CGGM-2 model as a candidate for IGRF-14

Based on more than 6-years of data from the China Seismo-Electromagnetic Satellite (CSES), we built the second generation CGGM model, called CGGM-2, which we wish to submit as a candidate model for IGRF-14. This note is to provide the necessary information requested to accompany this model.

1) Team

Lead Institution:
National Institute of Natural Hazards (NINH), Ministry of Emergency Management of China (MEMC)

Secondary Lead Institution:
Université Paris Cité, Institut de Physique du Globe de Paris (IPGP)

Participating Institutions:
National Space Science Center, Chinese Academy of Sciences (NSSC)
Space Research Institute, Austrian Academy of Sciences (IWF)
Danish Technical University (DTU)

Team leader: Yanyan Yang (NINH)

Secondary team leaders: Zhima Zeren (NINH), G. Hulot (IPGP)

Team members: Jie Wang (NINH), Nils Olsen (DTU), Siquan Yang (NINH), Bin Zhou (NSSC), Magnes Werner (IWF), Binjun Cheng (NSSC), Jianpin Huang (NINH),L. Chauvet (IPGP), Jian Lin (NINH), Dapeng Liu (NINH), Hengxin Lu (NINH), Na Zhou (NINH), Rui Yan (NINH), Zhengxia Zhang (NINH), Wei Chu (NINH), Song Xu (NINH), He Huang (NINH), Dehe Yang (NINH), Jing Cui (NINH), Qiao Tan (NINH), Wenjing Li (NINH) , Feng Guo (NINH)

2) Data used

Data source: Only the CSES 1 Hz magnetic field data were used in the modeling. CSES is a sun-synchronous (descending/ascending local time is 14:00/2:00) LEO Satellite (507 km altitude). The onboard High-Precision Magnetometer payload (HPM, Cheng et al. 2018) consists of two fluxgate magnetometers (FGM_S1 and FGM_S2) and a coupled dark state magnetometer (CDSM, Pollinger et al. 2018). All instruments are located on a deployable boom with two hinges. The CDSM provides the scalar data for both science applications and calibrations of the FGMs. These instruments are jointly providing magnetic field vector and scalar measurement from DC to 15Hz.

Data duration The first data is from 06/03/2018 and the last data is from 12/09/2024.

Data type

- *Type 1*: CDSM scalar and FGM vector magnetic field data with geographic latitudes between 65°S and 65°N. These data are from Level 2 scientific data (Yang et al., 2021b). The vector data used are from the FGM_S1 instrument, as this instrument proved to be better suited for the modeling (Yang et al., 2021a). In addition, quaternions providing the attitude of the satellite using star cameras located on the body of the satellite are also used. The data were calibrated by the method introduced in Zhou et al., (2018) and provided by NINH. During the long-term turn-off (Feb to April 2020) and drifting (June 2021 to Feb 2022) of CDSM, the FGM in-orbit calibration parameters are obtained by interpolation from other days (Yang et al., 2023). Then, the scalar field was calculated from the modulus of the magnetic field vector. Both vector and scalar data are subsampled 1 point every 60 points to avoid over-representation along tracks and large calculation works.

- *Type 2:* Additional high latitude CDSM scalar data for North and South geographic latitudes higher than 65°. These data were not contained in Level 2 scientific data since they are affected by magnetically noisy operations/maneuvers at high latitude orbital segments. To build an IGRF candidate model entirely based on CSES data, however, a special effort has been made to make such data available (and our previous CGGM modeling also has proven that these datasets still can be

used in the modeling). To avoid the most spurious data, these processed data were next screened to ensure that no intensity data departed from the CHAOS7-x7.18 (Finlay et al., 2021) model by more than 100 nT. 1 point every 15 points are subsampled for these data.

All data of both types collected in this way were next made available to the modeling team, which next screened and selected the data in the way described below.

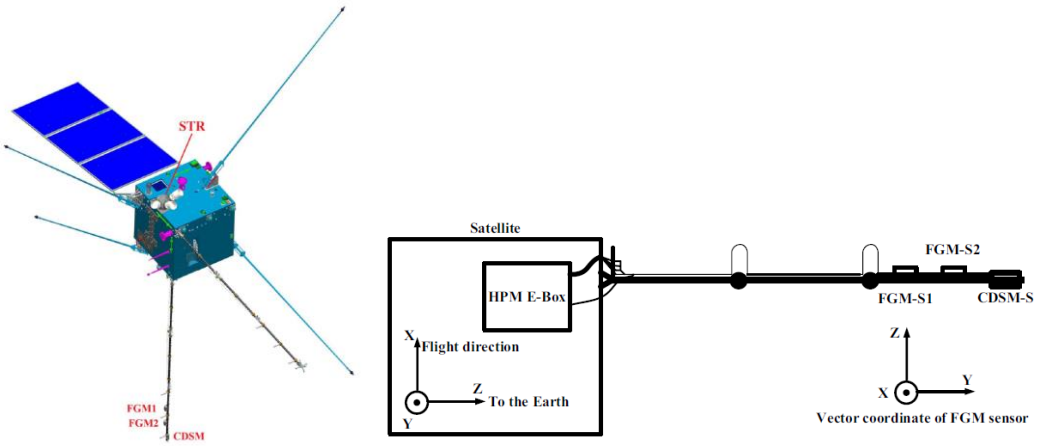

Figure 1: *Location of the various magnetometers and star sensor on the CSES satellite. (from Yang et al., 2021a; 2021b).*

Data selection: Data selection criteria are similar to those we used for our previous CGGM model (see Yang et al., 2021a). In particular, considering the possible boom deformation, we again consider vector data only within (absolute) QD latitudes of less than 20°, while scalar data for all latitudes are being used. As in Yang et al. (2021a), we also assume a fairly large uncertainty (100 arc seconds) on the attitude provided by the quaternions.

