

FULL PAPER

Open Access



# The CSES global geomagnetic field model (CGGM): an IGRF-type global geomagnetic field model based on data from the China Seismo-Electromagnetic Satellite

Yanyan Yang<sup>1</sup>, Gauthier Hulot<sup>2\*</sup>, Pierre Vigneron<sup>2</sup>, Xuhui Shen<sup>1\*</sup> , Zeren Zhima<sup>1</sup>, Bin Zhou<sup>3</sup>, Werner Magnes<sup>4</sup>, Nils Olsen<sup>5</sup>, Lars Tøffner-Clausen<sup>5</sup>, Jianpin Huang<sup>1</sup>, Xuemin Zhang<sup>6</sup>, Shigeng Yuan<sup>7</sup>, Lanwei Wang<sup>1</sup>, Bingjun Cheng<sup>3</sup>, Andreas Pollinger<sup>4</sup>, Roland Lammegger<sup>8</sup>, Jianpin Dai<sup>9</sup>, Jun Lin<sup>10</sup>, Feng Guo<sup>1</sup>, Jingbo Yu<sup>11</sup>, Jie Wang<sup>1</sup>, Yingyan Wu<sup>6</sup>, Xudong Zhao<sup>12</sup> and Xinghong Zhu<sup>7</sup>

## Abstract

Using magnetic field data from the China Seismo-Electromagnetic Satellite (CSES) mission, we derive a global geomagnetic field model, which we call the CSES Global Geomagnetic Field Model (CGGM). This model describes the Earth's magnetic main field and its linear temporal evolution over the time period between March 2018 and September 2019. As the CSES mission was not originally designed for main field modelling, we carefully assess the ability of the CSES orbits and data to provide relevant data for such a purpose. A number of issues are identified, and an appropriate modelling approach is found to mitigate these. The resulting CGGM model appears to be of high enough quality, and it is next used as a parent model to produce a main field model extrapolated to epoch 2020.0, which was eventually submitted on October 1, 2019 as one of the IGRF-13 2020 candidate models. This CGGM candidate model, the first ever produced by a Chinese-led team, is also the only one relying on a data set completely independent from that used by all other candidate models. A successful validation of this candidate model is performed by comparison with the final (now published) IGRF-13 2020 model and all other candidate models. Comparisons of the secular variation predicted by the CGGM parent model with the final IGRF-13 2020–2025 predictive secular variation also reveal a remarkable agreement. This shows that, despite their current limitations, CSES magnetic data can already be used to produce useful IGRF 2020 and 2020–2025 secular variation candidate models to contribute to the official IGRF-13 2020 and predictive secular variation models for the coming 2020–2025 time period. These very encouraging results show that additional efforts to improve the CSES magnetic data quality could make these data very useful for long-term monitoring of the main field and possibly other magnetic field sources, in complement to the data provided by missions such as the ESA Swarm mission.

**Keywords:** CSES, CGGM, IGRF, Geomagnetism, Space magnetometry

## Introduction

The International Geomagnetic Reference Field (IGRF) is a series of mathematical models used to describe the large-scale internal part of the geomagnetic field. The building of these models is an international endeavour carried out under the auspices of the International Association of Geomagnetism and Aeronomy (IAGA). Every

\*Correspondence: gh@ipgp.fr; shenxh@seis.ac.cn

<sup>1</sup> National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China

<sup>2</sup> Université de Paris, Institut de Physique du Globe de Paris, CNRS, 75005 Paris, France

Full list of author information is available at the end of the article

5 years, these models are updated after IAGA releases an open call to the international community to collect candidate models, which are next assessed and used to build the final official IGRF update (see, e.g., Macmillan and Finlay (2011) for more details about IGRF). The previous update (IGRF-12) was published in 2015 (Thébault et al. 2015a) and consisted in a series of snapshot models every 5 years between 1900 and 2015, and a predictive secular variation model, describing the expected average (linear) temporal variation of the field between 2015 and 2020. A new update (IGRF-13) is now in order, consisting in: (1) replacing the previous 2015 model by an improved 2015 model (taking into account data acquired since the last update), (2) providing a new model for epoch 2020, and (3) providing a new secular variation model to describe the expected average (linear) temporal variation of the field between 2020 and 2025. The corresponding call for candidate models has been issued in March 2019 by an IAGA dedicated task force, with an October 1, 2019 deadline.

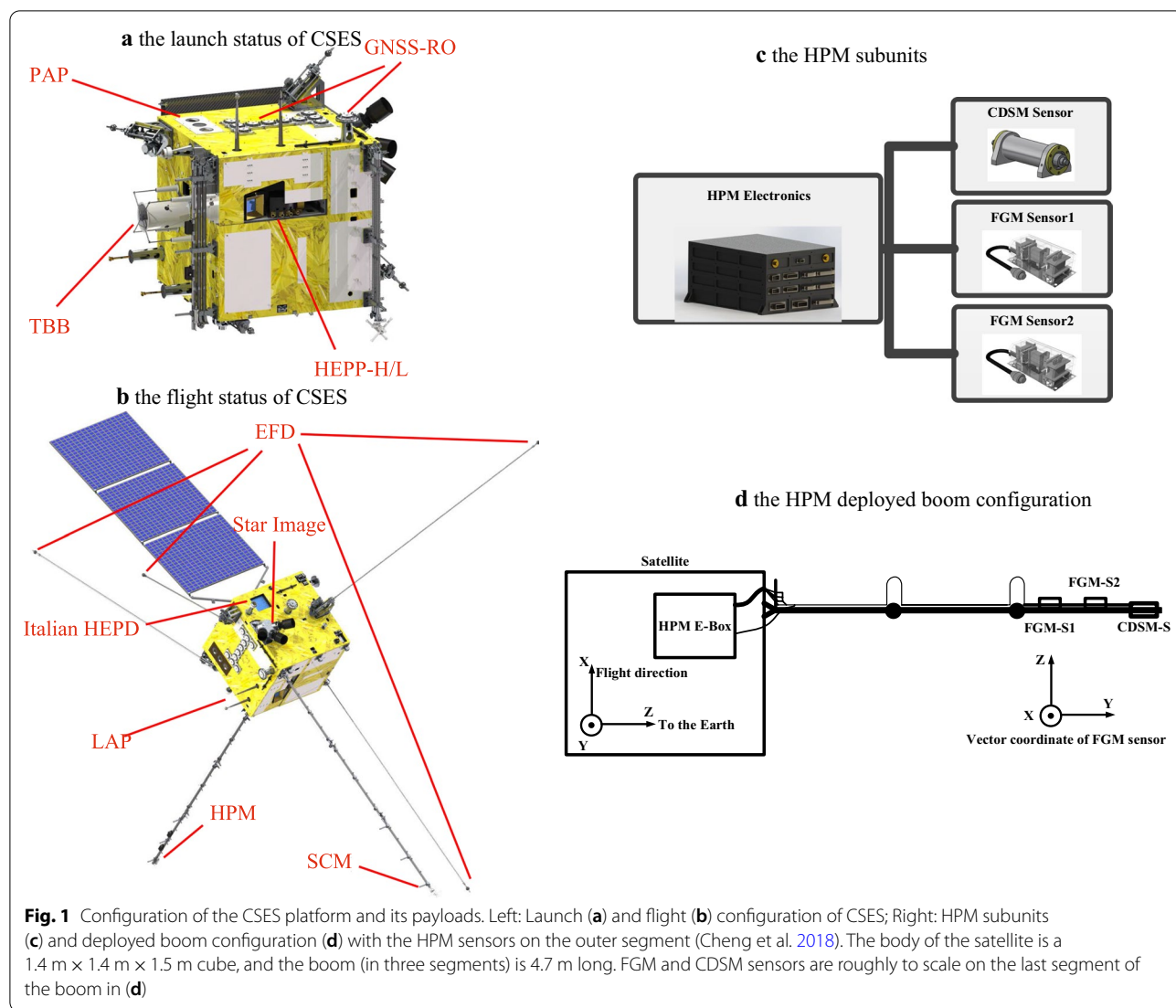
The present paper describes the way one such candidate model for epoch 2020 has been derived (and submitted to the call) using data from the China Seismo-Electromagnetic Satellite (CSES) mission, launched on February 2, 2018 (Shen et al. 2018). This candidate model, to which we will refer as the CGGM (CSES Global Geomagnetic Field Model) candidate model, is the first ever produced by a Chinese-led team. It also is the first produced by only relying on data from a Chinese satellite. It finally is the only 2020 IGRF candidate model not relying on any data from the ESA Swarm constellation (see Alken et al. 2020a).

The CSES satellite is orbiting on a Sun-synchronous low Earth circular orbit, at an altitude of about 507 km and with an inclination of  $97.4^\circ$ . It has a fixed 14:00 local time (LT) at descending node and a 5-day ground track recursive period. Its main scientific purpose is to acquire electric and magnetic field data, as well as plasma and high energetic particles data for the study of signals related to earthquakes, geophysics, and space science (see Shen et al. 2018). Nine payloads are operated on CSES (see Fig. 1). Six booms are used, one for the sensors of the High Precision Magnetometer (HPM) payload, one for a Search-Coil Magnetometer (SCM), and four for Electric Field Detectors (EFD). The other six payloads are assembled on the body of the satellite, where a set of three star imagers (STR) is also located to provide attitude restitution. These are a Plasma Analyzer (PAP), a Langmuir Probe (LAP), a High Energetic Particle Package from China (HEPP), a High Energetic Particle Detectors from Italy (HEPD), a GNSS Occultation Receiver (GNSS-RO), and a Tri Band Beacon (TBB), the latter being operated in coordination with ground receiver stations in Chinese

territory. Generally, except for some individual indicators, all payloads perform very well in orbit and meet their designed technical requirements (e.g., Yang et al. 2020 and references therein).

The main payload of interest for the present study is the HPM (Cheng et al. 2018) used to measure the magnetic field vector and intensity from DC to 15 Hz. As shown in Fig. 1c, d, the HPM consists of two fluxgate magnetometers (FGM-S1 and FGM-S2, to measure the magnetic field vector) and one coupled dark state magnetometer (CDSM, to provide the scalar data for both science applications and calibrations of the FGMs, Pollinger et al. 2018). All instruments are located on the last leg of a deployable boom with three hinges. FGM-S1 is the nearest to the satellite body (about 3.9 m) and CDSM is the farthest (about 4.7 m). The distance between sensors (FGM-S1 to FGM-S2 and FGM-S2 to CDSM) is about 0.4 m. This set-up was chosen to minimize perturbations among instruments and from the satellite body itself. It, however, has the drawback that the mechanical link between the FGM instruments providing the vector measurements on the boom and the STR providing the attitude restitution on the satellite body is complex and subject to possible deformation along the orbit. As we shall later see, this, indeed, is a significant limitation.

The preparation and production of the CGGM candidate model involved several steps, and the organisation of the present paper reflects these steps. We first introduce the characteristics of the CSES HPM data used in this study, as well as that of Swarm data used in preliminary modelling studies. We next describe early attempts to build main field models from CSES HPM data, which we compared to main field models built in a similar way from Swarm data. The purpose of this was to assess if CSES HPM data were of high enough quality to build a candidate model meeting the standards of IGRF. This revealed some significant limitations and guided us in our final modelling strategy. We then move on to describe the way a CGGM parent model was built, first describing the data selection strategy, next describing the model parameterization and optimization strategy, and providing key statistics. We also explain how this parent model was next used to build the CGGM IGRF 2020 candidate model. Finally, we describe the tests we carried out to assess the quality and limitations of this candidate model, and the way we derived realistic uncertainties for each Gauss coefficient. This information was provided with the CGGM candidate model on time for the October 1, 2019 deadline. We conclude with an a posteriori assessment of both this CGGM IGRF 2020 candidate model and the secular variation associated with the CGGM parent model. This assessment encouragingly reveals that, despite their current limitations, CSES data can already



be used to produce useful IGRF 2020 and 2020–2025 secular variation candidate models to contribute to the official IGRF-13 2020 and predictive secular variation models for the coming 2020–2025 time period.

**Data used in this study**

**CSES HPM data**

The CSES HPM data that we used are 1 Hz level 2 scientific HPM data (version 1.0). The data are calibrated using the procedure described in Zhou et al. (2018), Zhou et al. (2019), and Pollinger et al. (2020) and provided by the National Institute of Natural Hazards, Ministry of Emergency Management of China. For the purpose of this study, two distinct sets of level 2 data

were used, which we will refer to as Type 1 and Type 2 data.

Type 1 data are the nominal data of the mission, only provided for CSES geographic locations between 65°S and 65°N (i.e., not at high latitudes). The reason for this is that, as already noted, the CSES mission was not originally intended to provide data for main field modelling. The corresponding 1 Hz level 2 data are produced from the original 60 Hz FGM and 1 Hz CDSM data. These data are provided on a half orbit basis and calibrated in several steps (see Zhou et al. (2018), (2019) and Pollinger et al. (2020) for detailed explanations). The FGM (from both FGM, recall Fig. 1) and CDSM raw signals are first converted to physical quantity, using calibration parameters determined on ground before the launch of the satellite. The three axes of the two FGMs

not being strictly orthogonal, CDSM scalar measurements are next used to calibrate these FGM instruments in orbit, to correct for non-orthogonality, biases, and rescale each axis. The corresponding parameters are calculated separately for the day- and night-side and updated every day. Interferences from the satellite and other neighbouring sensors are also further removed. However, occasional significant disturbances from magnetotorquers (MT) and the TBB instrument could not be corrected for. These can be identified from the flags provided with the CSES level 2 data and then removed during the data selection. This then leads to scalar data from the CDSM on one hand, and to calibrated vector data from the FGM-S1 and FGM-S2 in their respective (orthogonalized) instrument reference frames, on the other hand. Although the CSES mission further provides 1 Hz Level 2 FGM-S1 and FGM-S2 data in the North East Centre (NEC) reference frame after an additional processing step, we do not use these in our modelling procedure. Rather, we directly take joint advantage of the 1 Hz Level 2 FGM data provided in the instrument frame, and of the 1 Hz quaternions describing the rotation to change from the STR reference frame to the Inertial Celestial Reference Frame (ICRF) frame of reference (STR data, also provided as a CSES product).

