Global geomagnetic model errors as a function of altitude and geomagnetic activity

Manoj Nair[1,2], Martin Fillion[3], Arnaud Chulliat[1,2], and Sam Califf[1,2]

1) Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

2) NOAA's National Centers for Environmental Information, Boulder, CO, USA

3) Institut de Physique du Globe de Paris, Paris, France

Corresponding author: Manoj Nair (manoj.nair@colorado.edu)

Key Points:

- Model errors decrease with altitude to a minimum around 400 kilometers, then increase again as external field effects become more dominant.
- Geomagnetic storms significantly increase model errors and reduce the maximum altitude at which the model remains accurate for navigation.
- This study provides the first error model for global magnetic model from the ground to ten thousand kilometers altitude.

Abstract

The errors of global geomagnetic models, such as the International Geomagnetic Reference Field (IGRF) and the World Magnetic Model (WMM), are well-characterized at the Earth's surface, but their error behavior as a function of altitude and geomagnetic activity has not been quantified. A lack of error characterization at operational altitudes and during space weather events, however, introduces unquantified risk for system designers and satellite operators. This study provides the first comprehensive, multi-source error analysis for the WMM from the surface to 10,000 km. We combine models for crustal, ionospheric, and magnetospheric fields to systematically quantify WMM performance across all levels of geomagnetic activity (G0-G4). Our results define the operational envelope where the WMM meets its Military Specification (MilSpec) accuracy. We demonstrate that while the model is robust at high altitudes (up to1,900 to 10,000 km, depending on the component) during quiet conditions, its validity ceiling can drop by thousands of kilometers during moderate to severe storms. We present maximum altitude validity maps and error profiles that provide a quantitative framework for mission planning, and real-time risk assessment. These model-based findings are validated using in-situ measurements from the CHAMP, Swarm and CryoSat-2 satellites. This work provides a resource for engineers and operators to assess and

mitigate risks associated with using standard geomagnetic models for high-altitude applications during space weather events.

Plain Language Summary

Satellites that provide navigation, communication, and imaging services often use the Earth's magnetic field to determine their orientation and position. They often rely on a standard model called the World Magnetic Model. While this model is highly accurate at ground level, its reliability for satellites at high altitudes, especially during geomagnetic storms caused by the Sun, has been a major unknown. This uncertainty creates risks for satellite operations.

To address this, our study created the first complete "error map" for the World Magnetic Model, detailing its performance from the Earth's surface up to 10,000 kilometers. By combining several advanced computer models and checking our results against real-world measurements from three satellite missions, we identified how and where errors occur.

We found that the model's performance changes significantly with altitude and is strongly affected by space weather. During calm conditions, the model is reliable up to very high altitudes. However, during a strong geomagnetic storm, the maximum safe altitude where the model can be trusted can drop by thousands of kilometers.

1. Introduction

Global geomagnetic models, such as the International Geomagnetic Reference Field (IGRF, Alken et al., 2020) and the World Magnetic Model (WMM, Chulliat et al., 2025) represent the Earth's main magnetic field, primarily generated by the planet's core, and account for its slow temporal changes known as secular variation. The Earth's magnetic field is influenced by various sources, including magnetic minerals in the crust and electric currents in the oceans, ionosphere, and magnetosphere. When comparing magnetic field measurements with predictions from a core-field model, three types of errors must be considered: measurement errors, modeling errors of commission, and modeling errors of omission. Measurement errors encompass inaccuracies inherent to the measurement process, including instrument limitations and distortions caused by nearby man-made objects. Modeling errors of commission arise from the model's inability to perfectly represent the complex geophysical phenomena it aims to model, primarily the Earth's main field and its secular variation. These errors are influenced by factors such as the availability of global satellite magnetic measurements used in developing the model or simplifying assumptions in the modeling process. Modeling errors of omission, on the other hand, stem from the models' exclusion of natural magnetic field contributions not originating from the Earth's core. These primarily include crustal fields from magnetized rocks in the Earth's outer layer and disturbance fields caused by electric currents in the ionosphere and magnetosphere. The separation of commission and omission errors is useful for both model developers, who need to improve the core model, and for operators, who need to understand the impact of ignored sources.

In the context of this paper, we study the WMM model. However, the results are equally applicable to other core-field models such as the IGRF. The WMM is a spherical harmonic model of the

Earth's main magnetic field, up to degree and order 12. It is developed jointly by the US NOAA National Centers for Environmental Information, University of Colorado's Cooperative Institute for Research in Environmental Sciences (CIRES) and the UK's British Geological Survey (BGS). The WMM is used for navigation, attitude and heading reference systems, and is updated every five years to account for changes in the Earth's magnetic field.

2. Background

Early geomagnetic models were limited by sparse ground observatory data until satellite magnetometer data (e.g., the 1960s - 70s POGO series) enhanced global coverage. A key evaluation (Peddie and Zunde, 1985) occurred after the MAGSAT mission (1979–1980), when the IGRF1980 model - derived largely from satellite data - was shown to be as accurate near the surface as models based solely on ground observatory measurements. As data quality and modeling techniques improved, reported global root-mean-square (RMS) errors fell significantly; for instance, Lowes (2000) found that while average errors for modern satellite-era epochs are on the order of tens of nanoteslas (nT), older models can show differences reaching a few hundred nT, especially when extrapolated near the end of their validity periods. For context, these errors are compared against a main field that typically ranges from 30,000 to 50,000 nT.

In a first systematic study of the errors of core-field models at ground levels, Chulliat et al. (2015, updated in 2020 and 2025) estimated the uncertainty associated with the WMM using an extensive compilation of independent geomagnetic measurements spanning multiple spatial and temporal scales. They used independent global magnetic models, a global network of 175 geomagnetic observatories (2000-2018), and an extensive database of marine and aeromagnetic surveys (6.8 million measurement points, 2000-2014). Crustal field contributions were evaluated through direct comparison with survey data and historical analysis of past WMM versions against observatory records. External field effects were assessed through analysis of observatory hourly means, revealing latitudinal dependence particularly near the geomagnetic equator and auroral zones. The combined uncertainty was computed as the root sum square of commission and omission errors.

The uncertainty computed by Chulliat et al. (2015, updated in 2020 and 2025) demonstrates compliance with the military specification requirements for the WMM throughout the model's operational lifetime. The Military Specification (latest version, MIL-PRF-89500B, Department of Defense, 2019) sets out threshold accuracy requirements - often referred to as "MilSpec" - that the WMM must meet over its 5-year operational lifespan. These MilSpec limits are expressed as RMS error for each magnetic element and serve as upper bounds on the acceptable discrepancy (in a 1-sigma sense) between the WMM and the true geomagnetic field. While the MilSpec provides a "fixed performance threshold" - the model's errors must not exceed this specification during the 5-year validity period - the WMM error estimates offer a "data-driven estimate" of the WMM's current or expected accuracy in practice. Table 1 provides WMM MilSpec and error estimates that consider both modeling errors of commission and omission, representing a combined assessment of the model's accuracy in predicting magnetic elements. The accuracy of the WMM deteriorates toward the end of each five-year epoch due to accumulated errors in forecasting secular variation (SV) (Table 1). Note that SV error is a failure of model to perfectly forecast the core field's evolution and it is traditionally classified as a commission error, despite the fact that the non-linear component of SV is not represented in the WMM parametrization. This deterioration is particularly pronounced near the magnetic poles, where the horizontal field component weakens. Very close

to each magnetic pole, within regions referred to as Blackout Zones (BoZ, characterized by a horizontal intensity of Earth's magnetic field (H) of less than 2000 nT), where magnetic navigation becomes unreliable, data were excluded from the WMM Declination error analysis by Chulliat et al. (2015, updated in 2020 and 2025).

		X (nT)	Y (nT)	Z (nT)	H (nT)	F (nT)	I (°)	D (°)	GV_N (°)	GV_S (°)
1	Military specification MIL-W-89500B	140*	140*	200*	200	280	1.00	1.00	1.00	1.00
2	Commission error at 2020.0	3	3	5	3	4	0.01	0.01	0.04	0.02
3	Commission Error at 2025.0	46	53	84	47	64	0.11	0.20	0.49	0.21
4	Combined error at 2025.0	135	101	168	134	144	0.23	0.42	0.83	0.70
5	WMM Error Model	131	94	157	128	148	0.21	0.42°		

Table 1. World Magnetic Model 2020 (WMM2020) performance specifications and error estimates for various magnetic components. Includes the Military Specification requirements (MIL-W-89500B), estimated Commission Errors for epoch 2025.0, estimated Combined Errors (root-sum-square of commission and omission errors) for epoch 2025.0, and the WMM Error Model values. Reproduced from Chulliat et al. (2020). Thresholds shown for X, Y, and Z (marked with *) are from an earlier version of the military standard. The geomagnetic components F, H, D and I are defined in section 4. GVN and GVS are defined in section 6.

Beggan (2022) estimated the uncertainty of the IGRF based on the globally averaged misfit of the model to ground-based measurements at repeat stations and observatories between 1980 and 2021. The study found that the 68.3% confidence interval is 87 nT in the North (X) component, 73 nT in the East (Y) component, and 114 nT in the vertical (Z) component at the ground level. At higher altitudes, errors in geomagnetic models are increasingly influenced by external sources such as ionospheric and magnetospheric currents. Matteo and Morton (2011) conducted a study to evaluate the accuracy of the IGRF model using scalar intensity measurements from multiple Low Earth Orbit (LEO) satellites spanning altitudes from 350 km to 1,000 km. Their work was motivated by the need to characterize Global Positioning System (GPS) range measurement error due to higher-order ionosphere group delay and carrier advances. The study used data from four satellites: the Upper Atmosphere Research Satellite (UARS); the Ørsted satellite; the CHAMP satellite; and the SAC-C satellite. The results indicated that the IGRF model predicted the total magnetic field within 1% of measured values 92.80% of the time. However, accuracy varied with altitude and geomagnetic activity levels. Errors tended to increase during periods of higher geomagnetic activity, as measured by the Kp index, a planetary-scale measure of geomagnetic disturbances (Matzka et al., 2021). The study found that during quiet conditions ($Kp \leq 2$), model errors were minimal, but during active conditions ($Kp > 5$), deviations from IGRF predictions increased significantly.

Most prior work has quantified these errors only at or near the Earth's surface and often only in scalar total-field terms. Yet, applications like satellite attitude control, magnetic navigation

(MagNav), and GNSS-based positioning rely on accurate three-component fields at orbital altitudes. MagNav determines position by measuring variations in the Earth's crustal magnetic field and comparing them to known magnetic anomaly maps (Canciani and Raquet, 2016). This technique leverages the relatively stable and unique magnetic signatures of the Earth's crust, making it a viable alternative or supplement to GPS, especially in environments where GPS signals are unreliable or unavailable. MagNav systems have been successfully demonstrated at operational altitudes ranging from approximately 150 meters to 6 km (Muradoglu el al., 2025). Moreover, increasing space weather variability means that external disturbance fields (excluded from the models) play a larger role at higher altitudes. Quantifying how WMM/IGRF errors evolve with altitude – and identifying where they exceed mission requirements – is thus crucial for improving navigation accuracy, guiding future model development, and ensuring the reliability of magnetic-field-based systems in space and military operations. Recent advances in empirical and physics-based magnetic field models now enable comprehensive estimation of omission errors from the Earth's surface to at least 10,000 km altitude – described in detail in section 3.