The detailed data selection criteria are summarized as follows:

Selection common to all data:

- Sun angle at least 10° below the horizon
- Magnetically quiet conditions (based on RC<2nT/h and Kp<2+) are required
- All data were screened to remove any data not satisfying the criteria that differences between each datum and the prediction from CHAOS-7x18 should be less than 100 nT (scalar or norm comparison)

Additional selection for scalar data:

- A dedicated Flag signaling when magnetorquers are on was provided with the data and used to avoid data at times of magnetorquer activation for all type 1 data (Flag MT should be 0). This flag was not used for type 2 data, as all data appeared to be affected.
- Em < 0.8 mV/m for high latitude scalar data

Additional selection for vector data:

- Only vector data for (absolute) QD latitudes less than 20° are selected
- Data with disturbance from Tri-Band-Beacon (TBB) payload is removed (by requesting Flag_TBB=0)
- Scalar residuals (difference between scalar and modulus of the vector) must be less than 2 nT
- The vector residual with CHAOS-7x18 should be less than 30 nT, to remove some large bias data

This resulted in the selection of **509396 scalar data** and **120905x 3 vector data,** distributed in time and latitude as illustrated in Figure 2.

3) Parent model parameterization:

Compared with the previous CGGM model Yang et al. (2021a), the model parameterization for CGGM-2 followed the CHAOS-4 approach (Olsen et al. 2014) and further considered the crustal field and non-linear

variation of the main field. The internal field coefficients are solved to degree and order 45, with the time dependent internal coefficients to n=13. To describe the temporal variation of the main field for more than 5 years, 6 order B-splines are used. The modelling thus involves a total of **4624** coefficients corresponding to:

- Time-varying internal field up to degree and order 13 (included), using a 6 order B-spline, with 1-year knot separation (at t=2019.0, 2020.0, 2021.0, 2022.0, 2023.0, 2024.0 for the interior knots) and fivefold knots (so six external knots) at the endpoint t=2018.0 and t=2025.0, thus 7 splines in total. This led to 7x13x(13+2) =**1365** coefficients.
- Static field between degree and order 14 (included) and degree and order 45 (included). This led to an additional 45 x (45+2) −13x(13+2) = **1920** coefficients
- External field modeled as in Hulot et al. (2015, where details can be found):
    - Remote magnetospheric sources : $q_n^{0,GSM}$ in GSM frame, with n=1,2 -> 2 coefficients
    - Near magnetospheric ring current :
      Up to degree and order 2 in SM frame -> 8 coefficients
      Δq estimated every 5 days for q10 (470 time segments, -> 470 coefficients)
      Δq estimated every 30 days for qs11 (80 time segments, -> 2x80 = 160coefficients)

  Leading to a total of 2+8+470+160 = **640** coefficients
- Euler angles (rotation between FGM_S1 and STR reference frames) estimated every 10 days: 233 time segments, -> 3x233= **699** coefficients