Type 2 data are additional scalar data later made available, motivated by the need to also have access to scalar high-latitude data for the purpose of building a global field model. These additional 1 Hz scalar data were only made available for North and South geographic latitudes higher than  $65^\circ$  (and sometimes only at even higher latitudes, see Fig. 8 below). They underwent the same calibration procedure as Type 1 data. However, these data not originally being intended to be produced by the CSES mission, they suffer from a number of specific issues. In particular, the way the CSES mission is being operated implies that most magnetically noisy operations and manoeuvres take place during these high-latitude orbital segments. In addition, these data were found to suffer from timing inaccuracy. As a result, Type 2 data underwent additional non-nominal dedicated processing, starting from available satellite low-level data and using GPS time to timestamp the data.

All data of both types collected in this way were made available to the modelling team, which next screened and selected the data in the way we later describe.

#### Swarm data

For the purpose of investigating the ability of CSES to provide enough adequate data for building an IGRF model, a number of preliminary tests were done by also using Swarm data. These data were Level 1b 1 Hz

magnetic data version 0505/0506 from Swarm Alpha between August 01 and September 30, 2018, at a time when this satellite was orbiting at a similar local time as the CSES mission. Note that none of these data were used in the building of the CGGM final parent and candidate models.

#### Auxiliary data

In addition to the satellite data described above, we also relied on the planetary (3 h) geomagnetic Kp index (see Bartels 1949 and e.g., Menvielle and Berthelier 1991), the so-called Ring Current index RC introduced by Olsen et al. (2014), and  $E_m$ , the weighted average over the preceding hour of the merging electric field at the magnetopause (see, e.g., Kan and Lee 1979).

#### Early modelling attempts

In the early phase of this study, only Type 1 CSES HPM data were available. These data only cover geographic latitudes between  $65^\circ$  S and  $65^\circ$  N. To be able to build preliminary main field models, it was, therefore, decided to complement this data set with scalar data from the Swarm Alpha satellite. The goal was to test the value of the Type 1 CSES HPM data for such modelling purposes. The strategy we adopted was to focus on a simple modelling strategy only using 2 months of data (August–September 2018) when CSES Type 1 data were available and Swarm Alpha was orbiting at a similar local time, providing high-latitude scalar data distribution roughly mimicking the scalar data distribution CSES Type 2 data could ultimately provide. The data selection and modelling strategy used was kept simple to match the only 2 months data availability, and inspired by standard data selection and modelling strategies, such as that used by Vigneron et al. (2015) and Hulot et al. (2015a).

#### CSES HPM data selection

Only Type 1 CSES HPM data between August 01 and September 30, 2018 were used. 1 Hz scalar data were taken from the CDSM instrument without any geographic restriction (except for the fact, of course that no Type 1 CSES HPM data were available at high geographic latitudes beyond  $65^\circ$  S and  $65^\circ$  N). Since FGM-S2 was further away from the satellite body (see Fig. 1), Type 1 CSES HPM data from this instrument were initially assumed to be of the best quality (rather than FGM-S1), and thus selected for providing the needed 1 Hz vector data (expressed in the instrument's reference frame). These were further selected according to Quasi-Dipole (QD) latitude (Richmond 1995), using two alternative choices (for testing purposes). A first selection involved selecting vector data at QD latitudes between  $-55^\circ$  and  $+55^\circ$  (to which we will refer as the  $55^\circ$ QD selection). A



second selection involved selecting vector data at QD latitudes between  $-20^\circ$  and  $+20^\circ$  (to which we will refer as the  $20^\circ$ QD selection). To avoid spurious data (due to interference by e.g., the TBB instrument) all vector and scalar data were also screened to ensure that no scalar data (or modulus of the vector data) departed from predictions by the CHAOS-6- $\times$ 8 model (latest version of the CHAOS-6 model of Finlay et al. (2016) available at the time) by more than 300 nT. Such pre-screening of data using a reasonable prior model is standard practice (see, e.g., Finlay et al. 2016; Vigneron et al. 2015; Hulot et al. 2015a) to remove the relatively few most obvious outliers without biasing the bulk of the data towards the chosen prior model (the choice of the 300 nT ensuring this). In addition, for both vector and scalar data, only night-side data were used, using classical criteria to avoid perturbations due to external sources (LT between 18:00 and 06:00,  $K_p < 2+$ ,  $RC < 2$ ). Finally, all data were decimated (one point every 2 min) to avoid noise correlation between consecutive data and oversampling along the satellite track, while keeping enough data, given the targeted level of modelling.

#### Swarm alpha data selection

Swarm Alpha data were used for two different purposes. The first was to provide the scalar data at QD latitudes poleward of  $\pm 55^\circ$  needed to complement the Type 1 CSES HPM data to be able to produce main field models. This first set of data was selected according to the same criteria as the Type 1 CSES HPM data, further requesting that  $E_m < 10$  mV/m, and decimated in the same way.

The second purpose was to provide additional data for building reference models entirely based on Swarm data, over the same August to September 2018 time period, sharing similar local time properties and selection criteria as the CSES data. In addition to the previous high QD latitudes scalar data, two additional Swarm Alpha data sets were thus prepared, including 1 Hz scalar and vector data (expressed in the Swarm Alpha VFM vector field magnetometer reference frame) and selected according to similar criteria as either the  $55^\circ$ QD selection (first data set) or the  $20^\circ$ QD selection criteria (second data set) described above for the CSES HPM data. These data were again decimated in the same way.

#### Model parameterization and optimization

Model parameterization was chosen to be the same for the four models we derived in this preliminary series of tests (one CSES model and one Swarm model for each  $55^\circ$ QD or  $20^\circ$ QD data selection). This parameterization is a simplified version of that used by Vigneron et al. (2015) and Hulot et al. (2015a). Simplification involved parameterizing the main field only up to spherical harmonic

(SH) degree and order 15, and only allowing for a linear secular variation (SV) up to degree and order 5. This maximum degree was chosen to account for the fact that only 2 months of data were considered, and that changes in the field due to higher degree SV during such a short period are below the resolution of the data and cannot be resolved. No special procedure was used to handle the crustal field signal above degree 15 (which is neither modelled, nor removed), since this signal also appears to mainly be beyond recovery with just 2 months of data. To describe the external (magnetospheric) and corresponding Earth-induced fields, we mainly followed the CHAOS-4 model parameterization (Olsen et al. 2014, also used by Hulot et al. 2015a). In practice, however, only simplified parameters to account for the remote magnetospheric sources and the near magnetospheric ring current were included. Using the notation of Olsen et al. (2014, see their Eqs. 4 and 5), remote magnetospheric sources (and their induced counterparts) are thus described by a zonal external field up to degree 2 in geocentric solar magnetospheric (GSM) coordinates (2 coefficients,  $q_1^{0,GSM}$  and  $q_2^{0,GSM}$ ), while the near magnetospheric ring current (and its induced counterpart) is described using solar magnetic coordinates (SM, see Hulot et al. 2015b, for definitions of the GSM and SM coordinate systems, and Maus and Luehr 2005 for the justification of such an approach). However, only a static field up to degree 2 ( $\Delta q_1^0$ ,  $\Delta q_1^1$ ,  $\Delta s_1^1$ ,  $q_2^0$ ,  $q_2^1$ ,  $q_2^2$ ,  $s_2^1$ ,  $s_2^2$ ) and a time-varying part proportional to the RC index for degree 1 ( $\hat{q}_1^0$ ,  $\hat{q}_1^1$ ,  $\hat{s}_1^1$ ) are assumed, leading to 11 parameters in total. Finally, only one set of Euler angles (assumed static throughout the two months time period considered) was also solved for to recover the unknown rotation between the vector instruments (FGM-S2 for CSES, VFM for Swarm) and the STR data provided by each mission. This choice was intended for potential issues with the stability of this rotation to best manifest themselves in the data residuals (see Figs. 6 and 7 and later discussion). In total, 306 parameters were thus solved for, 255 for the static Gauss coefficients, 35 for the linear SV, 13 parameters for the external field, and 3 for the Euler angles.

For solving the inverse problem, we relied on an iteratively reweighted least-squares algorithm with Huber weights (as in Olsen et al. 2014, see also e.g., Farquharson and Oldenburg 1998). The cost function to minimize is  $e^T C^{-1} e$ , where  $e = d_{obs} - d_{mod}$  is the difference between the vector of observations  $d_{obs}$  (in the reference frame of the instrument) and the vector of model predictions  $d_{mod}$ , and  $C$  is the data covariance matrix (updated at each iteration). No regularization was applied, but a geographical weight was introduced, proportional to  $\sin(\theta)$  (where  $\theta$  is the geographic co-latitude), to balance the geographical sampling of data. Both scalar data

and Huber weights make the cost function nonlinearly dependent on the model parameters. The solutions were, therefore, obtained iteratively, using a Newton-type algorithm.

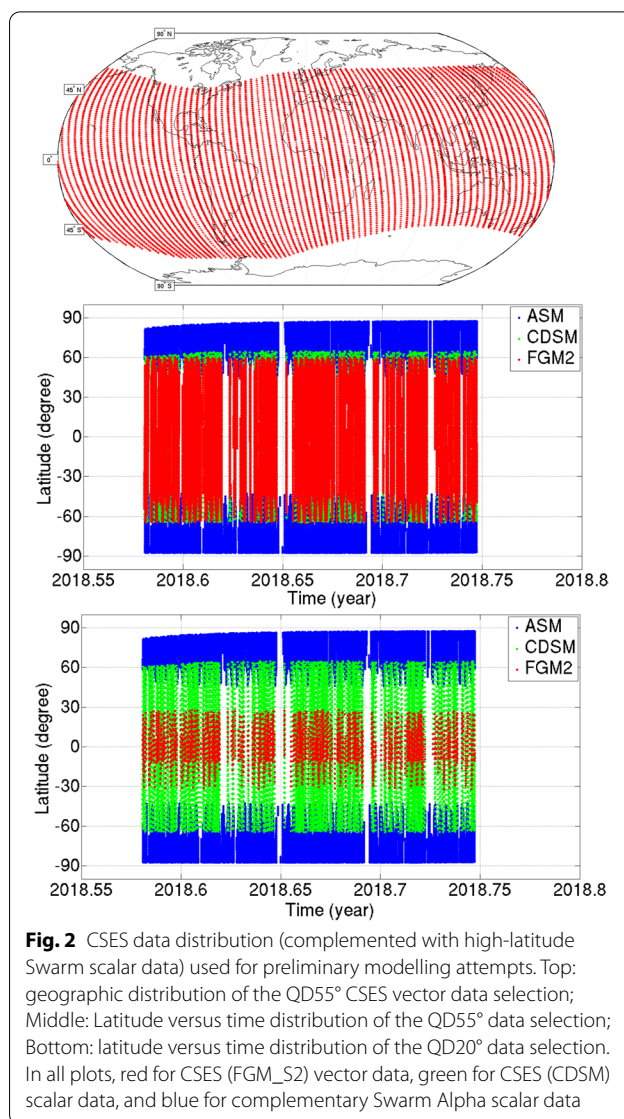
A priori data error standard deviations were set to 2.5 nT for both scalar and vector data in all cases (Swarm and CSES data). Attitude error was assumed isotropic (using the formalism of Holme and Bloxham (1996)). Different values were chosen for CSES (100 arcsecs) and Swarm (10 arcsecs), however. A much higher value was indeed required for CSES to account for the significantly lower quality of the mechanical link between the CSES STR reference frame and FGM reference frame (see below).

**Lessons learnt**

Four models were produced in total. Two were built using the Type 1 CSES HPM vector and scalar data from either the 55°QD or the 20°QD selection, complemented with high-latitude Swarm Alpha scalar (as described above). For brevity, we will refer to these as the 55°QD and 20°QD CSES models. Two additional Swarm reference models were otherwise built in the same way, using the 55°QD and 20°QD Swarm data selections (55°QD and 20°QD Swarm models). Figures 2 and 3 illustrate the corresponding data distributions. Comparing Figs. 2a, 3a (availability of vector data between -55° and +55° QD latitudes from, respectively, CSES and Swarm Alpha) reveals a significant difference between the CSES and Swarm data distributions.