Recent satellite missions - such as CHAMP, Swarm, CHAMP and CryoSat-2 - provide in-situ vector magnetic field measurements at multiple altitudes, offering critical "spot checks" along the modeled error profiles. These measurements serve as an independent benchmark to validate modeled altitude trends, particularly under varying geomagnetic conditions. Together, the synthesis of modeling and satellite observations offers a unique opportunity to quantify the limitations of core-field models like WMM and IGRF throughout near-Earth space, and to support their operational use in satellite attitude control, magnetic navigation, and space situational awareness.

3. Methodology

Our approach to estimating the omission errors involves evaluating a series of models that represent the crustal, ionospheric and magnetospheric fields on a three-dimensional grid, and computing the RMS across all altitudes and for various levels of geomagnetic activity. For each grid cell, this RMS value quantifies the typical magnitude of the unmodeled field from a given source. This is calculated from the temporal variability of a time series of model outputs for the dynamic ionospheric and magnetospheric fields, and from the spatial variability of the high-resolution field within the cell for the static crustal model (Table 3). The models were evaluated on 29 discrete height levels - each representing a spherical shell - using a 10°×10° grid over the Earth's surface on each shell. We used a 10°×10° grid to balance computational efficiency with the need to capture large-scale WMM error patterns. This matches the spatial scales of the contributing models: WMM (degree/order 12; ~30°/3,000 km), DIFI (degree 45; ~8°/440 km), AMPS (degree 13; 28°, ~500–3,000 km), and T89 (variations over several thousand km). While the EMM crustal model extends to degree 790 (~0.46°/51 km at equator), its small-scale features decay quickly with altitude, and become negligeable at the upper atmospheric heights of interest. All altitudes in this paper are specified with respect to the WGS84 ellipsoid, using a geodetic reference system for latitude, longitude, and height. The local coordinate triad is X, Y, and Z in North, East, and Center (NEC) directions, respectively. The height levels (in kilometers) used in the simulations are every 10 km from 0 to 100, every 100 km from 100 to 1000, and every 1000

km from 1000 to 10000. Further, the simulations were repeated for geomagnetic activity levels based on the NOAA Space Weather Scale for geomagnetic storms (Table 2), which categorizes geomagnetic storm intensities into six levels, from "Minor" (G1) to "Extreme" (G5). These levels are defined using the planetary Kp index, which represents the maximum peak-to-peak fluctuation of the horizontal magnetic field measured by 13 selected observatories in three-hour periods after removing a baseline (Matzka et al., 2021). Kp is further divided into thirds, represented by a number and a symbol ("-", "o", or "+"). The G0 level (representing Kp indices 0-4, or "Normal" conditions) was added for completeness and is not part of the original G scale. We chose to use the G scale instead of the Kp index directly because the G scale is widely used for operational forecasting by government agencies and industry.

Scale	Description	Physical measure	Average Frequency (1 cycle = 11 years)
G 5	Extreme	Kp = 9	4 per cycle (4 days per cycle, ~0.1%)
G 4	Severe	Kp = 8, including a 9-	100 per cycle (60 days per cycle, ~2%)
G 3	Strong	Kp = 7	200 per cycle (130 days per cycle, ~4%)
G 2	Moderate	Kp = 6	600 per cycle (360 days per cycle, ~9%)
G 1	Minor	Kp = 5	1700 per cycle (900 days per cycle, ~22%)
G 0	Normal	Kp = 0,1,2,3,4	

Table 2. Geomagnetic activity scale used in this study, adapted from the NOAA Space Weather Prediction Center (SWPC) G-scale definitions. The table relates the G levels (G1-G5) to the corresponding Kp index ranges and provides their average occurrence frequency per solar cycle (~11 years). The G0 level (representing Kp indices 0-4, or "Normal" conditions) has been added for completeness in this analysis and is not part of the official NOAA/SWPC G1-G5 storm scale.

To assess the accuracy of the WMM at different altitudes and under various geomagnetic conditions, this study employed a set of models, each with specific characteristics and functions. To evaluate model uncertainties across different geomagnetic activity levels, we employed a two-step approach for all external field models (T89, AMPS, DIFI). First, we analyzed the 1963-2020 OMNI database to determine representative average solar wind, interplanetary magnetic field (IMF), and F10.7 conditions for each G-scale by binning the historical data according to corresponding Kp levels. We then used these climatological driver conditions (as needed for each model) as static inputs for each G levels to evaluate each model for seasonal, local time and

Magnetic Local Time (MLT) dependencies by simulating outputs for the 15th of every month over one year, with 24 hourly evaluations per day (288 time points per G-level). The final RMS value for each grid cell was then computed from this time series, representing the typical variability of the magnetic field disturbance for that location and activity level.

3.1 Global Magnetospheric Field Model – T89 (Tsyganenko, 1989)

The Tsyganenko 1989 (T89) model is an empirical representation of the Earth's magnetospheric magnetic field, developed from an extensive set of satellite magnetic field observations (HEOS, IMP, and ISEE) spanning a range of geomagnetic activity levels, as measured by the Kp index. We selected T89 because our focus is on altitudes above low-Earth orbit - where more recent models such as CHAOS (Finlay et al., 2020) are less optimized - and because it offers a straightforward Kp-based parameterization that suits our operational scaling needs, unlike more complex models such as TS07D (Tsyganenko and Sitnov, 2007). The model accounts for major current systems, including the tail, ring, and magnetopause currents, providing a broad depiction of the magnetospheric field under varying conditions. We evaluated T89 at a spatial resolution of 10°×10° for three geomagnetic activity levels: low (Kp 0–4), moderate (Kp 5), and high (Kp >6) (see Table 3).

Model	**Grid**	**Area**	**Samples**	**Geomagnetic activity level**
Magnetospheric field: T89 (Tsyganenko, 1989)	10°x10°	Global	15th of every month for 1 year and 24 hours on each 15th day. RMS of values in each cell.	Kp 0-4, 5, >6 corresponding to G0, G1, G2+
Crustal Field: Enhanced Magnetic Model (EMM-2017) (https://www.ncei.noaa.gov/products/enhanced-magnetic-model)	10°x10° (RMS from 1°x1° grid)	Global	Evaluated on a 1°x1° grid. RMS of 100 values in each cell.	N/A
Ionospheric field (high latitude): Average Magnetic field and Polar current System (AMPS), Laundal et. al., 2018	10°x10°	Poleward of |45|° geographic latitude	15th of every month for 1 year and 24 hours on each 15th day. RMS of values in each cell.	Kp 0-4,5,6,7, 8 and 9- corresponding to G0, G1, G2, G3, G4, G5
Ionospheric field (low latitude): Dedicated Ionospheric Field Inversion (DIFI), Chulliat et al., 2016	10°x10°	Below |55|° geomagnetic latitude	15th of every month for 1 year and 24 hours on each 15th day. RMS of values in each cell.	Kp 0-4,5,6,7, 8 and 9- corresponding to G0, G1, G2, G3, G4

Table 3. Models used in the study to estimate the omission errors as a function of altitudes and geomagnetic activity levels.

3.2 Crustal Magnetic Field – Enhanced Magnetic Model (EMM)

The Enhanced Magnetic Model (EMM) is a high-resolution global model that represents the Earth's crustal magnetic field. It is derived from a combination of satellite, aeromagnetic, and

marine magnetic survey data, providing a detailed map of magnetic anomalies associated with the Earth's lithosphere. The EMM extends to spherical harmonic degree 790, enabling it to capture fine-scale features of the crustal magnetic field.

3.3 Ionospheric Magnetic Field Contributions

The Earth's ionosphere, ranging from approximately 80 km to 1000 km in altitude, hosts electric currents that generate magnetic fields influencing the geomagnetic environment. We used two separate models to evaluate ionospheric contributions at high and low latitudes: AMPS (Average Magnetic field and Polar current System, Laundal et al., 2018) and DIFI (Dedicated Ionospheric Field Inversion, Chulliat et al., 2016). We needed to account for the fact that high-latitude ionospheric currents move towards lower latitudes as geomagnetic activity increases. To accommodate this, we set the lower latitude limit of AMPS to 45 degrees Quasi-Dipole latitude (QDLat; Richmond, 1995). However, during quiet-time conditions, it is unclear whether AMPS is a reliable representation of ionospheric currents below 55 degrees QDLat. Therefore, we chose to evaluate DIFI up to 55 degrees QDLat. In the overlapping region (45°-55°), the contributions from both models are combined in quadrature, as shown in Equations 1-3.

3.3.1 AMPS (Average Magnetic field and Polar current System, Laundal et al., 2018)

The AMPS model is a global empirical representation of the average ionospheric magnetic field and current systems for specific Solar wind and interplanetary magnetic field conditions. Developed using magnetic field measurements from the CHAMP and Swarm satellites, AMPS describes both horizontal currents in the ionospheric E region and field-aligned currents at LEO satellite altitudes. Its design makes the model particularly effective for simulating currents and fields at high latitudes for various levels of geomagnetic activity.

3.3.2 DIFI (Dedicated Ionospheric Field Inversion, Chulliat et al., 2016)

The DIFI model offers a time-varying, spherical harmonic representation of the quiet-time ionospheric Sq (solar quiet) and equatorial electrojet (EEJ) magnetic fields below 55° geomagnetic latitude. It is based on a combination of magnetic field measurements from the Swarm satellite constellation and ground-based observatories. The model accounts for both the primary magnetic field, generated by electric currents in the ionosphere, and the induced magnetic field resulting from induced currents in the Earth's mantle. For this project, the most recent version available at the time, DIFI-7, was used. Note that DIFI model is not dependent on Kp or solar-wind parameters as it is a climatological model driven by solar EUV- ionization (F10.7) rather than geomagnetic disturbances. DIFI's output therefore static for a given G-level and represents a baseline quiet-time uncertainty.

4. Combined error estimation

The error-estimation calculations were performed independently on the 29 altitude shells, for which we note the corresponding altitudes h_k (with k = 1, 2, ….,29). For each altitude shell, the calculations were carried out on a grid defined by latitude ϕ_j and longitude θ_i. For each grid cell (h_k, ϕ_j, θ_I), the omission RMS errors for a given G scale are given by:

$$\delta X_{\text{omission}}(h_k, \phi_j, \theta_i) = \sqrt{\delta X^2_{\text{AMPS}}(h_k, \phi_j, \theta_i) + \delta X^2_{\text{EMM}}(h_k, \phi_j, \theta_i) + \delta X^2_{\text{T89}}(h_k, \phi_j, \theta_i) + \delta X^2_{\text{DIFI}}(h_k, \phi_j, \theta_i)} \quad (1)$$

$$\delta Y_{\text{omission}}(h_k, \phi_j, \theta_i) = \sqrt{\delta Y^2_{\text{AMPS}}(h_k, \phi_j, \theta_i) + \delta Y^2_{\text{EMM}}(h_k, \phi_j, \theta_i) + \delta Y^2_{\text{T89}}(h_k, \phi_j, \theta_i) + \delta Y^2_{\text{DIFI}}(h_k, \phi_j, \theta_i)} \quad (2)$$

$$\delta Z_{\text{omission}}(h_k, \phi_j, \theta_i) = \sqrt{\delta Z^2_{\text{AMPS}}(h_k, \phi_j, \theta_i) + \delta Z^2_{\text{EMM}}(h_k, \phi_j, \theta_i) + \delta Z^2_{\text{T89}}(h_k, \phi_j, \theta_i) + \delta Z^2_{\text{DIFI}}(h_k, \phi_j, \theta_i)} \quad (3)$$

Where, for any component *Q,* the omission error $\delta Q_{omission}$ at a given grid cell (h_k, ϕ_j, θ_i) and for each G scale, is obtained by adding the individual errors, for the corresponding components, estimated from each model (for example, δZ_{AMPS}) in quadrature.