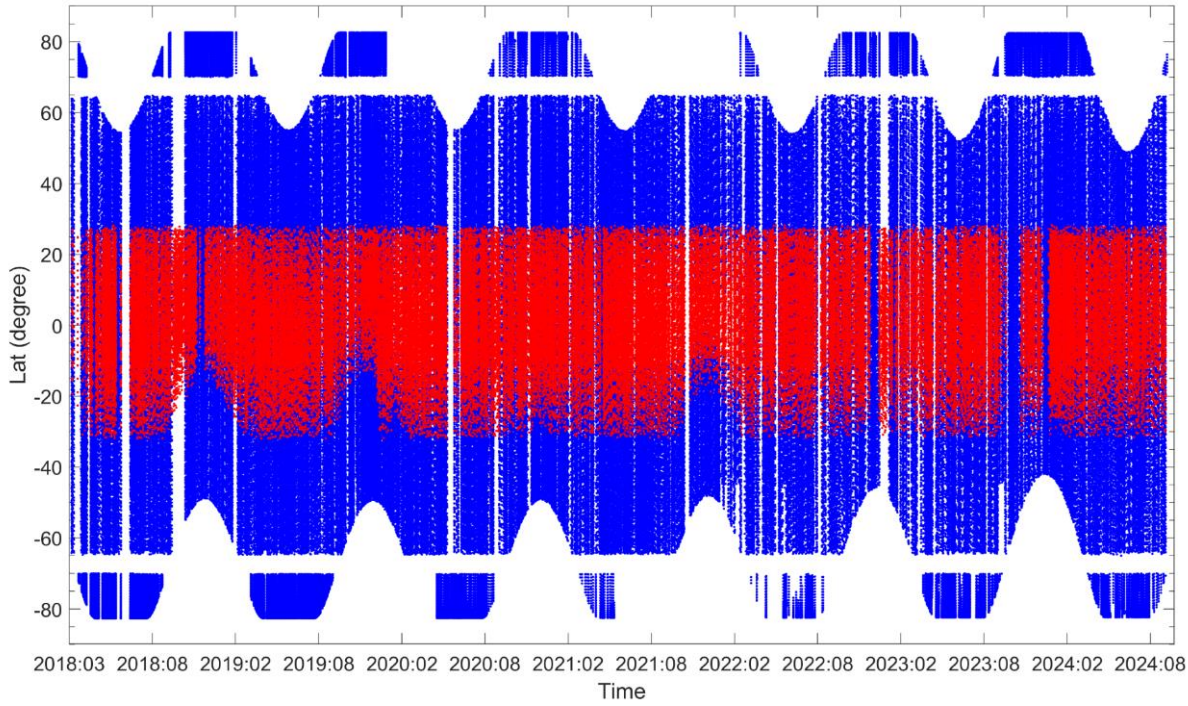

***Figure 2**: Data distribution as a function of time and latitude (blue: scalar data; red: vector data)*

4) Model estimation (inversion):

The model was computed as in Olsen et al., (2014) by using iteratively reweighted least-squares with Huber weights and temporal damping, to minimize the cost function:

$$\mathbf{e}^T \underline{\underline{\mathbf{C}}}^{-1} \mathbf{e} + \lambda_3 \mathbf{m}^T \underline{\underline{\Lambda}}_3 \mathbf{m} + \lambda_2 \mathbf{m}^T \underline{\underline{\Lambda}}_2 \mathbf{m},$$

In which $\mathbf{m}$ is the model vector contains parameters described in section 3) and $\mathbf{e} = \mathbf{d}_{obs} - \mathbf{d}_{\mathrm{mod}}$ is the residual between observations $\mathbf{d}_{obs}$ and model prediction $\mathbf{d}_{\mathrm{mod}}$. $\underline{\underline{\mathbf{C}}}$ is the data covariance matrix. λ_2 and λ_3 are damping parameters to control the strength of the regularization respectively for third and second time derivatives of the core field. In CGGM-2 modeling, we use the same λ_2 and λ_3 as Olsen et al., (2014), i.e., λ_2 =10 (nT yr^{-2}) $^{-2}$, λ_3 =0.33 (nT yr^{-3}) $^{-2}$ but with λ_3 =10 (nT yr^{-3})$^{-2}$ for g_1^0. Besides, a geographical weight was also introduced, proportional to sin(θ) (where θ is the geographic colatitude), to balance the geographical sampling of data (Hulot et al., 2015). Similar to previous CGGM model, anisotropic magnetic

errors due to attitude uncertainty were taken into account assuming an isotropic attitude error of 100 arcsecs to account for the limited quality attitude restitution. A priori data error variances were set to 2.5 nT for both scalar and vector data. The starting model used is a static model (CHAOS-4 up to degree and order 13 only for epoch 01/03/18), but this choice was found to not have any influence on the final model.

A total of six iterations were used, which was found to ensure convergence within the accuracy required. Resulting residual statistics are shown in Table 1.

Table 1. *Residual statistics for all data used to produce CGGM-2 parent model*

	Data Number	Raw residual mean	Raw residual RMS	Huber Weight mean（nT）	Huber Weight RMS（nT）
F_P	124517	0.40	10.55	-0.01	5.20
$F_{NP} + B_B$	630301	-0.01	5.22	-0.06	2.93
$F + B_B$	505784	-0.11	2.55	-0.07	2.21
B_B	120905	-0.56	3.46	-0.31	2.89
B_r	120905	-0.06	13.55	-0.07	9.16
B_θ	120905	0.62	6.80	0.58	5.83
B_φ	120905	0.03	14.18	0.14	6.46

Table 1*: Residual statistics for all data used to produce the parent model (using the same convention as in Hulot et al., 2015). "F_p" refers to scalar data above (absolute) QD latitude 55° (polar latitudes). "F_{NP}" refers to scalar data below (absolute) QD latitude 55° (non-polar latitudes). B_B refers to vector residuals projected along the field direction; "B_B" refers to data within (absolute) QD latitude 55°; F (polar) refers to scalar data above (absolute) QD latitude 55°.*

5) IGRF-14 candidate model generation

- DGRF-2020: computed from the parent model at epoch t=2020.0 for the degree n=1-13 part and output to 0.01nT.
- IGRF-2025: computed from the parent model at epoch t=2025.0 for the degree n=1-13 part and output to 0.01nT.
- SV-2025-2030: computed from the parent model at epoch t=2024.0 (rather than epoch t=2025.0 to avoid the possible end effects from the spline-model) for the degree n=1-8 part and output to 0.01nT.

