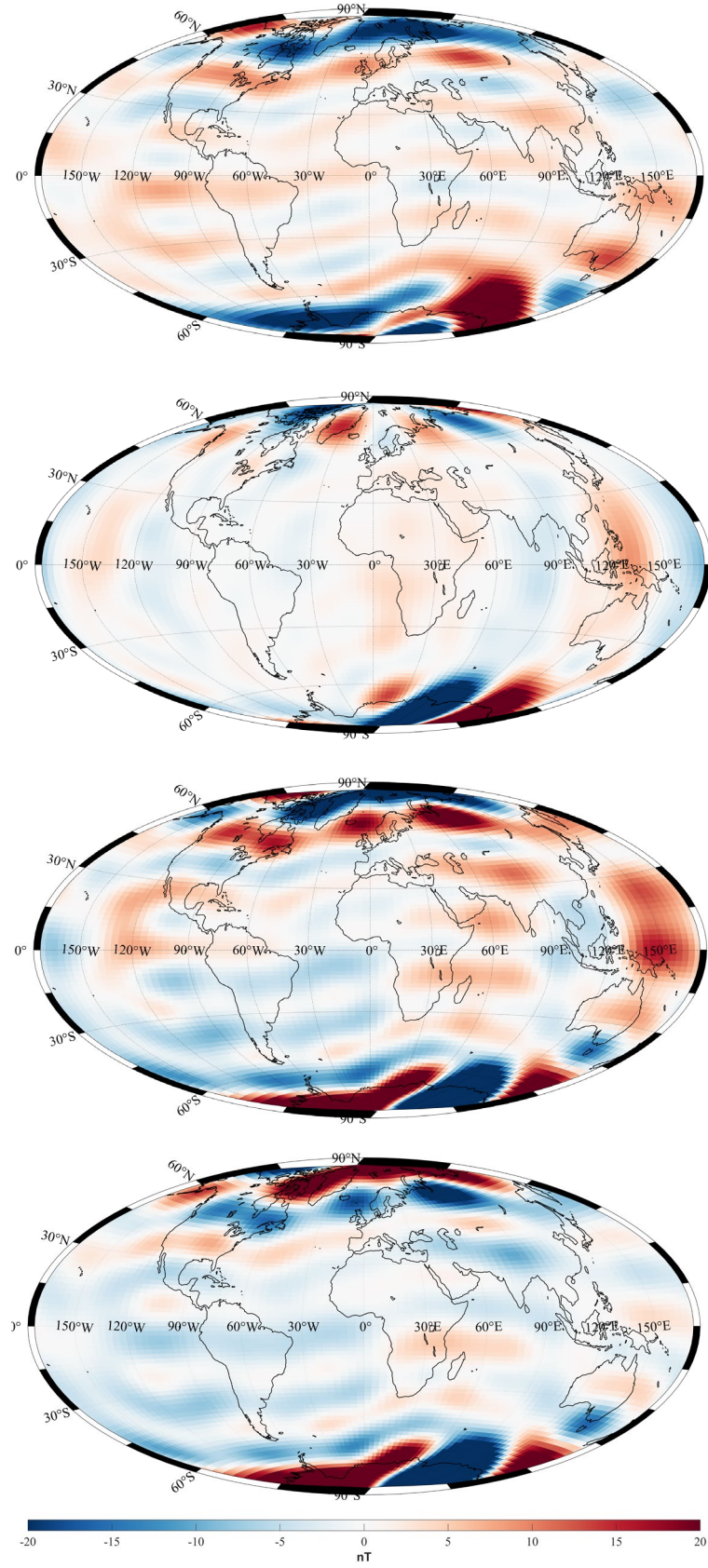
Figure 4. Differences between the DGRF-2020 model derived from WHUEMM-01 and the IGRF13 MF2020 Median Model, shown for the B_r, B_θ, B_φ, B_F components (from top to bottom, respectively).

No uncertainty estimates were provided with our candidate models, as we were unable to derive satisfactory estimates. The largest sources of error are likely due to two factors. First, biases introduced by unmodeled sources, such as the very dynamic polar electrojet activities, are difficult to quantify using formal model error covariance matrices and thus they cannot be reliably evaluated by constructing models from independent datasets using the same methodology. Second, the sequential inversion method we employed, which subtracts the lithospheric magnetic field and external magnetospheric effects by using the CHAOS model during the preprocessing stage, still results in significant errors in polar regions. This indicates that existing comprehensive models may be affected by significant coupling errors, particularly in high-latitude regions. Improving the accuracy of polar region modeling and achieving a precise separation of internal and external sources remains a substantial challenge for future work.

Acknowledgements

References

Anderson, O. D., Box, G. E. P., & Jenkins, G. M. (1978). Time Series Analysis: Forecasting and Control. The Statistician, 27(3/4), 265. https://doi.org/10.2307/2988198

Aster, R. C., Borchers, B., & Thurber, C. H. (2018). Parameter Estimation and Inverse Problems. Elsevier.

De Boor, C. (2001). A practical guide to splines. Applied Mathematical Sciences, 27.

Newell, P. T., Sotirelis, T., Liou, K., Meng, C.-I., & Rich, F. J. (2007). A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables. *Journal of Geophysical Research: Space Physics*, *112*(A1). https://doi.org/10.1029/2006JA012015

Olsen, N., Lühr, H., Finlay, C. C., Sabaka, T. J., Michaelis, I., Rauberg, J., & Tøffner-Clausen, L. (2014). The CHAOS-4 geomagnetic field model. *Geophysical Journal International*, *197*(2), 815–827. https://doi.org/10.1093/gji/ggu033

Olsen, N., Albini, G., Bouffard, J., Parrinello, T., & Tøffner-Clausen, L. (2020). Magnetic observations from CryoSat-2: calibration and processing of satellite platform magnetometer data. *Earth, Planets and Space*, *72*(1), 48. https://doi.org/10.1186/s40623-020-01171-9

Sabaka, T. J., Olsen, N., & Purucker, M. E. (2004). Extending comprehensive models of the Earth᾽s magnetic field with Ørsted and CHAMP data. *Geophysical Journal International*, *159*(2), 521–547. https://doi.org/10.1111/j.1365-246X.2004.02421.x

Stolle, C., Michaelis, I., Xiong, C., Rother, M., Usbeck, Th., Yamazaki, Y., et al. (2021). Observing Earth᾽s magnetic environment with the GRACE-FO mission. *Earth, Planets and Space*, *73*(1), 51. https://doi.org/10.1186/s40623-021-01364-w

Xiong, C., Lühr, H., & Fejer, B. G. (2015). Global features of the disturbance winds during storm time deduced from CHAMP observations. *Journal of Geophysical Research: Space Physics*,

120(6), 5137–5150. https://doi.org/10.1002/2015JA021302