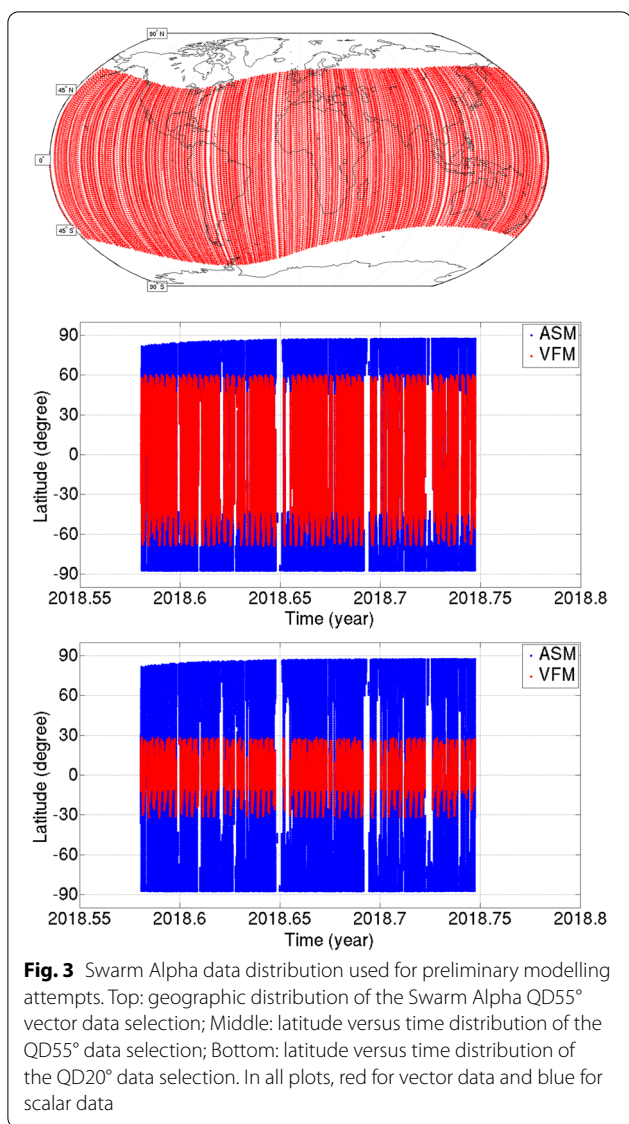
Whereas the Swarm Alpha orbit provides a nice global coverage of all longitudes over the 2 months considered, the 5-day revisiting period of CSES is responsible for a significantly poorer longitudinal distribution, leaving roughly 80 sectorial gaps. However, we note that by the Nyquist sampling criterion, 80 equally spaced bands in longitude should allow the recovery of sectorial dependence up to order 40. These gaps are thus expected to be narrow enough to only mildly affect the recovery of a global field model up to degree and order 15. Indeed, this does not turn out to be the most significant issue.

A much more significant issue is revealed by the comparison of the CSES and Swarm 55°QD models, as shown in Fig. 4. For SH degrees 1 to 4, the Lowes-Mauersberger spatial spectrum (Mauersberger 1956; Lowes 1966) of the differences between these two models at the Earth’s surface for central epoch of the models (September 1, 2018) is clearly much larger than that of the differences between the Swarm 55°QD model and the CHAOS-6- $\times$ 8 model for the same epoch. The latter spectrum provides a good indication of the limitation of using only 2 months of 55°QD selected data from a single satellite. Clearly, the CSES 55°QD model fails to properly determine the first four spherical harmonic degrees of



**Fig. 2** CSES data distribution (complemented with high-latitude Swarm scalar data) used for preliminary modelling attempts. Top: geographic distribution of the QD55° CSES vector data selection; Middle: Latitude versus time distribution of the QD55° data selection; Bottom: latitude versus time distribution of the QD20° data selection. In all plots, red for CSES (FGM\_S2) vector data, green for CSES (CDSM) scalar data, and blue for complementary Swarm Alpha scalar data

the field. Plotting the radial component of the difference between the predictions of the CSES and Swarm 55°QD models at the Earth’s surface (also shown in Fig. 4) makes it clear that this disagreement, reaching up to 70 nT at Earth’s surface, is mainly zonally distributed and not related to the sectorial gaps seen in Fig. 2. Its magnitude also makes it difficult to relate to differences in the magnetic field signals seen by the Swarm and CSES satellites, which share similar altitudes, or to some potentially poorly recovered secular variation, which cannot produce such differences between models built with only 2 months of data. Although one cannot exclude that this disagreement could be due to some other unidentified issue, the most likely possibility we identified is related to the mechanical link between the FGM\_S2 instrument (on



the last leg of the boom, see Fig. 1) and the STR (providing attitude information, but located on the body of the satellite). This link is prone to potential systematic deformation along the orbit. Recall, indeed, that our modelling procedure assumes this link to be strictly rigid throughout the 2 month period considered, whereas the design of the CSES HPM boom (three segments with three hinges) may not be capable of guaranteeing this.

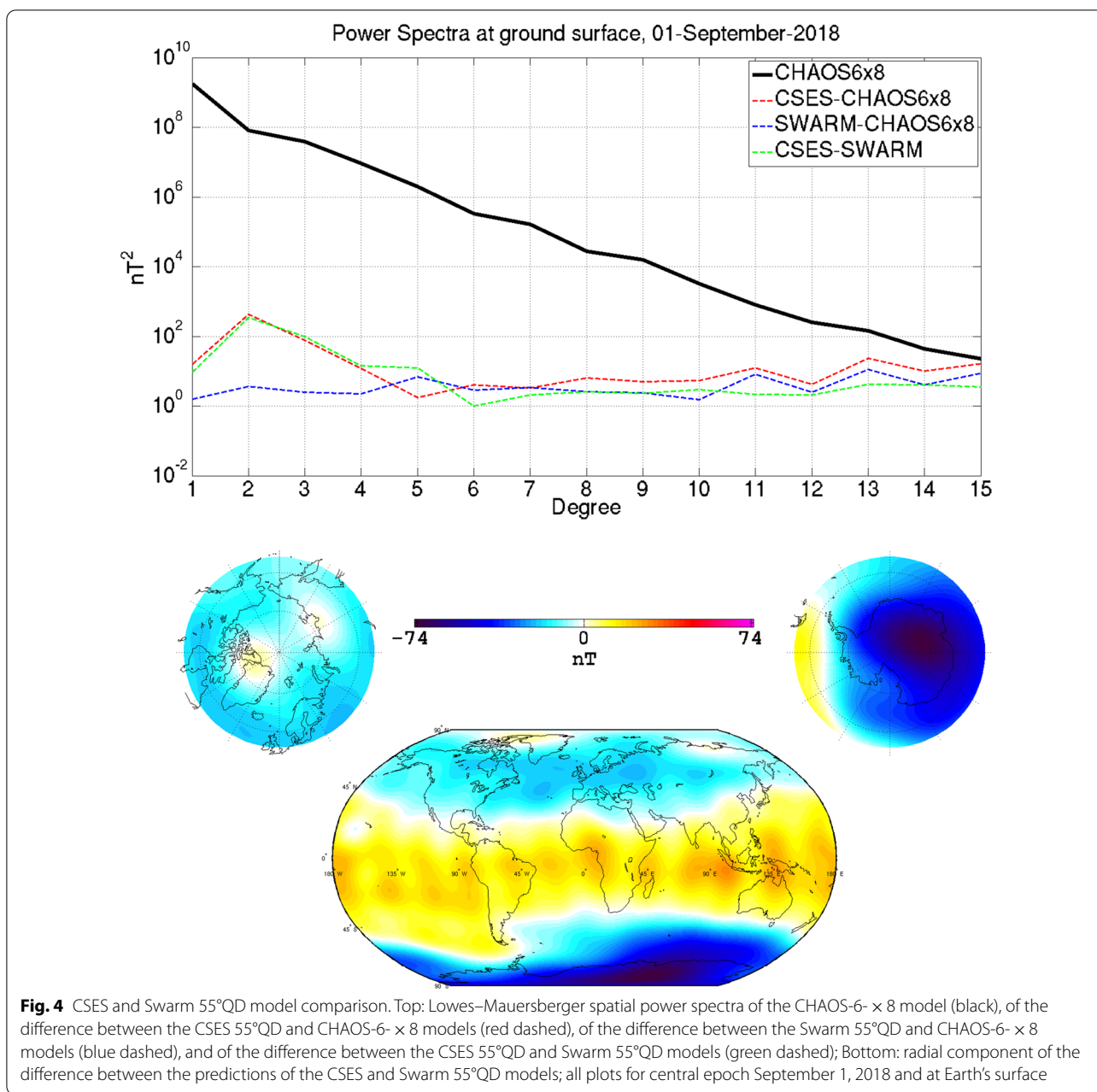
To check this possibility and attempt to improve the quality of the CSES model to be recovered, we relied on similar comparisons, now using the CSES and Swarm 20°QD models. These models are based on much less vector data, all concentrated in a 40°QD wide equatorial band along the magnetic equator. The hope was that the mechanical link (rotation matrix) between the FGM\_S2

and STR frames of reference would be stable enough along this equatorial part of the (night-side) orbit leg, and similar enough from one orbit to the next, to behave as if almost stiff. Ignoring all vector data was obviously not an option, since enough vector data close to the magnetic equator are mandatory, in particular to provide the knowledge of where this equator lies, a critical information (see Khokhlov et al. 1997, 1999) to avoid the recovered model being affected by the so-called Backus effect (Backus 1970, also known as the perpendicular effect, Lowes 1975). Figure 5, to be compared to Fig. 4, shows that this indeed brings improvement. The disagreement between the two CSES and Swarm 20°QD models for degrees 1 to 4 is much reduced. The reduced use of vector data comes at a slight cost, though, with a modest degradation of the recovery of the degree 5 SH component (see also the impact on the Swarm 20°QD model when compared to the CHAOS-6- × 8 model). Overall, nevertheless, the improvement is very substantial, as can also be seen in the map of the radial component of the difference between the predictions of the CSES and Swarm 20°QD models plotted at the Earth's surface (also shown in Fig. 5). Although the zonal effect is not entirely removed, it now leads to disagreements about three times less in magnitude, only reaching 25 nT at most at Earth's surface (note the difference in the colour scales used in Figs. 4, 5).

To further confirm that the issue in the CSES models is indeed likely linked to some deformation of the boom along the orbit, we finally computed the residuals between the CSES Type 1 vector data used and the predictions of the CHAOS-6- × 8 model (which includes both internal and magnetospheric source contributions, but not, e.g., in situ ionospheric currents crossed by the satellite). Should the CSES vector data be free of any slowly varying biases (such as produced by orbital boom deformation), these residuals would be expected to only reflect noise in the data and contributions of signals from sources not modelled by CHAOS-6- × 8. In contrast, if boom deformation occurs systematically along the orbit, significant signatures would be expected in the form of slowly varying biases as a function of latitude. Since it is known that no such effect is to be found on Swarm Alpha (see, e.g., Olsen et al. 2015; each Swarm satellite has its VFM rigidly linked to its set of STR on a specially designed optical bench), a simple way to check this is to plot the equivalent residuals between the Swarm Alpha vector data and predictions of the CHAOS-6- × 8 model. Both satellites orbiting at the same local time over the time period considered (therefore sensing similar unmodelled sources), the latter residuals are expected to provide a relevant baseline.

Residuals were computed in both the NEC and instrument frames of reference, taking advantage of the Euler

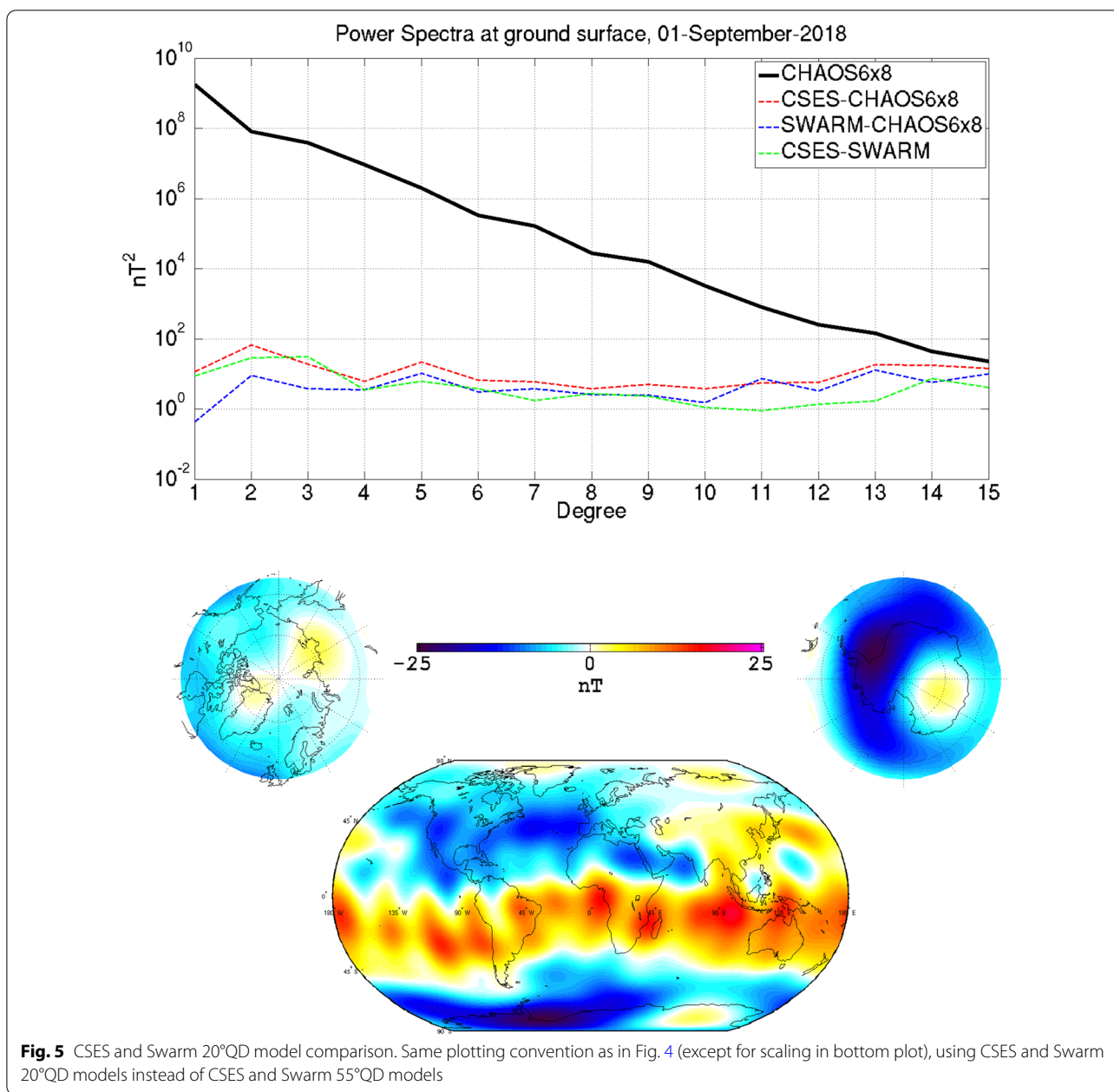




angles computed in the course of producing the 55°QD CSES and Swarm models to convert vector components from one frame to the other. Residuals in the NEC frame were computed using the Euler angles and quaternion information to rotate the vector data from the instruments frame to the NEC frame, before subtracting the predictions of the CHAOS-6- × 8 model (Fig. 6). Residuals in the instruments frame were computed using the quaternion information and Euler angles to rotate the predictions of the CHAOS-6- × 8 model before subtracting these from the vector data (Fig. 7).