6) Initial validation

To validate our candidate model, we decided to rely on some comparisons of our parent model with the CHAOS 7-x18 model of Finlay et al. (2021). These comparisons were first made at epoch times of respectively 2020.0 (for DGRF-2020) and 2024.0 (for SV-2025-2030). To estimate the likely quality of the IGRF-2025 model, we followed the same strategy as the one used by Yang et al., (2021a). Since our CGGM-2 parent model only uses data up to Sep 2024 (and the CHAOS 7-x18 only used data even before that), comparisons of predictions for epoch 2025.0 was not considered appropriate. In contrast, the CHAOS 7-x18 can be considered to provide a very reliable estimate of the main field for 15/11/2017 (2017.87), which is 111 days before the very first data used in our parent model. This is the same amount of time separating the last data used in our parent model and epoch 2025.0. Given the symmetry of the CSES data distribution we used (recall Figure 2), we consider this backward extrapolation test as a good test of how well our IGRF 2025.0 candidate model will likely perform.

Figure 3 illustrates the difference in the Br values predicted by the CGGM-2 parent model and CHAOS 7-x18 at Earth's surface, for epoch 2020.0. As can be seen, most differences are of zonal nature, with amplitudes of about 20nT. These features are similar to what was reported in the previous CGGM model and most likely

reflect some of the systematic boom deformation along the CSES orbits. In addition, high latitude scalar data (Type 2 data) contain magnetic field disturbances from the operation of magnetorquers (within ~20nT), which are also another main source leading to bias here. We note, however, that such differences remain within a reasonable level for typical IGRF candidate models (e.g., Alken et al., 2021). The backward extrapolated for epoch 2017.87 (Figure 4) indicates residuals within 24nT (Figure 4). Figure 5 shows the SV difference in the Br component given by the CGGM-2 parent model and CHAOS 7-x18 at Earth's surface, for epoch 2024.0. The result again shows zonal nature, but with amplitudes < 7nT/yr, which is reasonnably small.

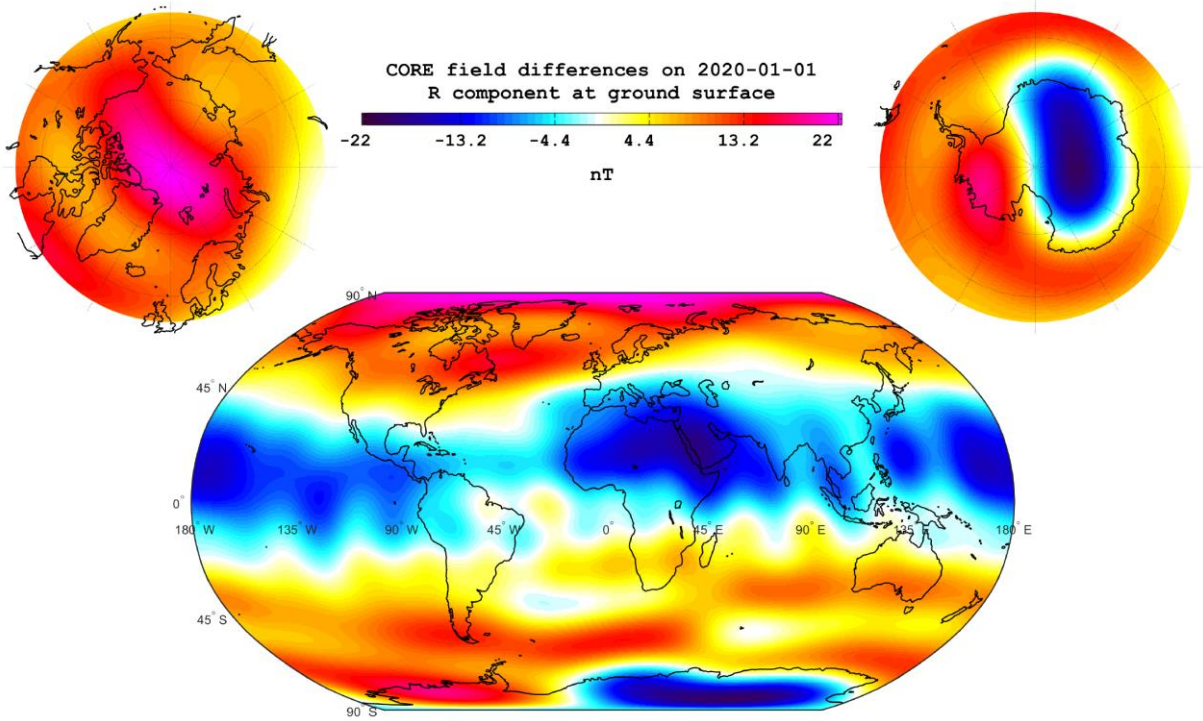

Figure 3: *Difference in the Br values predicted by the CGGM-2 parent model and CHAOS7-x18 at Earth's surface for epoch 1 Jan 2020.*