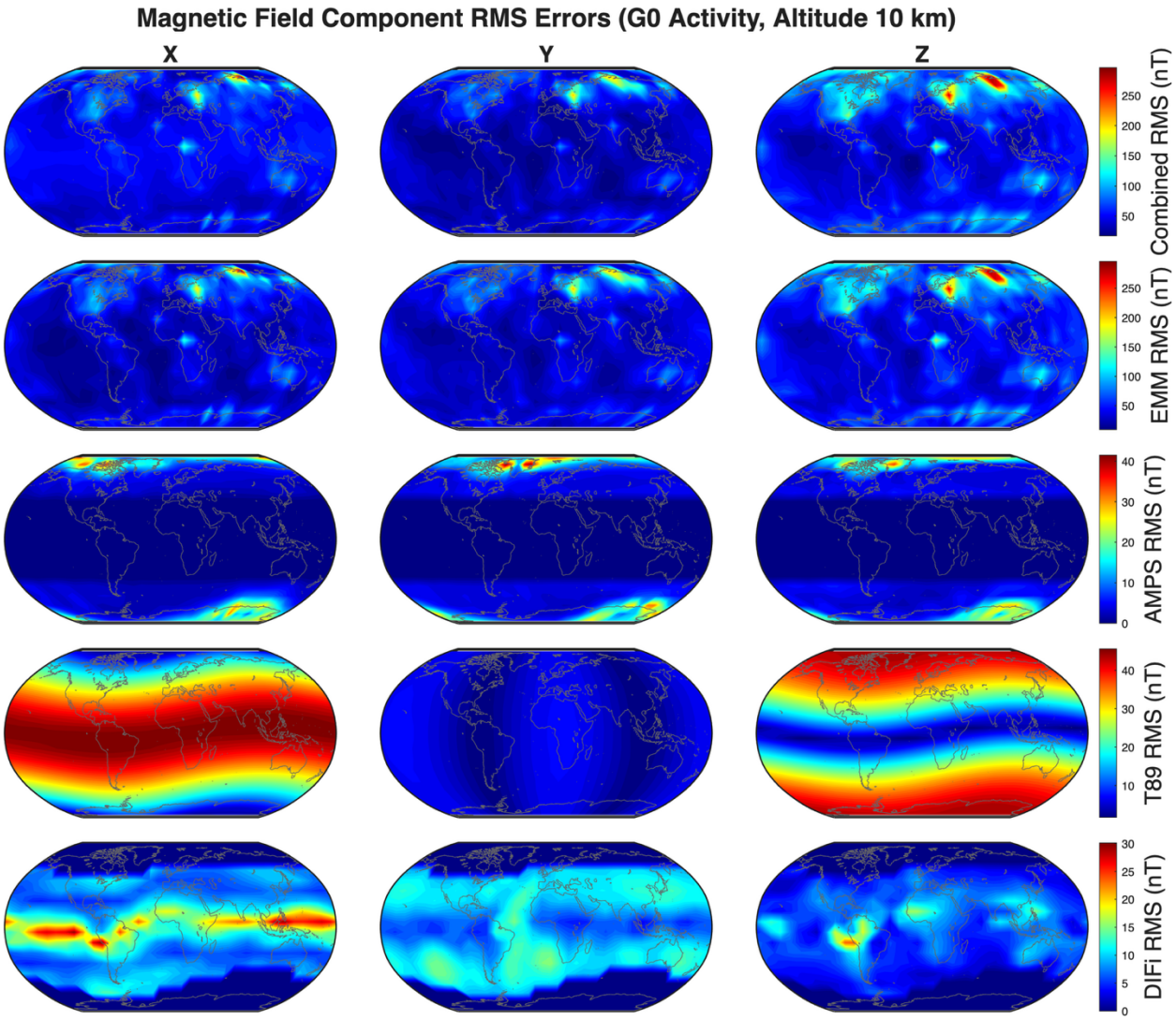

Figure 1a. Maps of the four individual models output for G0 at 10 km altitude.

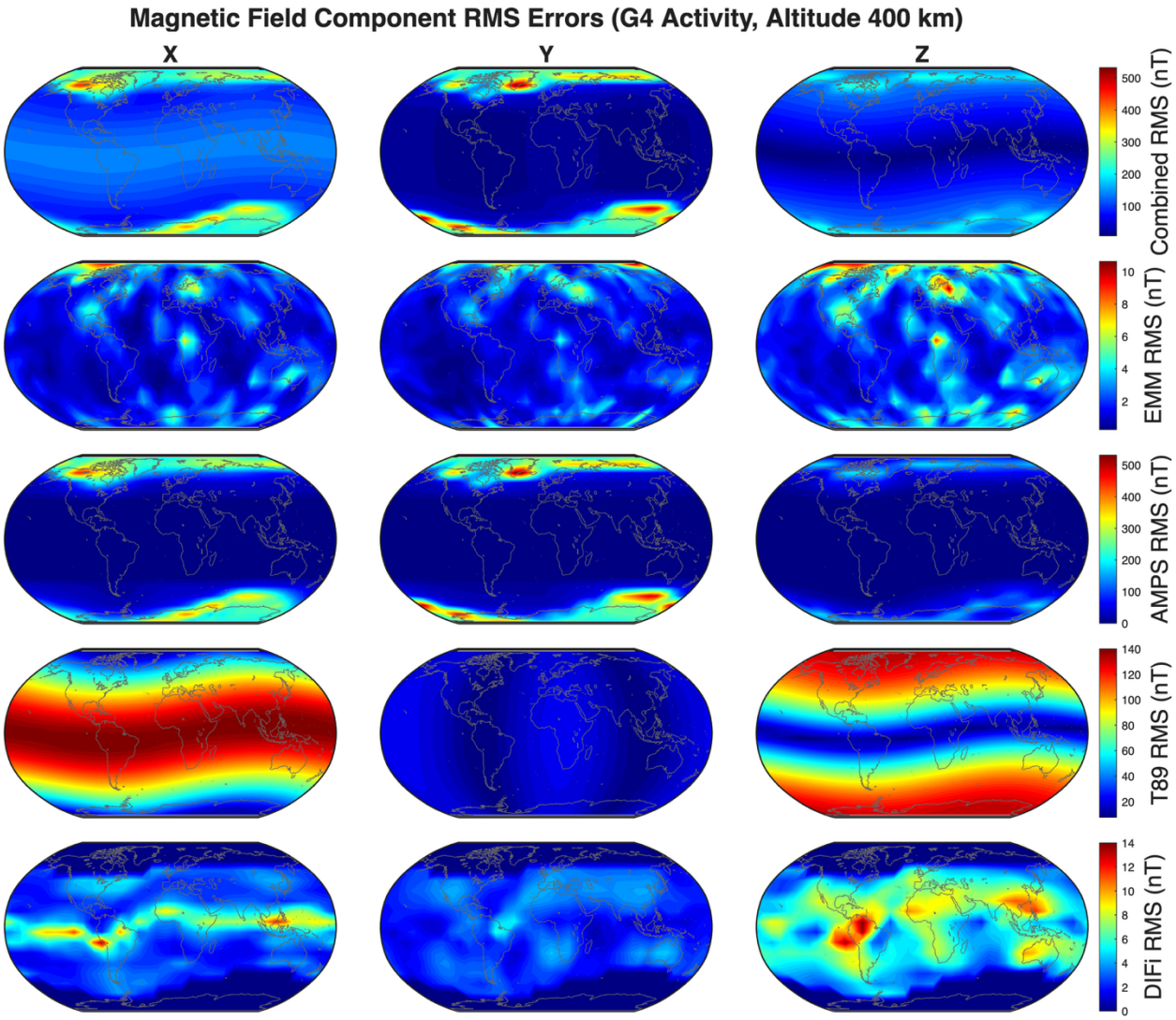

Figure 1b. Maps of the individual models' errors and their combined omission errors for G4 at 400 km altitude.

Figure 1 presents the spatial distribution of estimated WMM omission errors, contrasting quiet geomagnetic conditions (G0) at low altitude (10 km, Figure 1a) with severe storm conditions (G4) at Low Earth Orbit (LEO) altitude (400 km, Figure 1b). This comparison shows the total error, which is a combination of several error sources, as a function of altitude and geomagnetic activity, and it also reveals that the relative importance of the individual error sources varies under different conditions. In Figure 1a, the error distribution at 10 km under quiet conditions reveals significant contributions from both static crustal anomalies (EMM), evidenced by globally distributed patches of elevated error particularly over continental shields like the Eastern European Craton, and moderate external field variations, primarily confined to the equatorial (DIFI, T89) and high-latitude (AMPS, T89) regions. The combined RMS error is dominated by the geographically-fixed crustal sources. During severe storm conditions (G4) at 400 km altitude (Figure 1b), the relative importance of error sources shifts. The contribution from crustal fields (EMM) is substantially

diminished (from ~250 nT maximum RMS at 10 km to ~10 nT at 400 km) diminished due to the rapid decay of these potential fields with height, becoming a minor component of the total error despite the underlying anomalies remaining unchanged. Conversely, errors originating from external fields undergo significant amplification and spatial expansion due to increased geomagnetic activity. For instance, high-latitude ionospheric errors (AMPS) increase from a peak of ~40 nT under G0 conditions to over 500 nT during a G4 storm. Similarly, magnetospheric errors (T89) in the X component increase from ~45 nT to over 140 nT. High-latitude ionospheric errors (AMPS) intensify significantly, driven by strong storm-time auroral electrojets and field-aligned currents. Magnetospheric errors (T89) increase by more than an order of magnitude across most regions, especially in the X and Z components, reflecting the disturbed global current systems (e.g., ring current, magnetopause currents). In contrast, there is no strong enhancement in the DIFI model as it primarily characterizes low-latitude quiet-time ionospheric current systems such as the Sq and equatorial electrojet. These currents originate mainly from solar radiation and tidal drivers rather than direct solar wind forcing. Errors associated with this model decrease at orbital altitudes compared to near-surface levels (Figure 1a). This attenuation occurs largely because DIFI represents magnetic signals originating from the lower ionosphere (approximately 110 km altitude), and these signals naturally weaken with increasing distance from the source region (e.g., at 400 km).