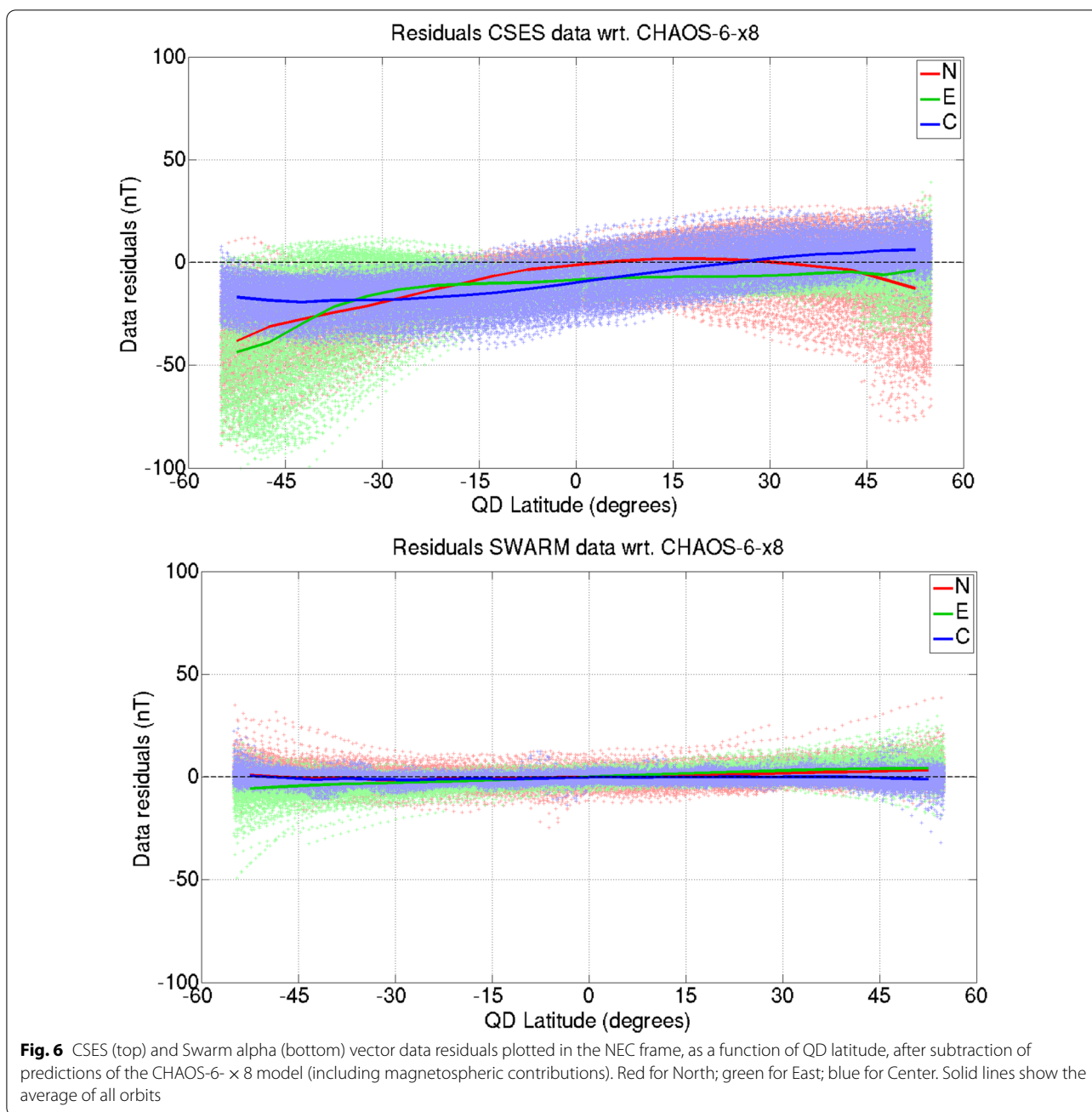
As can be seen, no significant bias can be found in the Swarm Alpha residuals, which also display a dispersion of the type expected for Swarm, for the quiet nighttime selection used in this study (see, e.g., Olsen et al. 2015). In contrast, strong varying biases can be found in the CSES residuals. These biases are strongest in the high southern latitudes, progressively decrease towards the equator, and are much less marked in the northern hemisphere. This North–South asymmetry, we note, is consistent with a similar asymmetry in the disagreements between the CSES and Swarm models (stronger in the





Southern hemisphere than in the Northern hemisphere, recall Figs. 4, 5). Since CSES orbits at a fixed 14h00 LT at descending node, this evolution follows the path of the satellite on its night leg of the orbit, from South to North. It shows that the bias is maximum every time CSES moves away from the Sun at the end of the day-side orbit leg during which the boom has been presumably heated, than starts decreasing as the satellite begins its journey northwards in the dark, allowing the boom to progressively cool down. This thus strongly suggests that the bias signature is indeed related to some thermal

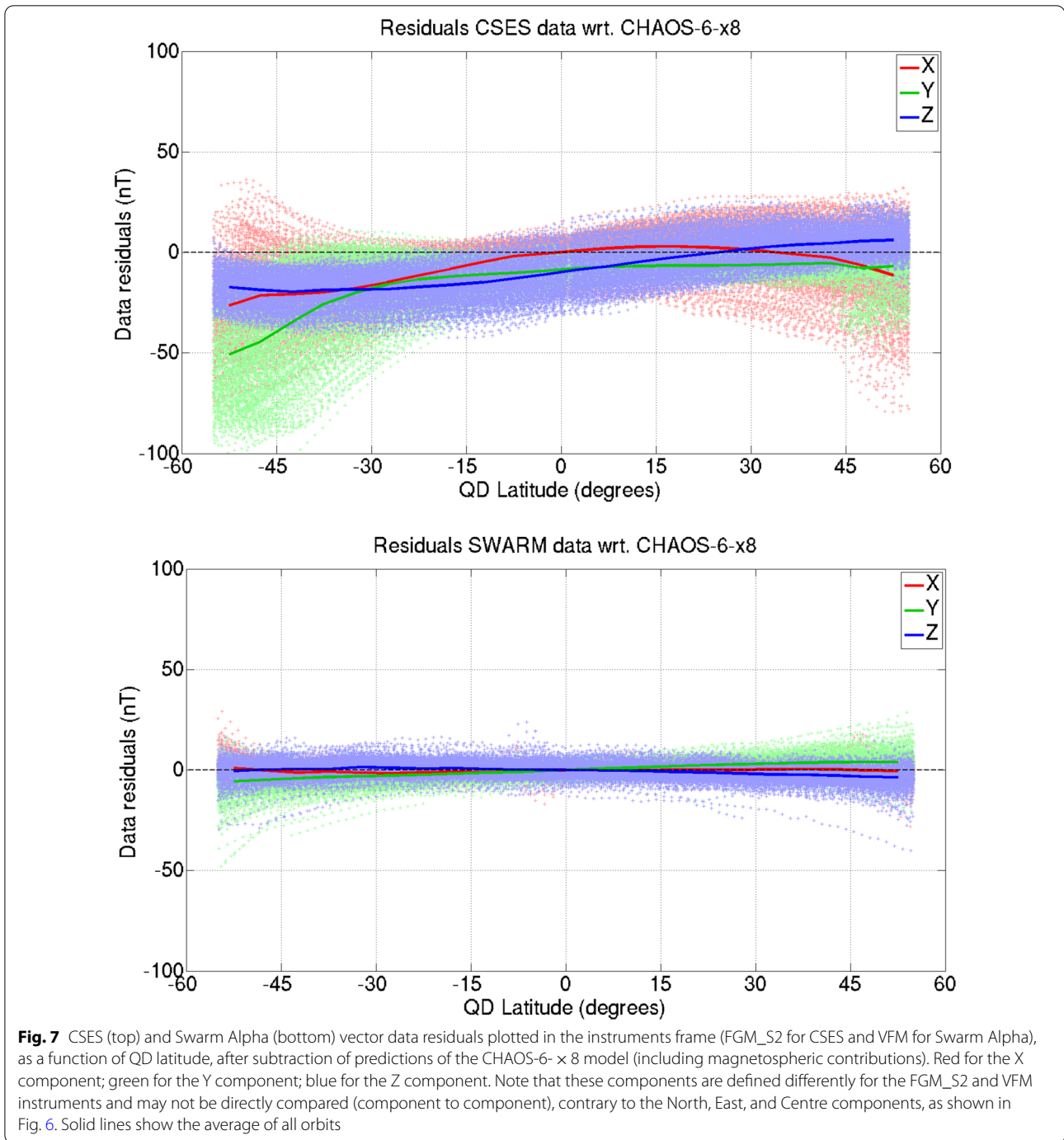
boom deformation, which builds up on the dayside leg of the orbit, then thermally relaxes on the night-side leg, settling back to a roughly stable state by the time the satellite reaches the equator on this night side. This evolution also shows that the most problematic CSES vector data are those from the southernmost part of the (night-side) orbit. These data being dismissed in the 20°QD data selection, it naturally explains why the 20°QD CSES model appears to be of much better quality than its 55°QD equivalent.



Last but not least, Figs. 6, 7 also clearly show that the dispersion in the CSES residuals is much larger than that in the Swarm Alpha residuals. It is highly doubtful that this could be the result of different natural unmodelled signals seen by the two satellites. The intrinsic noise level affecting the FGM\_S2 measurements (due to the instrument, the satellite and the rest of the payload) having been shown to be roughly comparable to that affecting the Swarm Alpha VFM instrument (Zhou et al. 2019), this, we practically, attributed to the impact of the

not-so-stiff boom and possibly also errors in the attitude restitution provided by the STR through the quaternions (though independent checks of these STR data, not reported here, suggest that this source of error is much less significant, except possibly on some specific days, see below). This noise level is the reason we assumed a fairly large error of 100 arcsecs for the attitude when computing CSES models.

A number of important lessons were thus learnt from the above preliminary modelling attempts. One is that



the a priori unfavourable 5 days recursive period of CSES, which introduces longitudinal gaps in the data distribution (see Fig. 2), does not appear to be critical for IGRF modelling purposes. Another one, unfortunately much more critical, is that the mechanical link between the FGM (on the last leg of the three hinges boom) and the STR (on the body of the satellite) appears to be

problematic. The boom seems to suffer from systematic thermal deformations along the orbit of CSES, which affect the recovery of the attitude of the vector data provided by the FGM. This deformation could be roughly characterized, and the issue appears to mainly affect data from the southernmost part of the night-side leg of the CSES orbits needed for IGRF modelling purposes.

Nevertheless, a simple workaround could be found, which consisted in selecting vector data only within a  $40^\circ\text{QD}$  band centred on the magnetic equator (the  $20^\circ\text{QD}$  selection), and assuming an attitude error of 100 arcsecs in the inversion procedure. The timeline imposed by the IGRF deadline of October 1, 2019 did not allow us to test more advanced strategies, and this is the strategy we therefore used to also produce the CGGM parent model as described below. One significant change we made, however, is that we decided not to use the vector data provided by the FGM\_S2 instrument, in favour of the vector data provided by the FGM\_S1 instrument. This choice was justified by the fact that this instrument being closer to the satellite (recall Fig. 1), boom deformation can be expected to be slightly attenuated, with the potential drawback of having slightly noisier data (because of the smaller distance to the satellite) being minor, since such noise level has not been identified as the limiting factor.

### CGGM parent model and IGRF 2020 candidate model construction

We now move to the description of the way the CGGM parent model was built and next extrapolated in time to build the CGGM IGRF 2020 candidate model, taking advantage of the lessons learnt during our early modelling attempts, and of a much-increased amount of data. The CGGM parent model covers a longer time period and uses all CSES data available before the October 1, 2019 deadline. It also only uses CSES data, Type 2 scalar data covering high latitudes having been made available on time by CSES team for this purpose, to avoid having to rely on any Swarm (or other satellite) data, in contrast to what had been done for the previous preliminary modelling attempts. It finally uses a slightly more advanced data selection and modelling strategy (closer to that used by Hulot et al. 2015a) to reach the quality needed to next extract an IGRF 2020 candidate model meeting the requirements of the call.

#### Data selection

##### Temporal coverage

Data used (both Type 1 and Type 2) now cover almost 19 months, between March 03, 2018 and September 20, 2019.