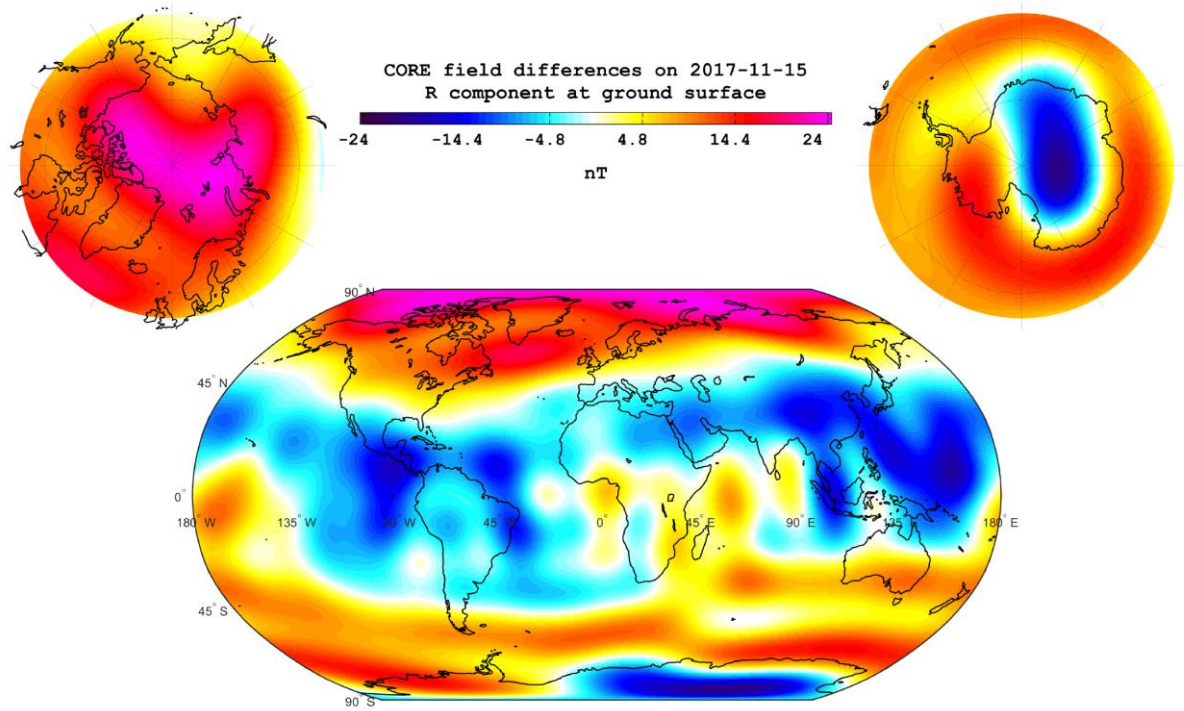

Figure 4: *Difference in the Br values predicted by the CSES parent model and CHAOS7-x18 at Earth's surface for backward extrapolated epoch 15/11/2017.*

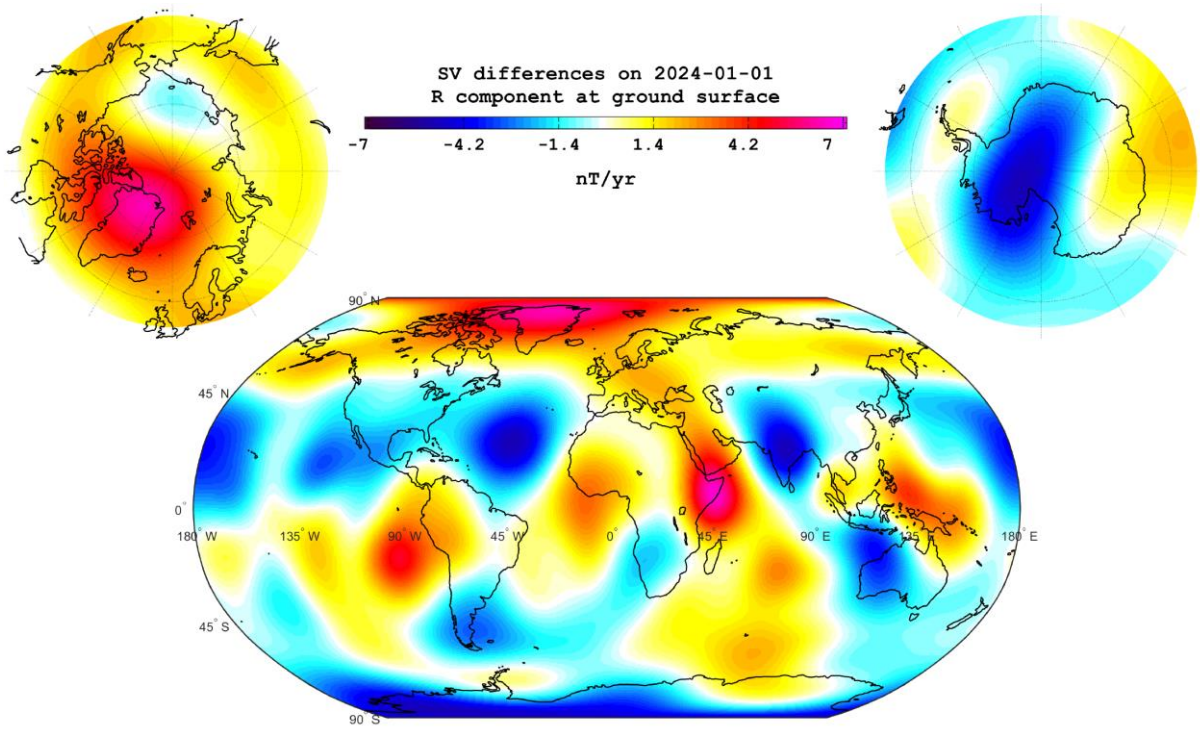

Figure 5: *The SV difference in the Br component predicted by the CGGM-2 parent model and CHAOS7-x18 at Earth's surface for epoch 1 Jan 2024.*