The errors in the *X (North)*, *Y (East)*, and *Z (Vertical, positive down)* components are propagated to the *F (Total Field)* = $\sqrt{X^2 + Y^2 + Z^2}$, *H (Horizontal Intensity)*=$\sqrt{X^2 + Y^2}$, *D (Declination)*= $\arctan\left(\frac{Y}{X}\right)$ and *I (Inclination)* = $\arctan\left(\frac{Z}{H}\right)$ components using simplified relationships (e.g., Chulliat et al., 2020, WMM2020 Technical report). Specifically, this study uses the following formulas to propagate errors:

$$\delta H_{omission}(k,j,i) = \sqrt{\delta X_{omission}(k,j,i)^2 \cos^2 D_{WMM}(k,j,i) + \delta Y_{omission}(k,j,i)^2 \sin^2 D_{WMM}(k,j,i)} \quad (4)$$

$$\delta D_{omission}(k,j,i) = \frac{\sqrt{\delta X_{omission}(k,j,i)^2 \sin^2\left(D_{WMM}(k,j,i)\right) + \delta Y_{omission}(k,j,i)^2 \cos^2\left(D_{WMM}(k,j,i)\right)}}{H_{WMM}(k,j,i)} \quad (5)$$

$$\delta F_{omission}(k,j,i) = \sqrt{\delta H(k,j,i)^2 \cos^2\left(I_{WMM}(k,j,i)\right) + \delta Z_{omission}(k,j,i)^2 \sin^2\left(I_{WMM}(k,j,i)\right)} \quad (6)$$

$$\delta I_{omission}(k,j,i) = \frac{\sqrt{\delta H(k,j,i)^2 \sin^2\left(I_{WMM}(k,j,i)\right) + \delta Z_{omission}(k,j,i)^2 \cos^2\left(I_{WMM}(k,j,i)\right)}}{F_{WMM}(k,j,i)} \quad (7)$$

Where, F_{WMM}, D_{WMM}, I_{WMM} and H_{WMM} are the nominal WMM (epoch 2020, date of evaluation 2020-1-1) predictions at the center of each grid cell (h_k, ϕ_j, θ_i). Note that we chose a day at the start of the WMM2020 model so that the secular variation error is minimum. To mimic the total expected error of the WMM in an operational setting, we must account for both the omission errors due to unmodeled magnetic sources and the commission errors associated with the WMM itself. The commission errors primarily reflect uncertainties arising from secular variation (SV) forecasts,

data coverage gaps, and other model-specific limitations. Specifically, the commission errors are based on the maximum secular variation errors observed near the end of the WMM2015 epoch, ensuring a *conservative* estimate of the WMM's total error budget. These numbers are provided in row 3, Table 1. By combining the omission errors with these peak commission errors, we capture the full range of potential discrepancies that a user might encounter over the model's lifespan. Specifically, for any derived component Q, the total error δQ_{final} at a given grid cell (h_k, ϕ_j, θ_i) and for each G scale, is obtained by adding the omission error $\delta Q_{\text{omission}}$ and the commission error $\delta Q_{commission}$. Note that there is no altitude and location dependency for commission errors – they are determined as global average values for each component.

$$\delta Q_{final}(k,j,i) = \sqrt{\delta Q^2_{\text{omission}}(k,j,i) + \delta Q^2_{commission}} \tag{8}$$

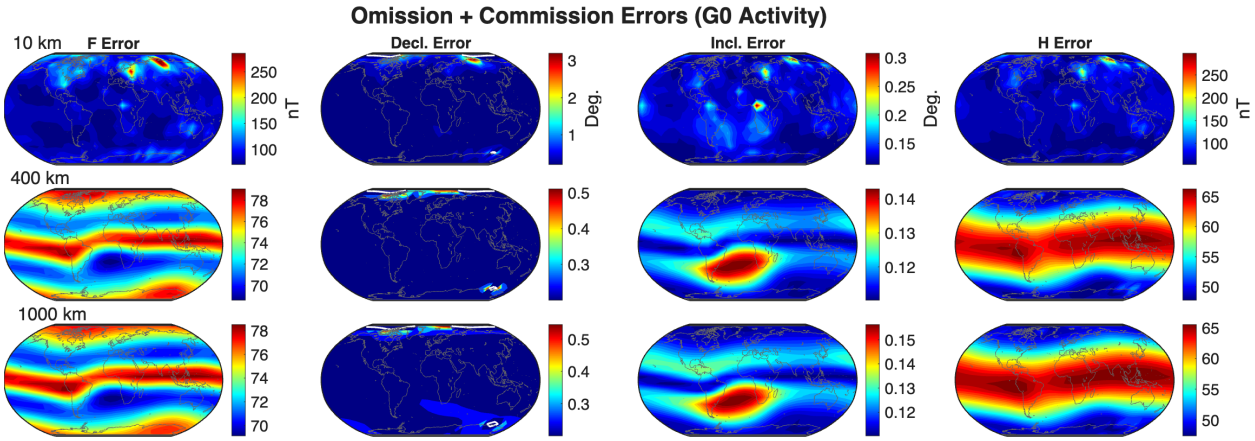

Figure 2a. Estimated final errors (δQ_{final}) of WMM from four models used in the study for G0 conditions and at 10, 400 and 1000 km altitudes.

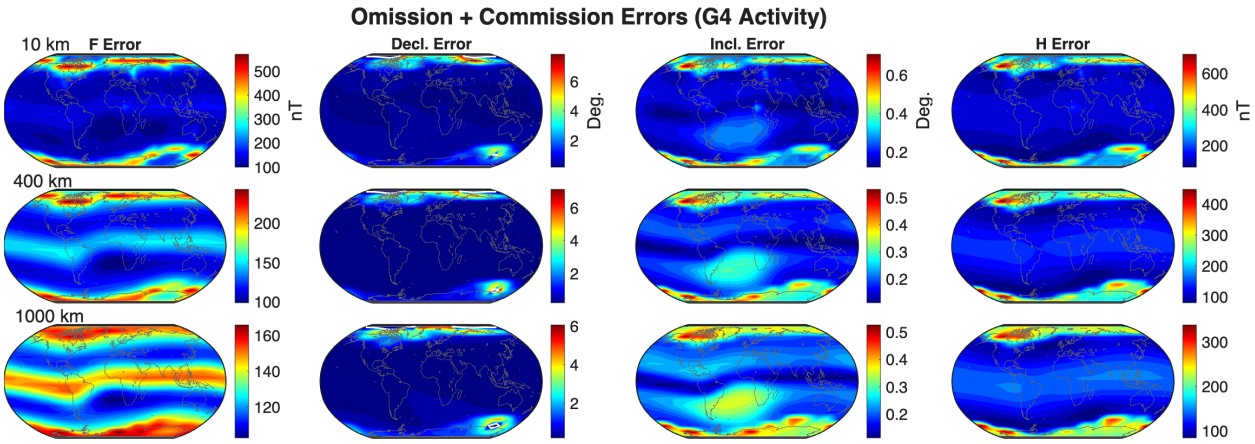

Figure 2b. Estimated final errors (δQ_{final}) of WMM from four models used in the study for G4 conditions and at 10, 400 and 1000 km altitudes.

Figure 2a displays the spatial distribution of the estimated final WMM errors (omission plus commission) for the Total Field (F), Declination (D), Inclination (I), and Horizontal Intensity (H) components under quiet geomagnetic conditions (G0) at 10 km, 400 km, and 1000 km altitudes. White contours on Declination maps indicate the Blackout Zone (BoZ), where the World Magnetic Model predicts horizontal intensity below 2000 nT; declination errors within these regions were excluded from the global error analysis. At 10 km, errors in F, I, and H are strongly modulated by crustal magnetic anomalies, with notable peaks like the elevated F and I errors over the Bangui anomaly region in central Africa. Declination errors at this altitude are characteristically largest in the polar regions due to geometric effects associated with weak horizontal fields near the poles. With increasing altitude to 400 km and 1000 km, the influence of crustal sources diminishes, leading to a reduction in the overall magnitude of F and H errors. However, broader patterns associated with quiet-time external fields (magnetospheric and ionospheric) become clearer and remain relatively consistent between 400 km and 1000 km, especially in the F, I and H components, reflecting the distant nature of magnetospheric sources and the stable structure of the quiet ionosphere at these heights. The polar maxima in D persist, albeit with slightly reduced magnitude compared to 10 km. In contrast, Figure 2b shows the error distribution during severe storm conditions (G4), revealing a significant increase in error magnitudes across all components and altitudes compared to the quiet G0 scenario. At 10 km, while the underlying crustal error contribution remains the same, it is largely overshadowed by significantly amplified errors from storm-enhanced ionospheric and magnetospheric currents, particularly evident in the high-latitude regions for all components. At LEO altitudes (400 km and 1000 km), the error is overwhelmingly dominated by intense and spatially extensive patterns driven by geomagnetic storms. Features prominent during quiet times, such as the specific signature of the Bangui anomaly in F and I, become masked by these large-scale storm-induced errors. The Declination errors in the polar regions are substantially larger than in G0, indicating that intense activity-related currents exacerbate the baseline geometric effect. Although F and H errors still show some decrease with altitude during the G4 conditions, the overall error levels remain extremely high and are governed by the intense external field contributions, which show similar broad spatial structures between 400 km and 1000 km.

5. Altitude limit maps

To determine the maximum altitudes at which the World Magnetic Model (WMM) errors remain within the specified Military Specification (MilSpec) and WMM Error Model thresholds, the study employed a systematic approach based on global error grids. The errors in the magnetic field components were analyzed as a function of altitude, and relative errors were calculated to account for the decreasing strength of the main magnetic field with increasing altitude. The relative error (error as a percentage of the main field) provides a measure of model accuracy at different altitudes. For the total magnetic field strength (F), the maximum altitude k at which the WMM errors remain within the MilSpec threshold was determined using the following condition:

$$\frac{\delta F(i,j,k)}{F(k)_{\mathrm{WMM}}} \leq \frac{\delta F_{\mathrm{MilSpec}}}{F(\mathrm{WGS84})_{\mathrm{WMM}}} \quad \text{and} \quad \frac{\delta F(i,j,k+1)}{F(k+1)_{WMM}} > \frac{\delta F_{MilSpec}}{F(WGS84)_{WMM}} \tag{9}$$

Here, $\delta \mathrm{F}(\mathrm{I},\mathrm{j},\mathrm{k})$ presents the total error (omission + commission, δQ_{final}) in the total magnetic field strength at a specific grid cell (i,j) and at altitude level k while $\mathrm{F}(\mathrm{k})_{\mathrm{WMM}}$ is the average value of

the WMM-predicted total magnetic field strength over the altitude shell at level k. The term $\delta F_{MilSpec}$ denotes the MilSpec error limit for the total magnetic field strength (Table 1), and $F(WGS84)_{WMM}$ is the average value of the WMM-predicted total magnetic field strength over the WGS84 ellipsoid (ground level). The process was repeated for all the components and for all the geomagnetic activity levels. Similar analyses were carried out to determine the upper altitude limits of the WMM Error Model. The global averages were calculated by weighing the quantities by the cosine of the latitude to account for varying areas of uniformly distributed grid cells with latitude and longitude across the Earth's surface. For the angular components of the magnetic field, such as Declination (D) and Inclination (I), the errors do not decrease with altitude because these components are not directly influenced by the weakening of the main magnetic field. Therefore, the maximum altitude k for these components was determined using a simpler condition:

$$\delta D(i,j,k) < \delta D_{MilSpec} \quad \text{and} \quad \delta D(i,j,k+1) > \delta D_{MilSpec} \tag{10}$$