##### Geographic coverage

1 Hz scalar data (both Type 1 and Type 2) were taken from the CDSM instrument without any geographic restriction. 1 Hz vector data (only Type 1) were taken from the FGM\_S1 instrument (expressed in the instrument's reference frame) selected according to the  $20^\circ\text{QD}$

selection, i.e., only within the  $-20^\circ\text{QD}$  to  $20^\circ\text{QD}$  equatorial band.

#### Selection criteria common to both scalar and vector data

**Quality check:** Removal of data not satisfying the criteria that differences between each datum and the prediction from the CHAOS-6- $\times$ 9 model (latest version of the CHAOS-6 model of Finlay et al. (2016), which had then been made available) should be less than 100 nT (scalar or norm comparison for vector data). This more stringent criterion was found to be better suited to remove the occasional blatant outliers who are slightly more numerous within the 100 nT–300 nT range in the Type 2 data (which were not used in the preliminary study).

**Night-time selection:** Sun angle seen by the satellite required to be at least  $10^\circ$  below horizon for night-time selection (rather than LT selection, as was done for the preliminary models). This ensures better night-time selection and was found to be compatible with the available Type 2 data.

**Magnetically quiet conditions:** based on  $|d\text{RC}/dt| < 2\text{nT/h}$  and  $K_p < 2+$ .

#### Additional selection criteria for scalar data

A more stringent  $E_m < 0.8$  mV/m criterion than for the preliminary models was required for high-latitude scalar data. This was again found to be compatible with the available Type 2 data. A dedicated Flag signalling when magnetotorquers were activated on CSES was provided with the data and used to avoid data at times of magnetotorquer activation for all Type 1 data (Flag MT should be 0). This flag was not used for Type 2 data, as magnetotorquers are activated most of the time at high latitudes (as a result of the operating mode of the satellite). These perturbations, however, remain within 20 nT for these Type 2 data. Finally, decimation was applied to Type 1 data to avoid over-representation along tracks, but not to the much scarcer Type 2 data. That led to scalar data (for both Types 1 and 2 data) typically separated by 1 min.

#### Additional selection for vector data:

Scalar residuals (difference between scalar provided by CDSM and modulus of vector provided by FGM\_S1) were required to be less than 2.5 nT. In addition, 17 days of problematic vector data were discarded: 15 days in 2018 (May 4, 8, 12, 14, 18, 20, 27, 29–31; June 5, 12–14; September 24) and 2 days in 2019 (March 3, September 20). Given the selection criteria previously applied to the data, the issue during these days is most likely due to temporary problems with attitude restitution, leading these data to be incompatible with the rest of the dataset (recall, indeed, that we do not apply any specific selection criteria on STR data). Finally, decimation was also



applied (now keeping 1 point out every 15 s) to again avoid over-representation along track.

**Total amount of data selected:**

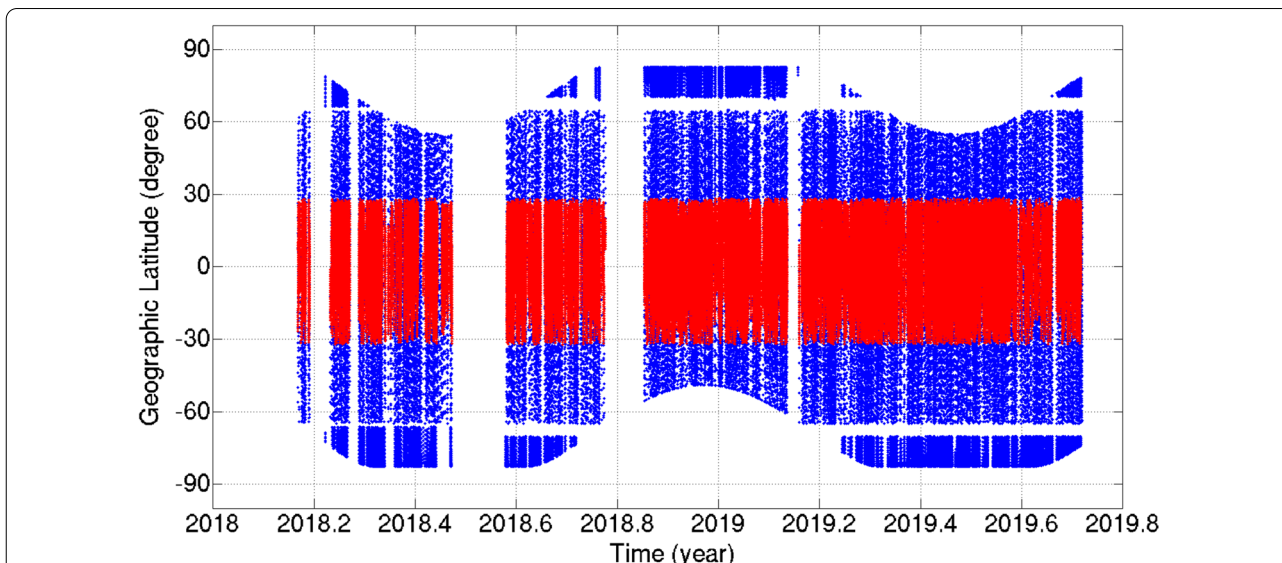
Overall, this selection procedure resulted in the selection of 92,068 scalar data (among which 62,715 data at absolute geographic latitudes higher than 65°) and 122,867 × 3 vector data, distributed in time and latitude, as illustrated in Fig. 8.

**CGGM parent model parameterization and optimization**

The model parameterization we chose to build the CSES parent model with is more sophisticated than the one used for the preliminary modelling attempts, given the longer time period to be modelled. It now is closer to that used by Hulot et al. (2015a). The main field is still modelled up to SH degree and order 15, but the linear SV is now modelled up to degree and order 8. As before (and again referring to the notation of Olsen et al. 2014), the remote magnetospheric sources (and their induced counterparts) are described by a zonal external field up to degree 2 in GSM coordinates (2 coefficients), while the near magnetospheric ring current (and its induced counterpart) is described by an external field up to SH degree and order 2 in SM coordinates. The latter, however, is now modelled in a more advanced way. SH degree 2 coefficients are still assumed static (5 coefficients:  $q_2^0, q_2^1, q_2^2, s_2^1, s_2^2$ ). SH degree 1 coefficients are still described by a fast time-varying part proportional to the RC index for degree 1 (3 coefficients:  $\hat{q}_1^0, \hat{q}_1^1, \hat{s}_1^1$ ), but their baselines are

no longer assumed static. For the zonal coefficient,  $\Delta q_1^0$  is now allowed to change every 5 days (since 98 time segments of 5 days are involved, this implies solving for 98 different coefficients), while for the sectorial coefficients,  $\Delta q_1^1$  and  $\Delta s_1^1$  are now allowed to change every 30 days (implying solving for  $2 \times 19 = 38$  different coefficients). Finally, Euler angles are now also allowed to change every 10 days, to account for possible long-term deformation of the mechanical link between the FGM\_S1 instruments and the STR (implying solving for  $3 \times 53 = 159$  different coefficients). In total, 640 parameters were thus solved for, 255 for the static Gauss coefficients, 80 for the linear SV, 146 parameters for the external field, and 159 for the Euler angles.

For solving the inverse problem, we relied on the same iteratively reweighted least-squares algorithm with Huber weights as for the preliminary models (again with no regularization, and using the same geographical weight). A priori data error standard deviations were slightly reduced to 2.2 nT for both scalar and vector data. Attitude error was again set to 100 arcsecs. For completeness, we also specify that CHAOS-4 (Olsen et al. 2014) up to degree and order 13 for epoch 01/03/18 was used as a (static) starting model for the iterative computation. This choice was made to ensure faster convergence of the iterative computation than just starting from a simple dipole field. It has been shown to have very little influence on the final model (see, e.g., Vigneron et al. 2015). Full convergence of the computation was then reached after eight iterations. Resulting residual statistics (using



**Fig. 8** Latitude versus time distribution of the selected CSES data used for building the CGGM parent model (red: FGM\_S1 vector data; blue: CDSM scalar data), note the gaps around 65°N and 65°S due to unavailability of CDSM data in this transition from Type 1 to Type 2 data available at the time of modelling

the same conventions as in Hulot et al. 2015a) are provided in Table 1.

### CGGM IGRF 2020 candidate model generation

The CGGM parent model provides a spherical harmonic estimate of the main field up to degree and order 15 for central epoch December 11, 2018, together with a spherical harmonic estimate of the average secular variation over the time covered by the data (March 2018–September 2019) up to degree and order 8.

The CGGM IGRF 2020 candidate model was simply extrapolated in time from the CGGM parent model up to degree and order 13, using the central epoch December 11, 2018 as initial point and the SV coefficients up to degree and order 8 to extrapolate the model to epoch January 1, 2020. No temporal extrapolation for spherical harmonic with degrees 9–13 was used (which were thus assumed identical to that inferred by the CGGM parent model for central epoch December 11, 2018). Although this will have undoubtedly introduced some additional source of error in the CGGM candidate model, this choice was made to keep with our original goal of building an IGRF candidate model entirely, and only, based on CSES data.

### Initial quality assessment

To validate the CGGM candidate model, we relied on some comparison of the predictions of the CGGM parent model with those of the CHAOS-6- $\times$ 9 model. This CHAOS-6- $\times$ 9 model was computed by DTU only using L1b Swarm data (plus data from earlier missions as well as data from ground observatories) and is therefore independent from the CGGM parent model (except, strictly speaking, for the very minor fact that data used for producing the CGGM parent model were first checked against the this CHAOS-6- $\times$ 9 model for rejecting very occasional extreme outliers). However, since it only uses data up to April 2019, a comparison of predictions for epoch 2020.0 was not considered appropriate. In contrast, CHAOS-6- $\times$ 9 could be considered to provide a very reliable estimate of the main field for two epochs of interest, December 11, 2018, which corresponds to the central time of the CGGM parent model, and November 20, 2017, which is 103 days before the very first data used to build the CGGM parent model. This is the same amount of time separating the last data used in the CGGM parent model and epoch 2020.0. Given the symmetry of the CSES data distribution we used (recall Fig. 8), we considered this backward extrapolation test as a good way to assess how well our CGGM IGRF 2020.0 candidate model could be expected to perform.

Figure 9 illustrates the differences between the CGGM parent and CHAOS-6- $\times$ 9 models at Earth’s surface, for

central epoch December 11, 2018. The radial component  $B_r$  of the difference between the two models reveals a mainly zonal signature, with amplitudes reaching 22nT. These differences are reminiscent of those we had found in our early modelling attempts when comparing CSES based and Swarm Alpha-based models, but appear to be slightly weaker (recall Fig. 5), despite the fact that we now also only use CSES (Type 2) scalar data at high latitudes. The spectral difference between the two models is also very similar, but again slightly weaker. These differences most likely reflect the issue we previously identified with CSES (and attributed to systematic boom deformation along the orbits), which our improved modelling strategy slightly better mitigates.

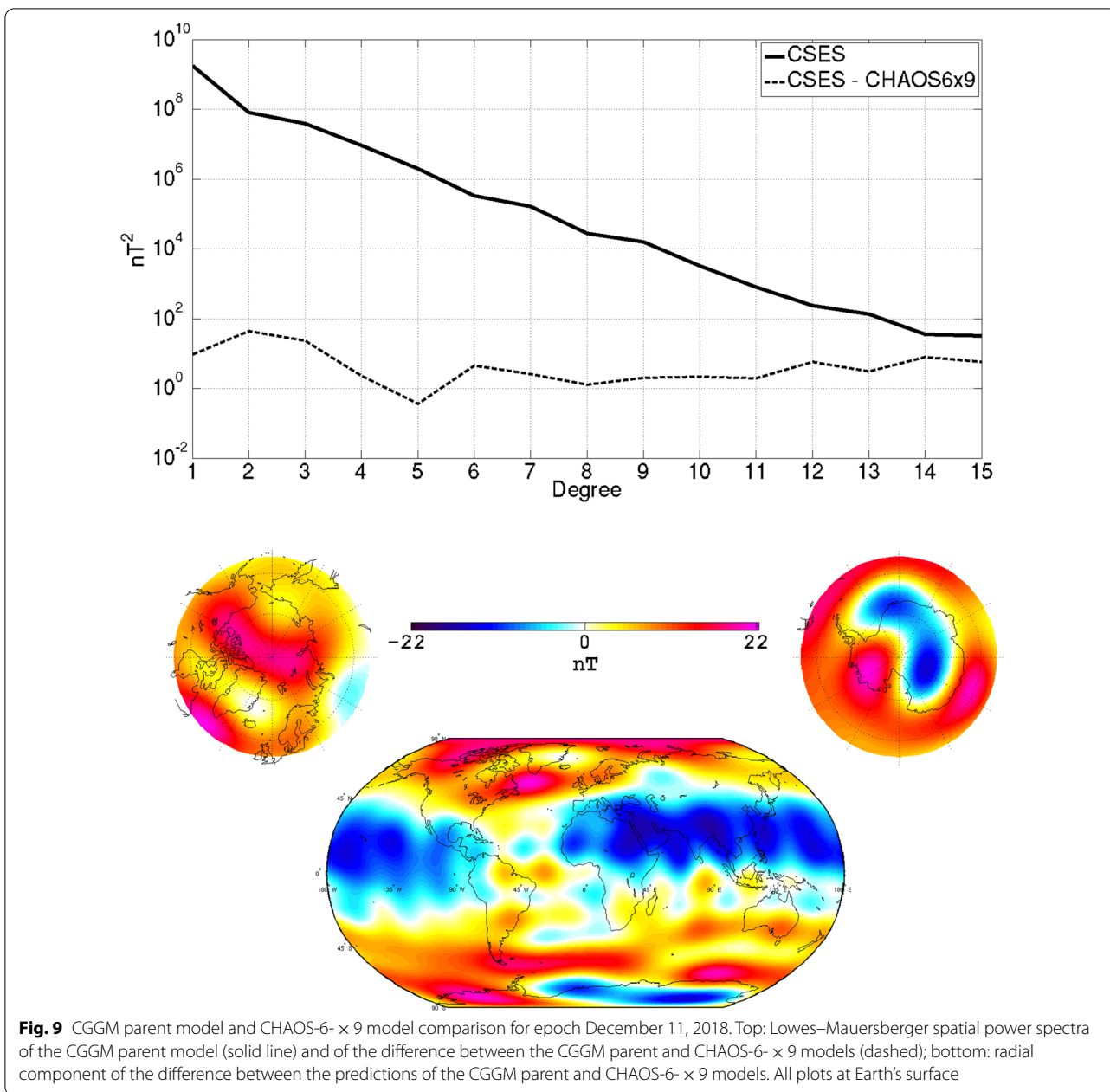
Figure 10 illustrates the same differences, but for the more relevant backward extrapolation to epoch November 20, 2017, reflecting the errors likely affecting the CGGM candidate model for epoch January 1, 2020. As expected, errors in the radial component are still mainly zonal, but peaking at 37 nT. Spectral differences are the largest for the three first degrees. They reach 20 nT<sup>2</sup> at degree 1, 50 nT<sup>2</sup> at degree 2, and 30 nT<sup>2</sup> at degree 3 while remaining below 10 nT<sup>2</sup> at all higher degrees, except for degree 9, which reaches 20 nT<sup>2</sup>. These differences are quite comparable (though more on the high side) to differences observed between the various IGRF 2015 candidate models that were submitted in 2015 (at a similar stage of IGRF model preparation), as can be checked by comparing Fig. 10 with Fig. 7 of Thébaud et al. (2015b).