To further assess the quality of the CGGM-2 candidate model, we also computed the Lowes-Mauersberger spectra of the differences between the field predictions of our parent model and the CHAOS7-x18 model for epoch 2020.0 (Figure 6, for assessing our DGRF-2020 candidate model), epoch 2017.87 (as a mesure of the likely quality of our IGRF-2025 candidate model) and of the SV predictions at epoch 2024.0 (Figure 8, for assessing our SV-2025-2030 candidate model). The spectra in Figure 6 lead to the same orders of magnitude for the disagreements between the two models as illustrated in Figure 3. It also highlights the fact that disagreements are strongest for degrees 2 and 3, corresponding to the large-scale zonal structures seen in Figure 3. Focussing on the backward extrapolated epoch 2017.87, the best representative of the likely error affecting our IGRF-2025 candidate model (which, we recall amounts to forwarding our CSES parent model over an identical period of time), we can see that this spectral error can be expected to reach 50 nT^2 for degree 2, 30 nT^2 for degree 3 while remaining well below 10 nT^2 for all higher degrees. We again interpret the strongest disagreements for degrees 2 to 3 as a likely consequence of some systematic boom deformation along the CSES orbits.

Despite the above deficiencies, the SV power spectra at epoch 2024.0 in Figure 8 indicate very good consistency between the SV predicted by the CGGM-2 and CHAOS 7-x18 models. The good behaviour of the SV predicted from CGGM-2 may well benefit from the special revisiting orbits of CSES, which can revisit the same place every 5 days. This result shows CSES's better ability to describe the magnetic field variation (despite its disadvantage in boom deformation and magnetic field disturbance at high latitudes).

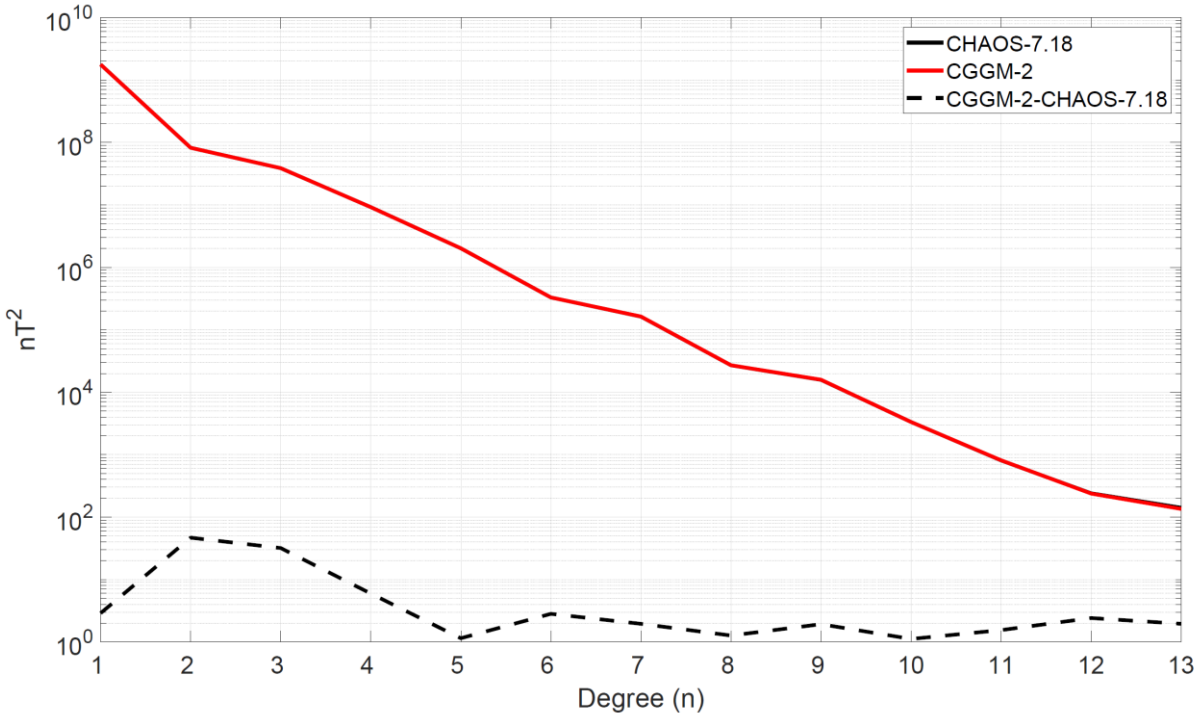

***Figure 6**: Spectra of our DGRF-2020 candidate model (CGGM-2, red line) and the difference with the CHAOS7-x18 model (black dashed line, referred to as CGGM-2– CHAOS-7.18) for epoch 1/1/2020, both at Earth's surface.*