Here, $\delta D(I,j,k)$represents the error in the Declination component at a specific grid cell (I,j) and altitude k and $\delta D_{MilSpec}$ is the MilSpec error limit for Declination (Table 1). A similar approach was used for the Inclination component and other angular components**.**

If the calculated error for a component never exceeds the specified threshold at any simulated altitude, we consider the model accurate for that component throughout the entire altitude range (up to 10,000 km). In certain locations, the error might exceed the threshold near the ground due to strong, localized crustal anomalies (like the Bangui anomaly in Central Africa). As altitude increases, the influence of such anomalies weakens, and the relative error may fall back within acceptable limits. However, at even higher altitudes, errors driven by external fields can increase, potentially causing the component to exceed the threshold again. In cases where a component's error exceeds the threshold at a low altitude but then falls back within limits before failing again at a higher altitude, the map displays the highest altitude reached before the final failure. The study generated maximum-altitude maps that indicate the altitudes up to which the WMM errors remain within the MilSpec and WMM Error Model thresholds. These maps were created by analyzing the global error grids and applying the conditions described above for each magnetic field component and for each geomagnetic condition. The results provide insights into the altitude-dependent performance of the WMM. The maps presented show the calculated upper altitude limit for each component, representing the maximum altitude at which the model's error remains below the specified threshold

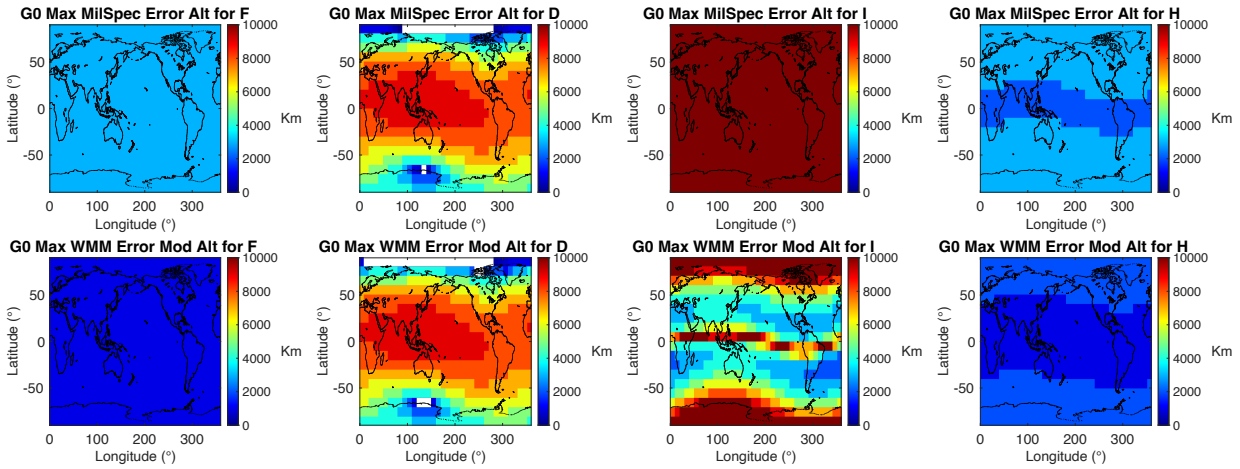

Figure 3a. Maximum altitude (km) up to which estimated World Magnetic Model (WMM) errors remain below specified thresholds for components F, D, I, and H during quiet (G0) geomagnetic conditions. Top row displays results using MilSpec thresholds; bottom row uses WMM Error Model thresholds. Color indicates the maximum altitude (up to 10,000 km), while white areas indicate regions where the threshold is not met within this altitude range.

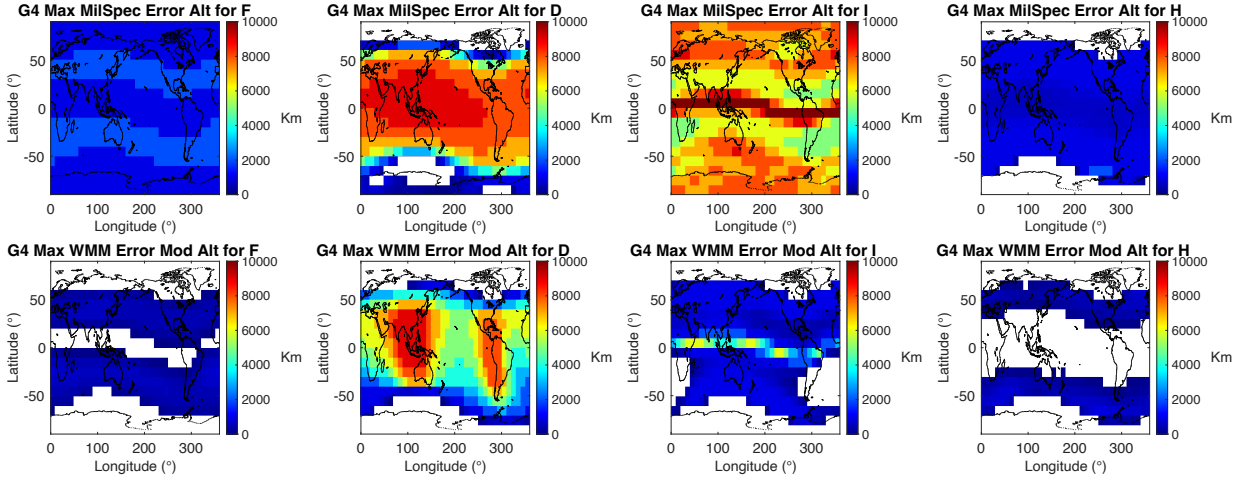

Figure 3b. Maximum altitude (km) up to which estimated World Magnetic Model (WMM) errors remain below specified thresholds for components F, D, I, and H during severe storm (G4) geomagnetic conditions. Top row displays results using MilSpec thresholds; bottom row uses WMM Error Model thresholds. Color indicates the maximum altitude (up to 10,000 km), while white areas indicate regions where the threshold is not met within this altitude range.

Figure 3 presents maps indicating the maximum altitude up to which the WMM meets specified error thresholds—the operational Military Specification (MilSpec, top row) and the more stringent WMM Error Model (bottom row)—for the F, D, I, and H components. Figure 3a illustrates these limits for quiet G0 geomagnetic conditions. Under the MilSpec threshold (Fig. 3a, top row), WMM performance is generally robust: the Inclination (I) component remains valid throughout the entire 10,000 km altitude range studied across the entire globe. The Total Field (F) meets condition up to 3,000 km. Declination (D) meets the specification up to high altitudes (8,000 km+) over most mid- and low-latitudes, but validity is significantly reduced near the poles, primarily due to

geometric effects and they are invalid for all altitudes near the geomagnetic poles. The Horizontal intensity (H) typically meets the MilSpec threshold up to intermediate altitudes, generally around 3,000-4,000 km. Applying the stricter WMM Error Model threshold (Fig. 3a, bottom row) substantially lowers these validity ceilings for G0: F validity drops globally to approximately 2,000 km, while H validity is highest near the poles (~2,000-3,000 km) but falls below 1,000 km in a band around the dip equator. For the angular components under this stricter threshold, D shows larger regions of reduced validity near the poles, while I remains valid to high altitudes (8,000 km+) near the poles and the dip equator but is limited to intermediate altitudes (~4,000-6,000 km) in the mid-latitudes.

In contrast, Figure 3b shows the corresponding maximum altitude maps for severe G4 storm conditions, revealing a reduction in the operational ceiling of the WMM compared to quiet times shown in Figure 3a. Under the MilSpec threshold (Fig. 3b, top row), the F validity is reduced to 1,000-2,000 km. The H validity is reduced to 700 – 2000 km, with noticeable gaps in the northern and southern polar regions. Under the WMM Error Model threshold during G4 (Fig. 3b, bottom row), large portions of the globe fail to meet the criteria even at low altitudes (indicated by white areas), particularly for the F and H components; H fails almost globally. While D and I retain some validity, the maximum altitudes are significantly lowered compared to G0; for I, validity up to 10,000 km is lost, and large areas, especially in the northern hemisphere, fail entirely. For D, the regions of failure expand far beyond the typical polar geometric limitations seen in G0, covering extensive high- and mid-latitude zones, clearly driven by geomagnetic activity.

6. Altitude profiles

To complement the spatial error maps and provide a global overview of how WMM errors evolve with altitude and geomagnetic activity, global altitude error profiles were computed for each magnetic field component (X, Y, Z, H, F, D and I) and for each geomagnetic activity level (G0-G4). These profiles illustrate the average error magnitude as a function of altitude from the Earth's surface up to 10,000 km. The altitude error profiles are derived from the global grids of final WMM errors, $\delta Q_{\text{final}}(h_k, \phi_j, \theta_i)$, calculated as described in Section 4. For each component Q, altitude level h_k, and G scale level, a single global average RMS error value, denoted as $\overline{\delta Q_{final}(h_k)}$, is computed. For most components (specifically X, Y, Z, H, F, D, I), this global average is calculated by averaging the RMS error values across all grid cells (i,j) at altitude h_k , weighted by the cosine of the latitude ϕ_j :

$$\overline{\delta Q_{\text{final}}}(h_k) = \frac{\sum_{i,j} \delta Q_{\text{final}}(h_k, \phi_j, \theta_i) \cos(\phi_j)}{\sum_{i,j} \cos(\phi_j)} \quad (11)$$

A distinction is made between intensity components (F, H, X, Y, Z) and angular components (D, I, GVN (Grid Variation North), GVS (Grid Variation South)) due to their different behavior with altitude relative to the main field strength.