These encouraging comparisons led us to conclude that despite the limitations of the current quality of CSES vector data (limited by the boom deformation issue), and CSES scalar data at high-latitude data (see corresponding residual statistics in Table 1), the CGGM candidate model could, indeed, be proposed as an IGRF 2020 candidate model.

### Computation of uncertainties for each Gauss coefficient

Realistic uncertainties affecting the Gauss coefficients of the CGGM candidate model were informally requested in addition to the coefficients of the model for submission to the IGRF call. These uncertainties were computed by again assuming that the observed disagreements between the CSES parent model backward extrapolated to epoch November 20, 2017 and the CHAOS6- $\times$ 9 model computed at the same epoch, are representative of the uncertainties affecting the coefficients of the CGGM candidate model. For each degree  $n$ , we computed the following root-mean-square quantity:

$$\sigma_n = \sqrt{(2n + 1)^{-1} \sum_{m=0}^n ((dg_n^m)^2 + (dh_n^m)^2)}, \quad (1)$$



**Fig. 9** CGGM parent model and CHAOS-6- × 9 model comparison for epoch December 11, 2018. Top: Lowes–Mauersberger spatial power spectra of the CGGM parent model (solid line) and of the difference between the CGGM parent and CHAOS-6- × 9 models (dashed); bottom: radial component of the difference between the predictions of the CGGM parent and CHAOS-6- × 9 models. All plots at Earth’s surface

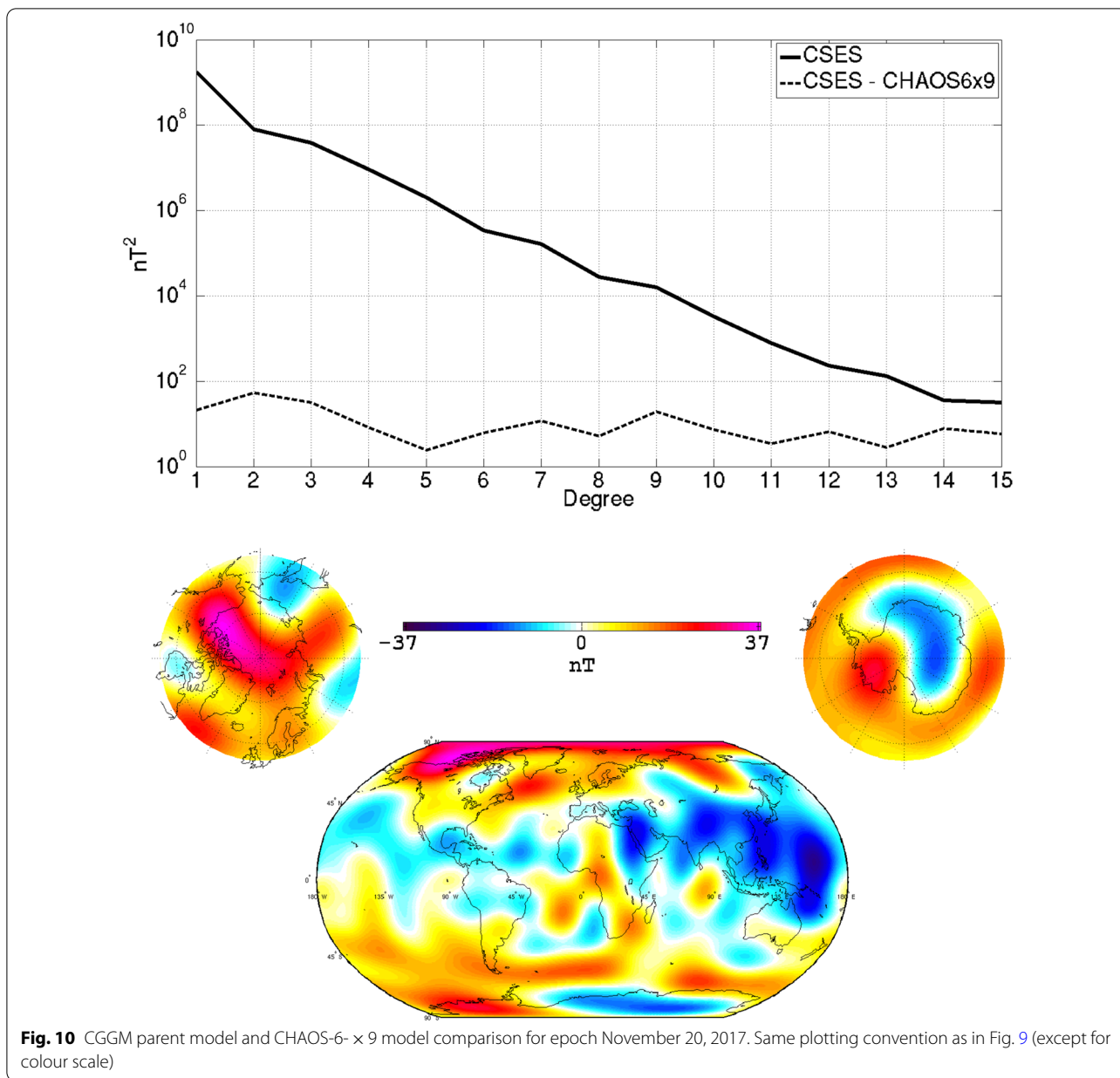
where  $dg_n^m$  and  $dh_n^m$  are the differences in the  $g_n^m$  and  $h_n^m$  Gauss coefficients between the two models. We then simply assigned this  $\sigma_n$  as our best estimate of the errors (one sigma type) affecting each Gauss coefficient of degree  $n$ .

This quantity should only be considered as a rough indicator. In particular, it likely underestimates uncertainties affecting zonal coefficients (i.e.,  $g_n^0$  Gauss

coefficients), by probably a factor 2 (at least for degrees 1–3, recall Fig. 10; see also Lowes and Olsen, 2004).

**A posteriori quality assessment and conclusion**

All candidate models provided in response to the IGRF-13 call having been made available after the October 1, 2019 deadline, and the final IGRF-13 series of models having since been released, we finally looked into the way



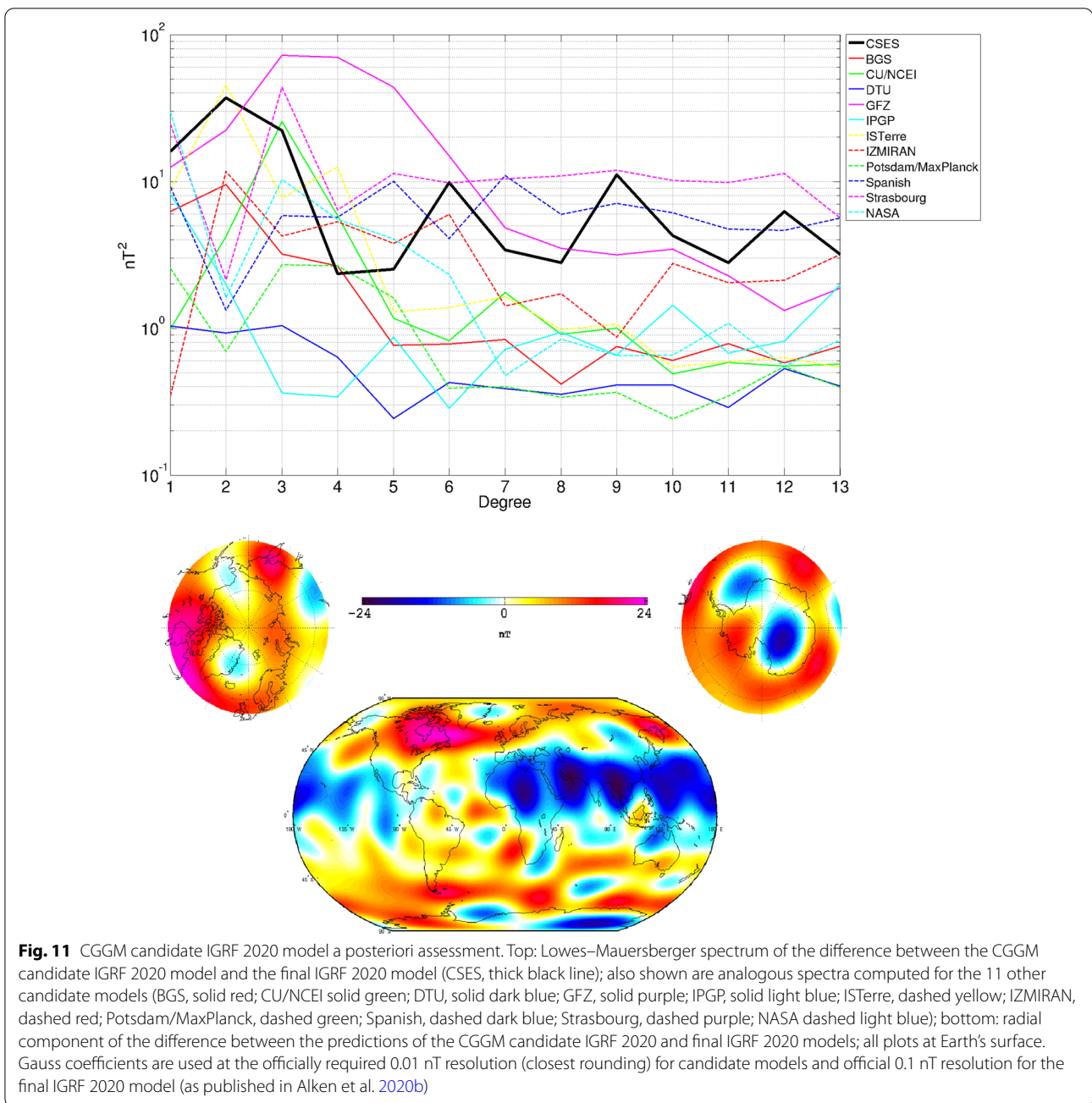
**Fig. 10** CGGM parent model and CHAOS-6- x 9 model comparison for epoch November 20, 2017. Same plotting convention as in Fig. 9 (except for colour scale)

the CGGM candidate model compares with these models. Eleven IGRF 2020 candidate models were submitted, in addition to the CGGM candidate model, and all 12 models have been used to produce the final IGRF 2020 model, which thus is a model combining all candidate models (Alken et al. 2020b). The detailed way this was done can be found in Alken et al. (2020a), and the way each candidate model was prepared can be found in a series of papers to also be found in the present issue. We here refer to these models as the BGS model (Brown et al. 2020), the CU/NCEI model (Alken et al. 2020c), the DTU model (Finlay et al. 2020), the GFZ model (Rother et al.

2020), the IPGP model (Ropp et al. 2020), the ISTerre model (Huder et al. 2020), the IZMIRAN model (Petrov and Bondar 2020), the Postdam/MaxPlanck model (Bärenzung et al. 2020), the Spanish model (Pavón Carrasco et al. 2020), the Strasbourg model (Wardinski et al. 2020), and the NASA model (Sabaka et al. 2020).