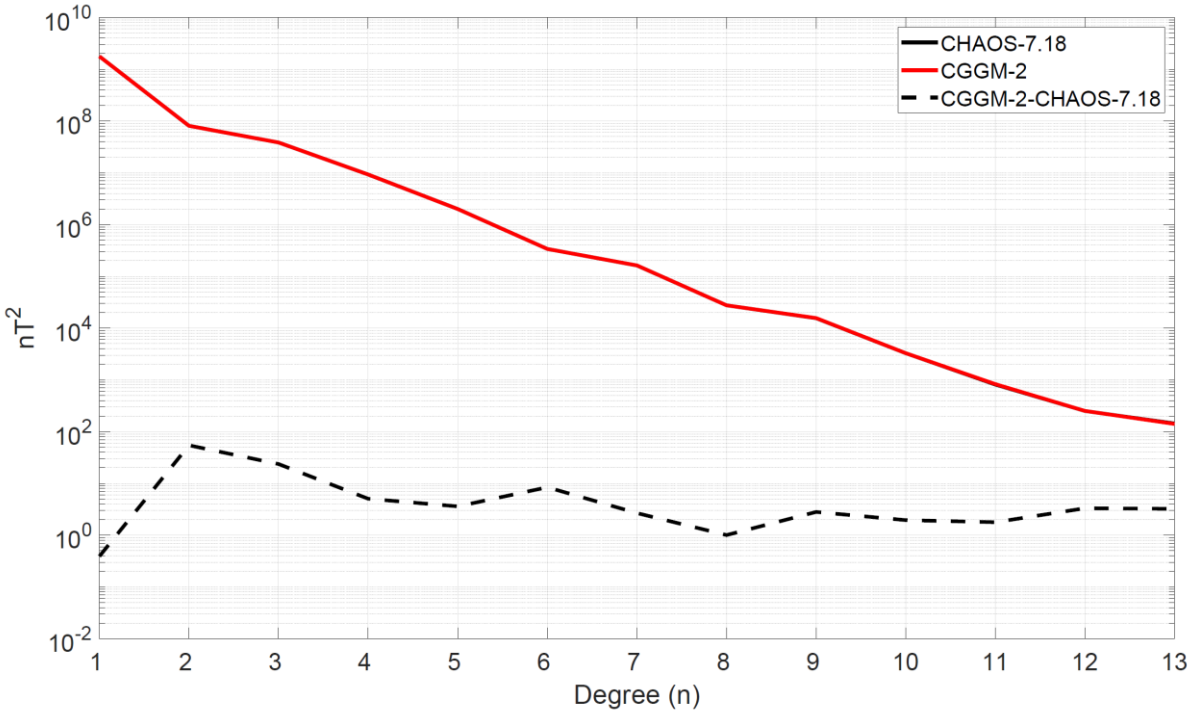

***Figure 7**: Same as Figure 6, but for the CGGM-2 model (red line) and difference with the CHAOS7-x18 model (black dashed line, referred to as CGGM-2– CHAOS-7.18) backward extrapolated epoch 15/11/2017, for assement of errors on our IGRF-2025 candidate model*

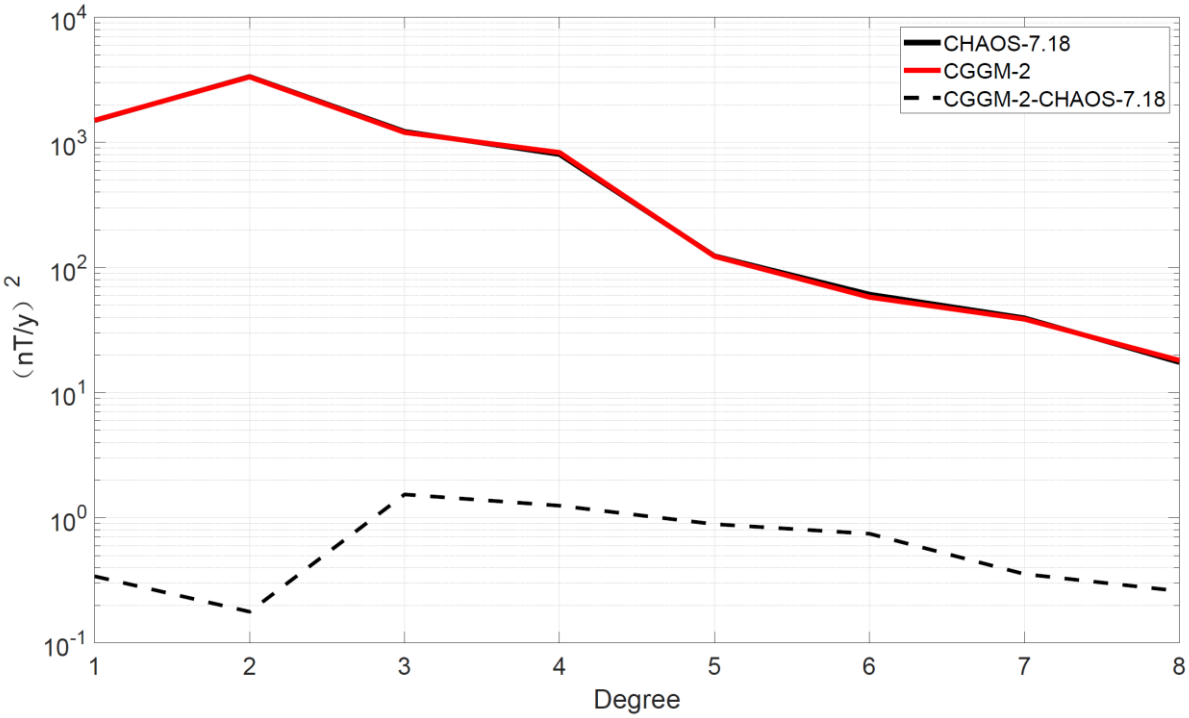

***Figure 8**: Same as Figure 6, but for the SV predicted in 2024, for assement of errors on our SV-2025-2030 candidate model*