- **Intensity Components:** The absolute errors for these components generally decrease with altitude, primarily due to the diminishing contribution of the crustal magnetic field. However, the WMM-predicted main field strength for these components also decreases with altitude. To provide a more informative measure of model accuracy relative to the

field being modeled, **relative error profiles** are generated. The relative error at altitude h_k, $\delta\mathrm{Qrel}(h_k)$, is calculated by dividing the globally averaged absolute RMS error $\delta\mathrm{Qfinal}(\mathrm{h_k})$ by the globally averaged WMM-predicted field strength $\mathrm{Q_{WMM}}(h_k)$at the same altitude, expressed as a percentage: $\delta Q_{rel}(h_k) = \frac{\overline{\delta Q_{final}}(h_k)}{\overline{Q_{WMM}}(h_k)} \times 100\%$

where $\overline{\mathrm{Q_{WMM}}}(\mathrm{h_k})$ is computed similarly to the average error using cosine latitude weighting. On these relative error plots, the MilSpec and WMM Error Model thresholds ($\delta\mathrm{Q_{Thresh}}$)are represented as vertical lines. The position of these lines corresponds to the ratio of the ground-level threshold value to the globally averaged WMM field strength at the ground level ($h_k = 0$), expressed as a percentage: Threshold Position (%) = $\frac{\delta Q_{Thresh}}{\overline{Q_{WMM}}(h_k=0)} \times 100\%$

- **Angular Components:** The magnitudes of errors in angular components do not inherently decrease significantly with altitude in the same way as intensity components when considered in absolute angular units (degrees). Therefore, **absolute error profiles** are presented for these components, plotting the globally averaged RMS error $\delta Qfinal(h_k)$ (in degrees) directly against altitude. Relative error calculations are not necessary for these components.
- **Grid Variation Components:** In polar regions, where magnetic declination changes significantly with longitude, the WMM specification ((MIL-PRF-89500B) defines auxiliary angles called Grid Variation North (GVN) and Grid Variation South (GVS). GVN and GVS are subsets of grid variation (or grivation), which represents the angle from grid north to magnetic north. GVN = D - λ for latitudes (φ) greater than 55°, and GVS = D + λ for latitudes (φ) less than -55°, where D is magnetic declination and λ is longitude. For the Grid Variation components, the global averages were calculated without cosine latitude weighting because they are exclusively polar components designed to provide a representative regional average without suppressing the areas near the pole.

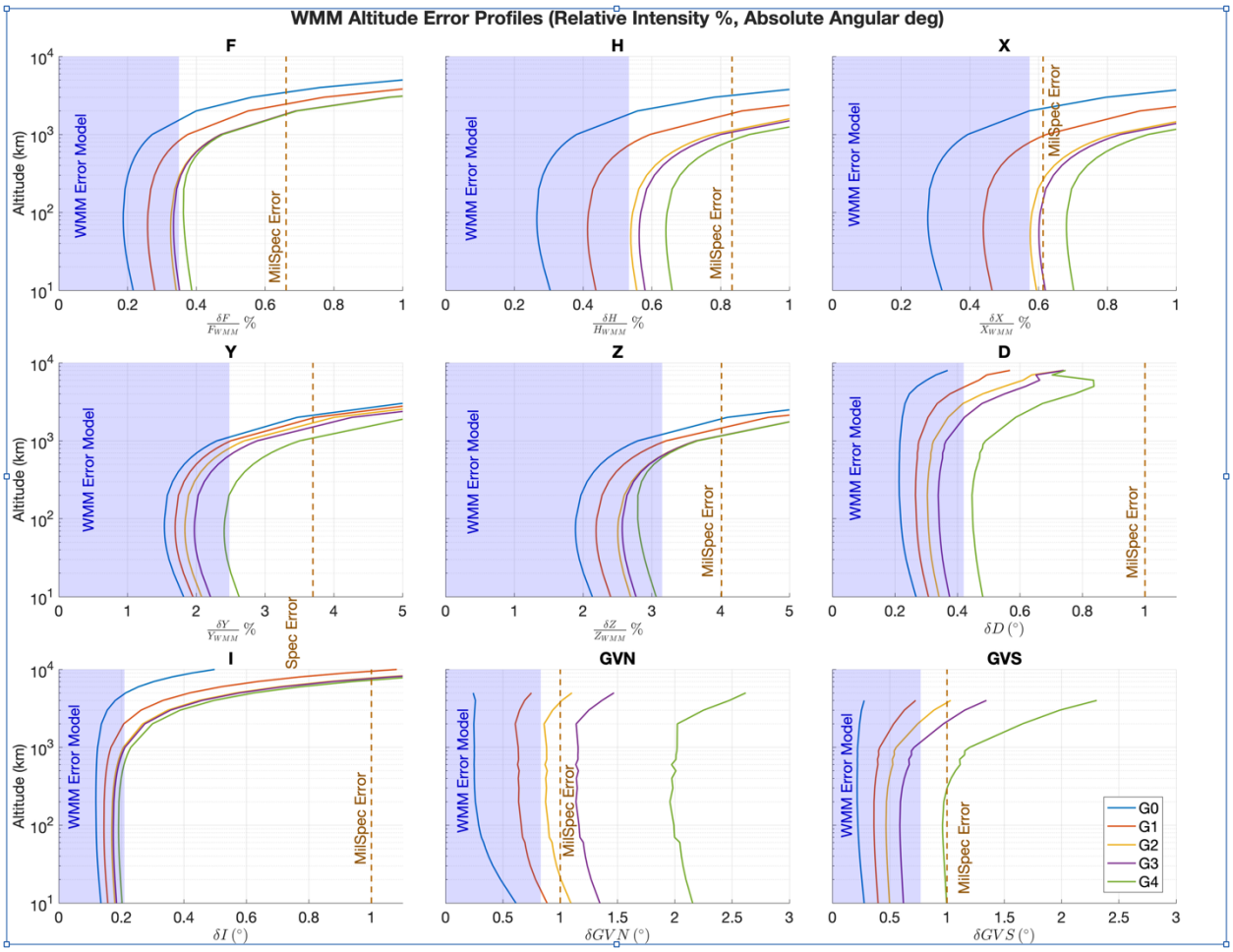

Figure 4. Altitude profiles of globally averaged WMM errors for different geomagnetic activity levels (G0-G4, indicated by colored lines). Intensity components (F, H, X, Y, Z) are plotted as relative error $\frac{\delta Q}{Q_{WMM}}$ %. Angular components (D, I, GVN, GVS) are plotted as absolute error (δQ ∘)). The dashed orange line represents the MilSpec error threshold, and the right boundary of the shaded blue region represents the WMM Error Model threshold.

The global average WMM error profiles, plotted as a function of altitude for different geomagnetic activity levels (G0-G4), reveal distinct behaviors for intensity and angular components (Figure 4). For intensity components (F, H, X, Y, Z), relative errors (plotted as $\frac{\delta Q}{Q_{WMM}}$ %) typically start low near the surface, decrease to a minimum at altitudes of a few hundred kilometers, often reach a minimum around 400-500 km as crustal influences wane faster than the main field, and then increase again at higher altitudes due to the growing relative influence of external fields. This results is generally consistent with the observations from Matteo and Morton (2011) where they find the RMS misfit (CHAMP – IGRF) relative to the IGRF predictions is highest at the lowest altitudes (~0.25% at 300 km) and steadily decreases with height for the entire range of the CHAMP data, reaching its minimum value of less than 0.1% at the highest altitude shown, around 460 km. Geomagnetic activity profoundly impacts these errors; profiles shift significantly towards higher error percentages as activity increases from G0 to G4. This directly translates to a reduced

operational altitude range, as shown in the maximum altitude data. For example, under the MilSpec threshold, the maximum valid altitude for F drops from 3500 km during G0 to 1900 km during G3/G4, while H drops from 3200 km to 800 km. The X component appears particularly sensitive, exceeding the MilSpec threshold above G1 activity. The stricter WMM Error Model thresholds are met only up to much lower altitudes, with most intensity components exceeding these limits during G2 or G3 activity levels within the first ~1000-2000 km.

The absolute errors for angular components (D, I, GVN, GVS), plotted in degrees, generally show an increasing trend with altitude, particularly during active conditions (Figure 4). Inclination (I) errors remain below the 1° MilSpec threshold up to very high altitudes even during G4 conditions (7300 km), although this is a reduction from the 10,000+ km range during G0. However, under the stricter WMM Error Model threshold (0.21°), the maximum altitude for I drops rapidly from 4900 km (G0) and fails entirely during G4 conditions. The plots show that D errors generally remain below 1° (MilSpec) across most altitudes and activity levels shown, but cross the 0.42° (WMM EM) threshold at progressively lower altitudes as activity increases. GVN and GVS errors also clearly increase with activity, crossing both thresholds significantly sooner during G3/G4 compared to G0/G1, particularly below their respective altitude cutoffs (5000/4000 km). Overall, both the profiles and the threshold data demonstrate that increasing geomagnetic activity severely restricts the altitude range where WMM component errors remain within MilSpec limits and the expected ranges defined by the WMM Error Model .

Component	G0 altitudes (km)	G1 altitudes (km)	G2 altitudes (km)	G3 altitudes (km)	G4 altitudes (km)
F (280 nT)	0 to 3500	0 to 2500	0 to 1900	0 to 1900	0 to 1900
H (200 nT)	0 to 3200	0 to 1900	0 to 1200	0 to 1100	0 to 800
X (140 nT)	0 to 2200	0 to 1000	Exceeds	Exceeds	Exceeds
Y (140 nT)	0 to 2200	0 to 2000	0 to 1800	0 to 1600	0 to 1100
Z (200 nT)	0 to 1900	0 to 1500	0 to 1200	0 to 1200	0 to 1200
D (1 deg)	0 to 6000+	0 to 6000+	0 to 6000+	0 to 6000+	0 to 6000+
I (1 deg)	0 to 10000+	0 to 9500	0 to 7800	0 to 7700	0 to 7300
GVN (1 deg)	0 to 5000+	0 to 5000+	0 to 3800	Exceeds	Exceeds
GVS (1 deg)	0 to 4000+	0 to 4000+	0 to 3800	0 to 2200	0 to 300

Table 4. Maximum altitude (km) up to which the globally averaged WMM error for each component remains below the **MilSpec threshold** for different geomagnetic activity levels (G0-G4). Altitudes are estimated by interpolating the global average error profiles (Figure 4) and rounded to the nearest 100 km. 'Exceeds' indicates the threshold is surpassed at or near the lowest altitude considered (10 km). A '+' indicates validity extends beyond the maximum modeled altitude of 10,000 km.

Component	G0 altitudes (km)	G1 altitudes (km)	G2 altitudes (km)	G3 altitudes (km)	G4 altitudes (km)
F (148 nT)	0 to 1600	0 to 800	Exceeds	Exceeds	Exceeds
H (128 nT)	0 to 1900	0 to 700	Exceeds	Exceeds	Exceeds
X (131 nT)	0 to 2000	0 to 800	Exceeds	Exceeds	Exceeds
Y (94 nT)	0 to 1200	0 to 1000	0 to 800	Exceeds	Exceeds
Z (157 nT)	0 to 1300	0 to 1000	0 to 700	0 to 600	Exceeds
D (0.42 deg)	0 to 6000+	0 to 4900	0 to 3000	0 to 2000	Exceeds
I (0.21 deg)	0 to 4900	0 to 2000	0 to 1100	0 to 1000	Exceeds
GVN (0.83 deg)	0 to 5000+	20 to 5,000+	Exceeds	Exceeds	Exceeds
GVS (0.77 deg)	0 to 4000+	0 to 4000+	0 to 2200	0 to 1200	Exceeds

Table 5. Maximum altitude (km) up to which the globally averaged WMM error for each component remains below the **WMM Error Model value** for different geomagnetic activity levels (G0-G4). Altitudes are estimated by interpolating the global average error profiles (Figure 4) and rounded to the nearest 100 km. 'Exceeds' indicates the WMM Error Model value is surpassed at or near the lowest altitude considered (10 km). A '+' indicates validity extends beyond the maximum modeled altitude of 10,000 km

7. Comparison to satellite data

The models used to construct the altitude profiles introduced in sections 4-6 provide a way to estimate the error due to lithospheric, ionospheric, and magnetospheric fields in the WMM. They, however, have some limitations, as they do not account for all ionospheric and magnetospheric currents, and generally make simplifying assumptions and time-space variations of ionospheric and magnetospheric currents and fields. To validate the model-based findings from Section 3, we analyzed magnetic field measurements from four satellite missions: CHAMP, Swarm and CryoSat-2. The CHAMP, and Swarm satellites provide high-precision vector magnetic measurements calibrated with on-board independent scalar measurements (Olsen et al. 2001, 2013). CryoSat-2, on the other hand, is only equipped with a platform magnetometer originally designed for attitude control. To obtain the necessary accuracy for scientific use, CryoSat-2 data are calibrated using the CHAOS main field model (Olsen et al. 2020). All these satellites provide data at different altitudes with the appropriate precision to investigate the lithospheric, ionospheric and magnetospheric fields. Table 6 provides a summary of the satellite data used in this study. (Swarm C data were not used, as Swarm C orbits at the same altitude and close to Swarm A.).