A Lowes–Mauersberger spatial power spectrum of the difference between the CGGM candidate IGRF 2020 model and the now released official IGRF 2020 model, as well as analogous spectra for all other candidate models, are shown in Fig. 11 (top), together with a map of the radial component of the difference between





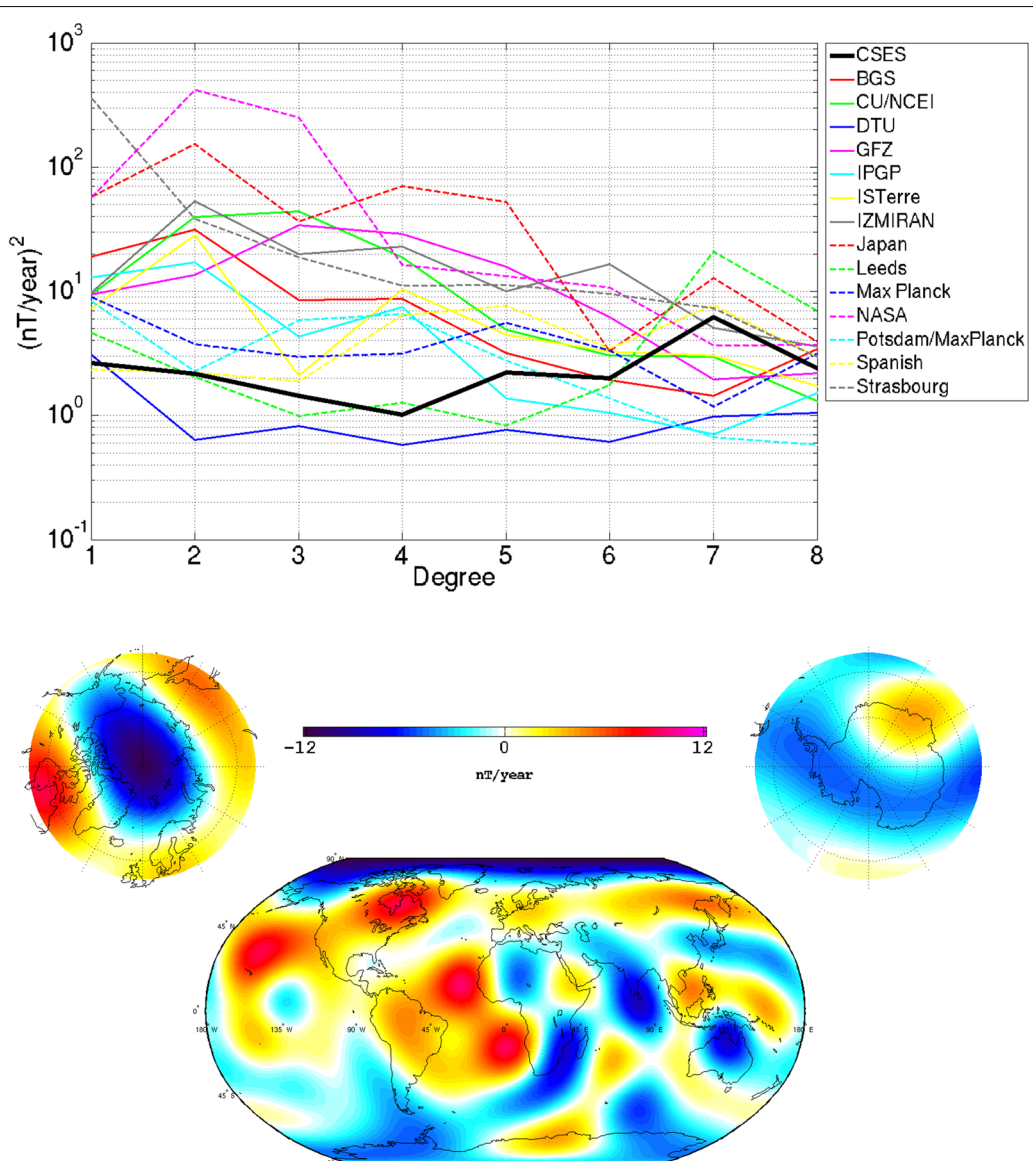
the predictions of the CGGM candidate and official IGRF 2020 models (bottom). Comparing this map with those shown in Figs. 9 and 10 reveals that these differences are much closer to the differences found when comparing the CGGM parent model to the CHAOS-6- × 9 model at central epoch December 11, 2018 (Fig. 9) than to those found when carrying the same comparison for the backward extrapolation to epoch November 20, 2017 (Fig. 10). This shows that the CGGM candidate model does much better than anticipated, despite all the issues identified

in the CSES data. The same conclusion holds when comparing Lowes–Mauersberger spectra. In particular, we note that the spectral comparison of the CGGM candidate model with the official IGRF 2020 model always lies within the envelope of the analogous spectral comparison for all the other IGRF 2020 candidate models.

The encouraging ability of the CGGM candidate to perform better than anticipated finally led us to also test how well the secular variation associated with the CGGM parent model (referred to as the CGGM SV model in what

follows) would have performed, had it been submitted as an IGRF-13 2020–2025 predictive SV candidate model. To test this, similar a posteriori comparisons were performed with the final IGRF-13 2020–2025 predictive SV model, and with the various IGRF-13 2020–2025 predictive SV candidate models that were used to build it, and which were produced by either the same teams as

the IGRF 2020 candidate models (BGS, CU/NCEI, DTU, GFZ, ISTERre, IZMIRAN, NASA, Potsdam/MaxPlanck, Spanish, and Strasbourg models), by other teams led by the same institutions (IPGP model, Fournier et al. 2020) or by teams led by other institutions. These, we refer to as the Japan model (Minami et al. 2020), the Leeds model (Metman et al. 2020), and the Max Planck model



**Fig. 12** Comparing the CGGM SV and final IGRF-13 2020–2025 predictive SV models. Top: Lowes–Mauersberger spectrum of the difference between the CGGM SV and the final IGRF-13 predictive SV models (CSES, thick black line); also shown are analogous spectra computed for the fourteen candidate SV models (BGS, solid red; CU/NCEI solid green; DTU, solid dark blue; GFZ, solid purple; IPGP, solid light blue; ISTERre, solid yellow; IZMIRAN, solid grey; Japan, red dashed; Leeds, dashed green; Max Planck, dashed dark blue; NASA, dashed purple; Potsdam/MaxPlanck, dashed light blue; Spanish, dashed yellow; Strasbourg, dashed grey); bottom: radial component of the difference between the predictions of the CGGM SV model and the final IGRF-13 2020–2025 predictive SV model; All plots at Earth’s surface. Gauss coefficients are used at the officially required 0.01 nT/yr resolution (closest rounding) for candidate models, and official resolution of 0.1 nT/yr for the final IGRF-13 predictive SV model (as published in Alken et al. 2020b)

**Table 1 Residual statistics for all data used to produce the CGGM parent model (using the same convention as in Hulot et al. 2015a)**

Residual statistics of the CGGM parent model			
	N	Mean (nT)	RMS (nT)
$F_p$	30246	-0.02	6.20
$F_{NP} + B_B$	184689	+0.29	2.60
$F + B_B$	214935	+0.26	3.16
$B_B$	122867	+0.47	2.67
$B_r$	122867	+0.06	8.08
$B_\theta$	122867	-0.70	4.84
$B_\varphi$	122867	+0.05	5.57

For each type of data, N is the number of data used, while Mean and RMS are the Huber-weighted misfit mean and Root-Mean-Square values (in nT);  $F_p$  refers to the misfit of the scalar data above (absolute) QD latitude  $55^\circ$  (polar latitudes),  $F_{NP}$  to the misfit of the scalar data below (absolute) QD latitude  $55^\circ$  (non-polar latitudes),  $F$  to the misfit of all scalar data,  $B_B$  to the misfit of the field component projected along the field direction (providing a measure of the misfit of the modulus of the vector data with respect to model prediction), and  $B_r$ ,  $B_\theta$ , and  $B_\varphi$  refer to the three geocentric vector field components. Note that vector residuals provided here are reconstructed residuals propagated from the vector residuals minimized in the reference frame of the instrument. Recall that vector data are only used for (absolute) QD latitude below  $20^\circ$

(Sanchez et al. 2020). A Lowes–Mauersberger spectral representation of how the CGGM SV model, and all these candidate models, compared to the official IGRF-13 2020–2025 predictive SV model is shown in Fig. 12 (top), together with a map of the radial component of the difference between the predictions of the CGGM SV and IGRF-13 2020–2025 predictive SV models (bottom). This comparison reveals that the agreement between the CGGM SV and IGRF-13 2020–2025 predictive SV models is now even better, and that the CGGM SV performs among the best when compared to the fourteen submitted 2020–2025 predictive SV candidate models. This is all the more remarkable that the CGGM model is based on a CSES data set completely independent from those used by all other candidate models (which all rely on Swarm data, sometimes also on other data, from, e.g., observatories), and that these CSES data still suffer from a number of issues (lack of boom rigidity for the Type 1 vector data, low quality of the high-latitude Type 2 scalar data).

The above a posteriori comparisons finally lead us to two very encouraging conclusions. One is that in principle, and despite their current limitations, CSES magnetic data can already be used to produce useful IGRF 2020 and 2020–2025 secular variation candidate models to contribute to the official IGRF-13 2020 and predictive secular variation models for the coming 2020–2025 time period.

The other is that now that the main issues affecting the CSES magnetic data have been identified, further improving the quality of its HPM data and making Type

2 scalar data systematically available would undoubtedly be worth the effort. This is quite a challenge. In particular, it would require developing an appropriate description of the way the boom deforms along the orbit of the satellite. However, the systematic nature of this deformation, most likely due to the fixed LT of the orbits of CSES (thus always exposed to the Sun in the same way along the orbits), could be taken advantage of. As suggested by the present study, such improved data could then very usefully contribute to the long-term monitoring of the main field and possibly other magnetic field sources, in complement to the data provided by missions such as the ESA Swarm mission.

#### Abbreviations

BGS: British Geological Survey; CDSM: Coupled dark state magnetometer; CGGM: CSES global geomagnetic field model; CSES: China Seismo-Electromagnetic Satellite; CU/NCEI: University of Colorado/National Centers for Environmental Information; DTU: Technical University of Denmark; EFD: Electric field detector; ESA: European space agency; FGM: Fluxgate magnetometer; GFZ: German Research Centre for Geosciences, Potsdam; GNSS-RO: GNSS occultation receiver; GSM: Geocentric solar magnetospheric; HEPD: High energetic particle detector; HEPP: High energetic particle package; HPM: High precision magnetometer; ICRF: Inertial celestial reference frame; IAGA: International Association of Geomagnetism and Aeronomy; IGRF: International geomagnetic reference field; IPGP: Institut de Physique du Globe de Paris; ISTerre: Institut des Sciences de la Terre; IZMIRAN: Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation; LAP: Langmuir probe; LT: Local time; MT: Magnetotorquer; NASA: National Aeronautics and Space Administration; NEC: North-East-Center; QD: Quasi-dipole; PAP: Plasma analyzer; SCM: Search-coil magnetometer; SH: Spherical harmonic; SM: Solar magnetic; STR: Star imagers; SV: Secular variation; TBB: Tri band beacon; VFM: Vector Fluxgate magnetometer on Swarm.

#### Acknowledgements

This work made use of the data from CSES mission, a project funded by the China National Space Administration (CNSA) and the China Earthquake Administration (CEA). The authors would like to thank Ciaran Beggan and an anonymous reviewer for their very useful comments that helped to improve this manuscript.

#### Authors' contributions

YYY prepared all HPM data, carried out initial model computations with PV, participated in all steps of the study, and contributed to the manuscript; GH designed the study, wrote the manuscript, and participated in all steps of the study; PV carried out all model computations discussed in the manuscript and participated in all steps of the study; XHS, the chief scientist of CSES, coordinated the cooperation of the research group and followed the entire study; ZMZR coordinated the needed HPM level 0 data from CRESDA and followed the entire study. BZ and WM led the design of the HPM instruments and its main processing chain; NO and LTC contributed to the initial data quality analysis and assisted in the modelling strategy choices; JPH and FG led the production of HPM data, with the help of JBY and JW; XMZ and LWW contributed to the development of CSES application center and determination of HPM level 2 products; SGY, JPD, and XHZ designed the satellite platform and helped in the analysis of star imager data; BJC carried out the performance verification of the HPM; AP and RL had the lead for the CDSM instrument development and carried out the performance verification of the CDSM data pre- and post-launch; JL contributed to the analysis of HPM Level 0 data; YYW and XDZ contributed to the data analysis. All authors read and approved the final manuscript.

#### Funding

This work was supported by the National Key R&D Program of China (Grant No. 2018YFC1503500), National Natural Science Foundation of China (Grant

No. 41904134), APSCO Earthquake Research Project Phase II and ISSI-BJ project: the electromagnetic data validation and scientific application research based on CSES satellite. GH and PV acknowledge travel support from the European Space Agency through Swarm DISC funding, and support for computation from CNES through the “Suivi et exploitation de la mission Swarm” project.

#### Availability of data and materials

CSES HPM data were obtained and are accessible from [www.leos.ac.cn](http://www.leos.ac.cn). Swarm data and Kp index values were downloaded from the FTP server <ftp.swarm-diss.eo.esa.int>. RC data were taken from <http://www.spacecenter.dk/files/magnetic-models/RC/current/>. The merging electric field Em was computed from Omni data downloaded from [ftp://spdf.gsfc.nasa.gov/pub/data/omni/high\\_res\\_omni/](ftp://spdf.gsfc.nasa.gov/pub/data/omni/high_res_omni/). CGGM model coefficients can be downloaded from <http://www.leos.ac.cn>, and are also available on request from the corresponding authors.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interest.

#### Author details

<sup>1</sup> National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China. <sup>2</sup> Université de Paris, Institut de Physique du Globe de Paris, CNRS, 75005 Paris, France. <sup>3</sup> National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China. <sup>4</sup> Space Research Institute, Austrian Academy of Sciences, 8042 Graz, Austria. <sup>5</sup> DTU Space, National Space Institute, Technical University of Denmark, 2800 Kongens Lyngby, Denmark. <sup>6</sup> Institute of Earthquake Forecasting, China Earthquake Administration, Beijing 100036, China. <sup>7</sup> DFH Satellite Co. Ltd., Beijing 100081, China. <sup>8</sup> Institute of Experimental Physics, Graz University of Technology, 8010 Graz, Austria. <sup>9</sup> Beijing Special Engineering Design and Research Institute, Beijing 100028, China. <sup>10</sup> China Centre for Resources Satellite Data and Application, Beijing 100094, China. <sup>11</sup> Hebei GEO University, Shijiazhuang 050031, China. <sup>12</sup> Institute of Geophysics, China Earthquake Administration, Beijing 100081, China.