7) Computation of realistic uncertainties on each Gauss coefficient

Following the method used in the previous CGGM model, the "realistic" uncertainties of our CGGM-2 candidate model are calculated by the following RMS quantity:

$$\sigma_n = \text{sqrt}\left[\left(\sum\{ dg^m{}_n{}^2 + dh^m{}_n{}^2\} \right) / (2n+1) \right]$$

where $dg^m{}_n$ and $dh^m{}_n$ are the differences in the $g^m{}_n$ and $h^m{}_n$ Gauss coefficients from our model and CHOAS 7-x18. In our submission, such uncertainties are provided both for DGRF-2020, IGRF-2025 and SV-2025-2030, in which $g^m{}_n$ and $h^m{}_n$ are calculated respectively at epoch 2020.0, 2017.87 and 2024.0.

We then simply assigned this σ_n as our best estimate of the errors (one sigma type) affecting each Gauss coefficient of degree n.

8) Conclusion

With the preliminary assessment provided in this note, we conclude that the CGGM-2 candidate model entirely based on CSES data could be used for contribution to the final IGRF-14 model.

9) References:

Alken P, Thébault E, Beggan C, Aubert J, Baerenzung J, Brown W, Califf S, Chulliat A, Cox G, Finlay CC, Fournier A, Gillet N, Hammer MD, Holschneider M, Hulot G, Korte M, Lesur V, Livermore P, Lowes F, Macmillan S, Nair M, Olsen N, Ropp G, Rother M, Schnepf NR, Stolle C, Toh H, Vervelidou F, Vigneron P, Wardinski I (2021) Evaluation of candidate models for the 13th International Geomagnetic Reference Field. Earth Planets Space. https ://doi. org/10.1186/s4062 3-020-01281 -4

Cheng B. J., Zhou B., Magnes W., et al., High precision magnetometer for geomagnetic exploration onboard of the China Seismo-Electromagnetic Satellite, *Sci China Tech Sci*, 61: 659–668, https://doi.org/10.1007/s11431-018-9247-6, 2018.

Finlay, C.C., Kloss, C., Olsen, N., Hammer, M. Toeffner-Clausen, L., Grayver, A and Kuvshinov, A. (2020), The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly, Earth Planets and Space 72, doi:10.1186/s40623-020-01252-9

Hulot, G., Vigneron, P., Léger, J.-M., Fratter, I., Olsen, N., Jager, T., Bertrand, F., Brocco, L., Sirol, O., Lalanne, X., Boness, A., Cattin, V., Swarm's absolute magnetometer experimental vector mode, an innovative capability for space magnetometry, *Geophys. Res. Lett.*, 42, https://doi.org/10.1002/2014GL062700, 2015.

Olsen N, Lühr H, Finlay CC, Sabaka TJ, Michaelis I, Rauberg J, Tøffner-Clausen L (2014) The CHAOS-4 geomagnetic field model. Geophys J Int 197:815–827. https ://doi.org/10.1093/gji/ggu033

Pollinger A., Lammegger R, Magnes W, et al., Coupled dark state magnetometer for the China seismo-electromagnetic satellite, *Meas. Sci. Technol*, 29, 095103, https://doi.org/10.1088/1361-6501/aacde4, 2018.

Yang, Y., Hulot, G., Vigneron, P., Shen, X., Zhima, Z., Zhou, B., et al. (2021a). The CSES global geomagnetic field model (CGGM): An IGRF-type global geomagnetic field model based on data from the China Seismo-Electromagnetic Satellite. Earth, Planets and Space, 73(1). 1–21. https://doi.org/10.1186/s40623-020-01316-w

Yang, Y., Zhou, B., Hulot, G., Olsen, N., Wu, Y., Xiong, C., et al. (2021b). CSES high precision magnetometer data products and example study of an intense geomagnetic storm. Journal of Geophysical Research: Space Physics, 126, e2020JA028026. https://doi. org/10.1029/2020JA028026

Yanyan Yang, Zeren Zhima, Xuhui Shen, Bin Zhou, JieWang, Werner Magnes, Andreas Pollinger,Hengxin Lu, Feng Guo, Roland Lammegger , Na Zhou, Yuanqing Miao, Qiao Tan and Wenjing Li (2023). An Improved In-Flight Calibration Scheme for CSES Magnetic Field Data. Remote Sens, 15, 4578. https://doi.org/10.3390/rs15184578

Zhou, B., Yang, Y. Y., Zhang, Y. T., Gou, X. C., Cheng, B. J., Wang, J. D., and Li, L., Magnetic field data processing methods of the China Seismo-Electromagnetic Satellite. *Earth Planet. Phys.*, 2(6), 455–461, http://doi.org/10.26464/epp2018043, 2018.