7.1 Satellite Data Sources

Parameter	CHAMP	Swarm (A & B)	CryoSat-2
Launch Date	July 15, 2000	November 22, 2013	April 8, 2010
Launch Agency	German Aerospace Center (DLR)	European Space Agency	European Space Agency
Operational Status	Operations ended 2010	Currently active	Currently active
Study Period	2005-01-01 to 2009-12-31	2015-01-01 to 2024-12-31	2015-01-01 to 2024-12-31

Average Altitude	345 km (2005-2010)	450 km (A), 511 km (B)	727 km
Primary Instruments	Fluxgate Magnetometer, Overhauser Magnetometer	Vector Field Magnetometer, Absolute Scalar Magnetometer	Platform Magnetometers

Table 6. Summary of key parameters for the satellite missions (CHAMP, Swarm A & B and CryoSat-2) providing in-situ magnetic field measurements used for validation in this study. Details include operational periods, average altitudes, and relevant instrumentation.

7.2 Satellite Data Processing Methodology

Satellite data from different satellite missions were preprocessed independently. A core field model was first subtracted from the data; we used the WMM2015v1 and WMM2020 for Swarm and CryoSat-2 data, for periods 2015-2019 and 2020-2024 respectively; WMM2005 for CHAMP (2005-2009). Next, the data were divided into separate groups corresponding to the geomagnetic activity levels G0, G1, G2, G3 and G4 (G5 is excluded here due to insufficient data available for this activity level). We identified and removed outliers by performing a generalized extreme Studentized deviate test for outliers (Rosner, 1983) for each activity level group, and separately for data above and below 55 degrees quasi-dipole latitude. This allows us to robustly account for the differences in amplitude of external fields as a function of geomagnetic activity and latitude, which would most likely lower the reliability of the test if all data were considered together. The RMSs were then computed for each activity level using data weighted by the cosine of latitude, again to account for the non-uniform spatial distribution of measurements.

Finally, secular variation (SV) forecasting errors inherent in the WMM residuals were inferred and accounted for by leveraging the high temporal resolution of the CHAOS-8.2 model. As CHAOS-8.2 incorporates recent satellite and observatory data and models time dependence using B-splines with 6-month knot spacing (Finlay et al., 2020), it provides a high-fidelity estimate of the recent main field evolution against which the WMM forecast can be compared. To obtain the inferred secular variation, we subtracted the CHAOS residuals (Measurements - CHAOS) from the WMM residuals (Measurements - WMM) in quadrature. We then removed this inferred secular variation from the WMM residuals to obtain satellite-derived omission errors. As a final setp, to create a total error estimate from the satellite data that is directly comparable to our modeled total error, this satellite-derived omission error was then combined in quadrature with the same peak global commission error values (Table 1, row 3) used in the model-based calculations.

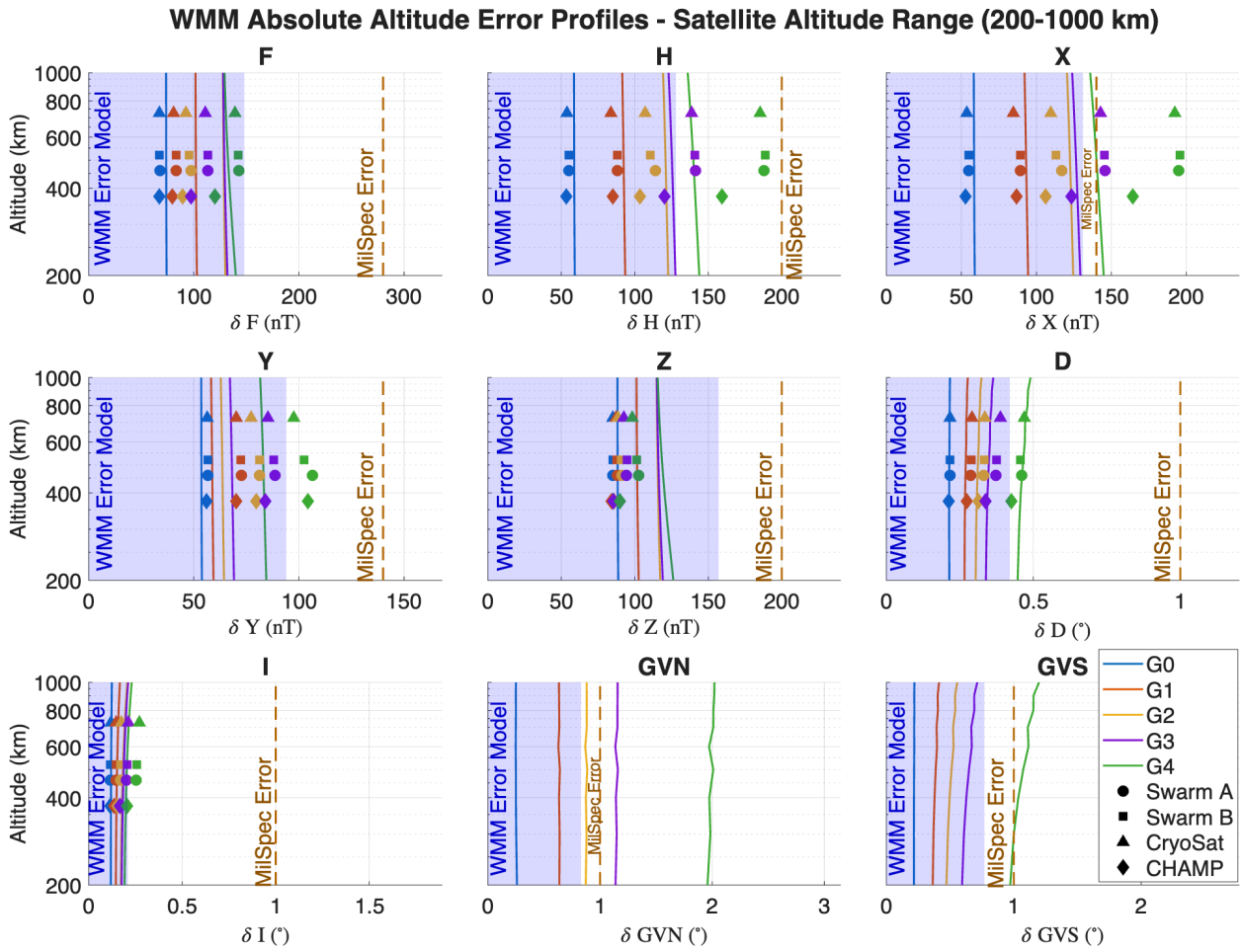

Figure 5. Comparison of modeled globally averaged absolute WMM error profiles (lines representing G0-G4 activity levels) with processed satellite error estimates (markers) for components F, H, X, Y, Z, D, I, GVN, and GVS. Satellite data points are shown at approximate average mission altitudes: CHAMP (diamonds), Swarm (A circles, B squares, and CryoSat-2 (triangles). Marker colors correspond to the geomagnetic activity level (G0-G4) of the data point. Dashed orange lines indicate MilSpec thresholds, and the boundary of the shaded blue region indicates WMM Error Model thresholds. Satellite error estimates were not made for GVN and GVS.

Figure 5 presents the modeled globally averaged absolute RMS error profiles for WMM components F, H, X, Y, Z, D, and processed error estimates from satellite observations. The satellite data includes measurements from CHAMP (diamonds; 2005-2010 altitude ~345 km, stars), Swarm (A, 450 km altitude, circles; B, 511 km altitude, squares), and CryoSat-2 (727 km, triangles, plotted at their respective altitudes. In general, satellite-derived errors show reasonable agreement with the model profiles near their orbital altitudes for most components. Specifically for F, H, X, Y, D, and I, the satellite data points typically are close to the model predictions, and consistently exhibit a clear increase in error magnitude and spread as geomagnetic activity increases from G0 (blue) to G4 (yellow/red), broadly validating the modeled dependence on disturbance levels. Notably, some G4 error estimates for X and H components lie outside of the average G4 model profile. These components are maximally sensitive to the external fields and the models may be underestimating the errors for these components for G4 situations.

A discrepancy between the models and satellite observations is apparent, however, for the Z component. While the model profiles predict a substantial increase in the globally averaged δZ RMS error with increasing geomagnetic activity, the satellite data points from all three missions (CHAMP, Swarm and CryoSat-2) show considerably less variation across the G0-G4 levels and cluster much closer to the modeled G0 profile. This suggests that the globally averaged RMS of the Z component variations, as sampled by the satellites along their orbits, is less sensitive to geomagnetic activity than predicted by the combination of external field models (T89 and AMPS) used in this study. It is possible that the combined external field models, particularly the empirical magnetospheric model (T89), overestimate the magnitude or spatial coherence of the Z component perturbation when averaged globally via RMS, leading to the observed difference. Despite this issue with the Z component, the overall agreement for the other components provides confidence in the model's ability to capture the general trends of WMM error with altitude and geomagnetic activity.

8. Discussion and conclusions

Our findings demonstrate a clear dependence of WMM errors on altitude. The errors generally decrease with altitude from the ground up to LEO satellite altitudes, primarily because the contribution from crustal magnetic fields diminishes rapidly with distance from the Earth's surface. This is consistent with previous studies that have estimated errors at ground level and LEO satellite altitudes. For instance, Beggan (2022) estimated ground-level uncertainties for the IGRF using ground-based measurements and found standard deviations of 144 nT for the North (X) component, 136 nT for the East (Y) component, and 293 nT for the vertical (Z) component. These values are substantially larger than typical RMS errors observed at LEO altitudes by studies with satellite data, such as Cilden-Guler et al. (2018) who found RMS errors for Swarm A on a quiet day (18 January 2014, X 25.8 nT, Y 7.1 nT, Z 18.6 nT, F 9.1 nT). This is within the range of 10-30 nT errors at LEO altitudes found in this study. We identified an altitude of approximately 400 km where relative errors are minimized for most components, representing a potential optimal altitude range in terms of relative WMM accuracy for LEO missions during quiet times.