Received: 12 March 2020 Accepted: 16 November 2020

Published online: 11 February 2021

#### References

- Alken P, Thébaud 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 (2020a) Evaluation of candidate models for the 13th international geomagnetic reference field. *Earth Planets Space* 203(3):1738–1751
- Alken P, Thébaud E, Beggan CD, Amit H, Aubert J, Baerenzung J, Bondar TN, Brown W, Califf S, Chambodut A, Chulliat A, Cox G, Finlay CC, Fournier A, Gillet N, Grayver A, Hammer MD, Holschneider M, Huder L, Hulot G, Jager T, Kloss C, Korte M, Kuang W, Kuvshinov A, Langlais B, Léger JM, Lesur V, Livermore PW, Lowes FJ, Macmillan S, Mound JE, Nair M, Nakano S, Olsen N, Pavón-Carrasco FJ, Petrov VG, Ropp G, Rother M, Sabaka TJ, Sanchez S, Saturnino D, Schnepf NR, Shen X, Stolle C, Tangborn A, Tøffner-Clausen L, Toh H, Torta JM, Varnier J, Vervelidou F, Vigneron P, Wardinski I, Wicht J, Woods A, Yang Y, Zeren Z, Zhou B (2020b) International geomagnetic reference field: the thirteenth generation. *Earth Planets Space* 67(1):79
- Alken P, Chulliat A, Nair M (2020c) NOAA/NCEI and University of Colorado candidate models for IGRF-13. *Earth Planets Space*. <https://doi.org/10.1186/s40623-020-01313-z>
- Backus G (1970) Non-uniqueness of the external geomagnetic field determined by surface intensity measurements. *J Geophys Res* 75:6339–6341. <https://doi.org/10.1029/JA075i031p06339>
- Baerenzung J, Holschneider M, Wicht J, Lesur V, Sanchez S (2020) The Kalmag model as a candidate for IGRF-13. *Earth Planets Space* 12:428
- Bartels J (1949) The standardized index, Ks, and the planetary index Kp. *IATME Bull* 12b. IUGG Publ Office, Paris, pp 97–112
- Brown W, Beggan CD, Cox G, Macmillan S (2020) The BGS candidate models for IGRF-13 with a retrospective analysis of IGRF-12 secular variation forecasts. *Earth Planets Space* 67(24):12
- Cheng BJ, Zhou B, Magnes W, Lammegger R, Pollinger A (2018) High precision magnetometer for geomagnetic exploration onboard of the China Seismo-Electromagnetic Satellite. *Sci China Tech Sci* 61(5):659–668. <https://doi.org/10.1007/s11431-018-9247-6>
- Farquharson CG, Oldenburg DW (1998) Non-linear inversion using general measures of data misfit and model structure. *Geophys J Int* 134(1):213–227. <https://doi.org/10.1046/j.1365-246x.1998.00555.x>
- Finlay CC, Olsen N, Kotsiaros S, Gillet N, Tøffner-Clausen L (2016) Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model. *Earth Planets Space* 68:112. <https://doi.org/10.1186/s40623-016-0486-1>
- Finlay CC, Kloss C, Olsen N, Hammer M, Tøffner-Clausen L, Grayver A, Kuvshinov A (2020) The CHAOS-7 geomagnetic field model and observed changes in the South Atlantic Anomaly. *Earth Planets Space* 17:28
- Fournier A, Aubert J, Lesur V, Ropp G (2020) A secular variation candidate model for IGRF-13 based on Swarm data and ensemble inverse geodynamo modelling. *Earth Planets Space* 72:136. <https://doi.org/10.1186/s40623-020-01309-9>
- Holme R, Bloxham J (1996) The treatment of attitude errors in satellite geomagnetic data. *Phys Earth Planet Int* 98:221–233. [https://doi.org/10.1016/S0031-9201\(96\)03189-5](https://doi.org/10.1016/S0031-9201(96)03189-5)
- Huder L, Gillet N, Finlay CC, Hammer MD, Tchoingui H (2020) COV-OBS.x2: 180 year of geomagnetic field evolution from ground-based and satellite observations. *Earth Planets Space* 411:20–26
- 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 (2015a) Swarm's absolute magnetometer experimental vector mode, an innovative capability for space magnetometry. *Geophys Res Lett* 42:1352–1359. <https://doi.org/10.1002/2014GL062700>
- Hulot G, Sabaka TJ, Olsen N, Fournier A (2015b) The present and future geomagnetic field. In: Gerald S (ed) *Treatise on geophysics*, vol 5. Elsevier, Oxford, pp 33–78
- Kan JR, Lee LC (1979) Energy coupling function and solar wind-magnetosphere dynamo. *Geophys Res Lett* 6:577–580. <https://doi.org/10.1029/GL006i007p00577>
- Khokhlov A, Hulot G, Le Mouél JL (1997) On the Backus effect I. *Geophys J Int* 130:701–703. <https://doi.org/10.1111/j.1365-246X.1997.tb01864.x>
- Khokhlov A, Hulot G, Le Mouél JL (1999) On the Backus effect II. *Geophys J Int* 137:816–820. <https://doi.org/10.1046/j.1365-246x.1999.00843.x>
- Lowes FJ (1966) Mean-square values on sphere of spherical harmonic vector fields. *J Geophys Res* 71:2179. <https://doi.org/10.1029/JZ071i008p02179>
- Lowes FJ (1975) Vector errors in spherical harmonic analysis of scalar data. *Geophys J R Astr Soc* 42(2):637–651
- Lowes FJ, Olsen N (2004) A more realistic estimate of the variances and systematic errors in spherical harmonic geomagnetic field models. *Geophys J Int* 157:1027–1044. <https://doi.org/10.1111/j.1365-246X.2004.02256.x>
- Macmillan S, Finlay C (2011) The international geomagnetic reference field. In: Manda M, Korte M (eds) *Geomagnetic observations and models*, vol 5. IAGA Special Sopron Book Series. Springer, Heidelberg, pp 265–276
- Mauersberger P (1956) Das Mittel der Energiedichte des geomagnetischen Hauptfeldes an der Erdoberfläche und seine säkulare Änderung. *Gerl Beitr Geophys* 65:207–215
- Maus S, Leuhr H (2005) Signature of the quiet-time magnetospheric magnetic field and its electromagnetic induction in the rotating Earth. *Geophys J Int* 162(3):755–763. <https://doi.org/10.1111/j.1365-246X.2005.02691.x>
- Menvielle M, Berthelier A (1991) The K-derived planetary indices: description and availability. *Rev Geophys* 29(3):415–432. <https://doi.org/10.1029/91RG00994>
- Metman MC, Beggan CD, Livermore PW, Mound JE (2020) Forecasting yearly geomagnetic variation through sequential estimation of core flow and magnetic diffusion. *Earth Planets Space* 68(1):112
- Minami T, Nakano S, Lesur V, Takahashi F, Matsushima M, Shimizu H, Nakashima R, Taniguchi H, Toh H (2020) A candidate secular variation model for IGRF-13 based on MHD dynamo simulation and 4DnVar data assimilation. *Earth Planets Space*. <https://doi.org/10.1186/s40623-020-01253-8>



- Olsen N, Hulot G, Lesur V, Finlay CC, Beggan CC, Chulliat A, Sabaka TJ, Floberghagen R, Friis-Christensen E, Haagmans R, Kotsiaros S, Lühr H, Tøffner-Clausen L, Vigneron P (2015) The Swarm Initial Field Model for the 2014 geomagnetic field. *Geophys Res Lett* 42:1092–1098. <https://doi.org/10.1002/2014GL062659>
- 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>
- Pavón-Carrasco FJ, Marsal S, Torta JM, Catalan M, Martín-Hernández F, Tordesillas JM (2020) Bootstrapping Swarm and observatory data to generate candidates for the DGRF and IGRF-13. *Earth Planets Space* 72:109
- Petrov VG, Bondar TN (2020) IZMIRAN sub-model for IGRF-13. *Earth Planets Space*. <https://doi.org/10.1088/1361-6501/aacde4>
- Pollinger A, Lammegger R, Magnes W, Hagen C, Ellmeier M, Jernej I, Leichtfried M, Kürbisch C, Maierhofer R, Wallner R, Fremuth G, Amtmann C, Betzler A, Delva M, Prattes G, Baumjohann W (2020) Coupled dark state magnetometer for the China seismo-electromagnetic satellite. *Measur Sci Tech* 29:9. <https://doi.org/10.1088/1361-6501/aacde4>
- Pollinger A, Amtmann C, Betzler A, Cheng B, Ellmeier M, Hagen C, Jernej I, Lammegger R, Zhou B, Magnes W (2020) In-orbit results of the Coupled Dark State Magnetometer aboard the China Seismo-Electromagnetic Satellite. *Geosci Instrum Method Data Syst* 9:275–291. <https://doi.org/10.5194/gi-9-275-2020>
- Richmond AD (1995) Ionospheric electrodynamics using magnetic Apex coordinates. *J Geomagn Geoelectr* 47:191–212. <https://doi.org/10.5636/jgg.47.191>
- Ropp G, Lesur V, Baerenzung J, Holschneider M (2020) Sequential modelling of the Earth's core magnetic field. *Earth Planets Space*. 68:102. <https://doi.org/10.1186/s40623-020-01230-1>
- Rother M, Korte M, Morschhauser A, Vervelidou F, Matzka J, Stolle C (2020) The Magnum core field model as a parent for IGRF-13, and the recent evolution of the South Atlantic Anomaly. *Earth Planets Space* 15:98
- Sabaka TJ, Tøffner-Clausen L, Olsen N, Finlay CC (2020) CM6: a comprehensive geomagnetic field model derived from both CHAMP and swarm satellite observations. *Earth Planets Space* 72:88
- Sanchez S, Wicht J, Baerenzung J (2020) Predictions of the geomagnetic secular variation based on the ensemble sequential assimilation of geomagnetic field models by dynamo simulations. *Earth Planets Space* 72:157
- Shen XH, Zhang XM, Yuan SG, Wang LW, Cao JB, Huang JP, Zhu XH, Piergiorgio P, Dai JP (2018) The state-of-the-art of the China Seismo-Electromagnetic Satellite mission. *Sci China Tech Sci* 61(5):634–642. <https://doi.org/10.1007/s11431-018-9242-0>
- Thébault E, Finlay CC, Beggan CD, Alken P, Aubert J, Barrois O, Bertrand F, Bondar T, Boness A, Brocco L, Canet E, Chambodut A, Chulliat A, Coisson P, Civet F, Du A, Fournier A, Fratter I, Gillet N, Hamilton B, Hamoudi M, Hulot G, Jager T, Korte M, Kuang W, Lalanne X, Langlais B, Léger JM, Lesur V, Lowes FJ, Macmillan S, Mandea M, Manoj C, Maus S, Olsen N, Petrov V, Ridley V, Rother M, Sabaka TJ, Saturnino D, Schachtschneider R, Sirol O, Tangborn A, Thomson A, Tøffner-Clausen L, Vigneron P, Wardinski I, Zvereva T (2015) International geomagnetic reference field: the 12th generation. *Earth Planets Space* 67:79. <https://doi.org/10.1186/s40623-015-0228-9>
- Thébault E, Finlay CC, Alken P, Beggan CD, Canet E, Chulliat A, Langlais B, Lesur V, Lowes FJ, Manoj C, Rother M, Schachtschneider R (2015) Evaluation of candidate geomagnetic field models for IGRF-12. *Earth Planets Space* 67:112. <https://doi.org/10.1186/s40623-015-0273-4>
- Vigneron P, Hulot G, Olsen N, Léger JM, Jager T, Brocco L, Sirol O, Coisson P, Lalanne X, Chulliat A, Bertrand F, Boness A, Fratter I (2015) International Geomagnetic Reference Field (IGRF) candidate model based on Swarm's experimental absolute magnetometer vector mode data. *Earth Planets Space* 67:95. <https://doi.org/10.1186/s40623-015-0265-4>
- Wardinski I, Saturnino D, Amit H, Chambodut A, Langlais B, Mandea M, Thébault E (2020) Geomagnetic core field models and secular variation forecasts for the 13th International Geomagnetic Reference Field (IGRF-13). *Earth Planets Space* 62:821
- Yang YY, Zeren ZM, Shen XH, Chu W, Huang JP, Wang Q, Yan R, Xu S, Lu HX, Liu DP (2020) The first intense geomagnetic storm event recorded by the China Seismo-Electromagnetic Satellite. *Space Weather* 18:43. <https://doi.org/10.1029/2019SW002243>
- Zhou B, Yang YY, Zhang YT, Gou XC, Cheng BJ, Wang JD, Li L (2018) Magnetic field data processing methods of the China Seismo-Electromagnetic Satellite. *Earth Planet Phys* 2(6):455–461. <https://doi.org/10.26464/epp2018043>
- Zhou B, Cheng B, Gou X et al (2019) First in-orbit results of the vector magnetic field measurement of the High Precision Magnetometer onboard the China Seismo-Electromagnetic Satellite. *Earth Planets Space* 71:119. <https://doi.org/10.1186/s40623-019-1098-3>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)