The study also confirms and quantifies the significant impact of geomagnetic activity on WMM errors, which increases with increasing activity. This finding is consistent with previous work, such as Matteo and Morton (2011), who observed that deviations between IGRF and satellite measurements increased during active conditions (Kp > 5). Similarly, Cilden-Guler et al. (2018) found that model differences from observations were significantly larger during active geomagnetic periods for both IGRF and T89 models, particularly during geomagnetic storms. Our systematic analysis across a range of activity levels (G0-G4) demonstrates that while the altitude-dependent error reduction persists during active times, the rate is more moderate, and the model's validity range is notably reduced. We find that the X, F, and H components are more sensitive to these external magnetic fields, which aligns with the expectation that these components are strongly influenced by disturbance currents in the magnetosphere and ionosphere.

The validation of our model-based findings using in-situ magnetic field data from CHAMP, Swarm and CryoSat-2 satellites provides empirical support for the observed altitude-dependent error trends near their orbital altitudes. These findings extend previous satellite-based validation efforts that might have focused primarily on scalar field intensity or limited altitude ranges. The

issue noted with the Z component perturbation possibly being overestimated by the model when globally averaged via RMS warrants further investigation and may reflect complexities in accurately modeling vertical external field components.

The results of this study have important implications for various space applications. For satellite attitude determination systems at LEO, which commonly utilize geomagnetic models like IGRF or WMM, understanding the altitude and activity dependence of model errors is critical. While the angle between model predictions and observations may remain acceptably small (e.g., less than 1 degree as noted by Cilden-Guler et al., 2018) during quiet periods, the significant increase in error magnitude during geomagnetic storms can compromise attitude accuracy. Our results provide quantitative data to inform error budgets for attitude determination systems at different altitudes and activity levels, enabling more robust system design and operational planning. They also underscore the value of employing geomagnetic models that incorporate external field effects, or developing mitigation strategies during active times if precise attitude knowledge is required.

Furthermore, our findings are relevant for magnetic navigation (MagNav) systems and GNSS-based applications operating at altitude. The accuracy of WMM predictions directly impacts navigation performance. The systematic characterization of WMM errors across altitudes and activity levels allows for better assessment of navigation accuracy limitations and the potential implementation of error correction techniques. For applications modeling ionospheric effects on radio signals, where the magnetic field vector is a fundamental parameter, understanding WMM error characteristics helps quantify the uncertainty introduced into these models.

While this study provides a comprehensive analysis, limitations exist, such as the complexity of accurately modeling all ionospheric and magnetospheric currents and their variations. A key aspect of our methodology is the use of empirical and climatological models that represent the *average* magnetic fields for a given level of geomagnetic activity. The results, therefore, reflect a statistical expectation of WMM error and not a forecast for a specific, dynamic storm event. The instantaneous errors experienced by a user during a particular storm could vary significantly from these averaged values. Future work could involve more extensive statistical validation using a wider range of satellite data across different orbits and environments, evaluating other global geomagnetic models, and refining error models to better capture the dynamic nature of external field contributions.

9. Acknowledgements

This research was supported by NOAA cooperative agreement NA22OAR4320151. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of NOAA or the U.S. Department of Commerce.

10. Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

11. Open Research

All models and satellite datasets used in this study are publicly available from the sources listed below. The final modeled-error grids and the satellite-based errors are available from the Zenodo database (Nair et al., 2025) .

- **World Magnetic Model (WMM) and Enhanced Magnetic Model (EMM):** The WMM and EMM coefficients, software, and documentation are available from the NOAA National Centers for Environmental Information (NCEI) website.
  - WMM: https://www.ncei.noaa.gov/products/world-magnetic-model
  - EMM: https://www.ncei.noaa.gov/products/enhanced-magnetic-model
- **Dedicated Ionospheric Field Inversion (DIFI):** The DIFI-7 model used in this study is available from the University of Colorado.
  - DIFI: https://geomag.colorado.edu/difi-7
- **Average Magnetic field and Polar current System (AMPS):** The AMPS model is available at https://klaundal.w.uib.no/.
- **Tsyganenko 1989 (T89) Model:** The T89 model is a standard magnetospheric field model available from public resources, including the NASA Goddard Space Flight Center Community Coordinated Modeling Center (CCMC, https://ccmc.gsfc.nasa.gov/).
- **Satellite Data:**
  - **Swarm and CryoSat-2:** Data from the European Space Agency (ESA) Swarm and CryoSat-2 missions are available through the ESA Earth Online portal (https://earth.esa.int/).
  - **CHAMP:** Data from the CHAMP mission is available at https://isdc.gfz.de/champ-isdc/access-to-the-champ-data/
- **Geomagnetic Indices:** The Kp index data used in this study are provided by the GFZ German Research Centre for Geosciences.
  - Kp Index: https://doi.org/10.5880/Kp.0001
- NOAA's Space-Weather Geomagnetic Scale is available at (https://www.swpc.noaa.gov/noaa-scales-explanation)

12. References

Alken, P., Thébault, E., Beggan, C. D., et al. (2021). International Geomagnetic Reference Field: The thirteenth generation. *Earth, Planets and Space*, *73*, 49. https://doi.org/10.1186/s40623-020-01288-x

Alken, P., Maus, S., Chulliat, A., & Manoj, C. (2015). NOAA/NGDC candidate models for the 12th generation International Geomagnetic Reference Field. *Earth, Planets and Space*, *67*, 68. https://doi.org/10.1186/s40623-015-0215-1

Bartels, J., Heck, N. H., & Johnston, H. F. (1939). The three-hour-range index measuring geomagnetic activity. *Terrestrial Magnetism and Atmospheric Electricity*, *44*(4), 411–454.

Beggan, C. D. (2022). Evidence-based uncertainty estimates for the International Geomagnetic Reference Field. *Earth, Planets and Space*, *74*, 17. https://doi.org/10.1186/s40623-022-01572-y

Canciani, A., & Raquet, J. (2016). Absolute positioning using the Earth's magnetic anomaly field. *Journal of the Institute of Navigation*, *63*(1), 111–126. https://doi.org/10.1002/navi.138

Chulliat, A., Brown, W., Alken, P., Beggan, C., Nair, M., Cox, G., Woods, A., Macmillan, S., Meyer, B., & Paniccia, M. (2020). *The US/UK World Magnetic Model for 2020-2025: Technical Report*. National Centers for Environmental Information, NOAA. https://doi.org/10.25923/ytk1-yx35

Chulliat, A., Vigneron, P., & Hulot, G. (2016). First results from the Swarm dedicated ionospheric field inversion chain. *Earth, Planets and Space*, *68*(1), 104. https://doi.org/10.1186/s40623-016-0481-6

Chulliat, A., Brown, W., Nair, M., Gomez Perez, N., Young, L.-Y., Watson, C., Boneh, N., Beggan, C., Meyer, B., & Paniccia, M. (2025). *The US/UK World Magnetic Model for 2025–2030: Technical report.* National Centers for Environmental Information, NOAA. https://doi.org/10.25923/prbc-s316

Chulliat, A., Macmillan, S., Alken, P., Beggan, C., Nair, M., Hamilton, B., Woods, A., Ridley, V., Maus, S., & Thomson, A. (2015). *The US/UK World Magnetic Model for 2015–2020: Technical Report*. National Geophysical Data Center, NOAA. https://doi.org/10.7289/V5TB14V7

Cilden-Guler, D., Kaymaz, Z., & Hajiyev, C. (2018). Evaluation of geomagnetic field models using magnetometer measurements for satellite attitude determination at low Earth orbits: Case studies. *Advances in Space Research*, *61*(1), 513–529. https://doi.org/10.1016/j.asr.2017.10.041

Department of Defense. (2019). *Performance specification—World Magnetic Model (WMM)* (Report No. MIL-PRF-89500B). National Geospatial-Intelligence Agency. https://www.ngdc.noaa.gov/geomag/WMM/data/MIL-PRF-89500B.pdf

Finlay, C. C., Kloss, C., Olsen, N., Hammer, M. D., Tøffner-Clausen, L., Gray, J., & Kuvshinov, A. (2020). The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly. *Earth, Planets and Space*, *72*, 156. https://doi.org/10.1186/s40623-020-01252-9

Lowes, F. J. (2000). An estimate of the errors of the IGRF/DGRF fields 1945–2000. *Earth, Planets and Space*, *52*(12), 1207–1211. https://doi.org/10.1186/BF03352353

Matteo, N. A., & Morton, Y. T. (2011). Ionosphere geomagnetic field: Comparison of IGRF model prediction and satellite measurements 1991–2010. *Radio Science*, *46*, RS4003. https://doi.org/10.1029/2010RS004529

Matzka, J., Bronkalla, O., Tornow, K., Elger, K., & Stolle, C. (2021). *Geomagnetic Kp index* (Version 1.0) [Dataset]. GFZ Data Services. https://doi.org/10.5880/Kp.0001

Muradoglu, M., Johnsson, M. T., Wilson, N. M., Cohen, Y., Shin, D., Navickas, T., Pyragius, T., Thomas, D., Thompson, D., Moore, S. I., Rahman, M. T., Walker, A., Dutta, I., Bijjahalli, S., Berlocher, J., Hush, M. R., Anderson, R. P., Szigeti, S. S., & Biercuk, M. J. (2025). *Quantum-assured magnetic navigation achieves positioning accuracy better than a strategic-grade INS in airborne and ground-based field trials*. arXiv. https://arxiv.org/abs/2504.08167

Nair, M., Chulliat, A., Fillion, M., & Califf, S. (2025). Model-based Error Grids and Satellite-derived errors for the World Magnetic Model (WMM) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.16954302

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*, 48. https://doi.org/10.1186/s40623-020-01171-9

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., & 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. https://doi.org/10.5047/eps.2013.07.001

Olsen, N., Risbo, T., Brauer, P., Merayo, J., Primdahl, F., & Sabaka, T. (2001). *In-flight calibration methods used for the orsted mission*. Technical University of Denmark.

Peddie, N. W., & Zunde, A. K. (1985). An assessment of the near-surface accuracy of the International Geomagnetic Reference Field 1980 model of the main geomagnetic field. *Physics of the Earth and Planetary Interiors*, *39*(4), 251–259. https://doi.org/10.1016/0031-9201(85)90056-1

Richmond, A. D. (1995). Ionospheric electrodynamics using magnetic apex coordinates. *Journal of Geomagnetism and Geoelectricity*, *47*(2), 191–212. https://doi.org/10.5636/jgg.47.191

Rosner, B. (1983). Percentage points for a generalized ESD many-outlier procedure. *Technometrics*, *25*(2), 165–172. https://doi.org/10.1080/00401706.1983.10487840

Tsyganenko, N. A. (1989). A magnetospheric magnetic field model with a warped tail current sheet. *Planetary and Space Science*, *37*(1), 5–20. https://doi.org/10.1016/0032-0633(89)90066-4

Tsyganenko, N. A., & Sitnov, M. I. (2007). Magnetospheric configurations from a high-resolution data-based magnetic field model. *Journal of Geophysical Research*, 112, A06225. https://doi.org/10.1029/2007JA012260