

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

**Please provide feedback**

Please support the ScholarWorks@UMBC repository by emailing [scholarworks-group@umbc.edu](mailto:scholarworks-group@umbc.edu) and telling us what having access to this work means to you and why it's important to you. Thank you.

# CM5, a pre-*Swarm* comprehensive geomagnetic field model derived from over 12 yr of CHAMP, Ørsted, SAC-C and observatory data

Terence J. Sabaka,<sup>1</sup> Nils Olsen,<sup>2</sup> Robert H. Tyler<sup>1,3</sup> and Alexey Kuvshinov<sup>4</sup>

<sup>1</sup>Planetary Geodynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. E-mail: [Terence.J.Sabaka@nasa.gov](mailto:Terence.J.Sabaka@nasa.gov)

<sup>2</sup>DTU Space, National Space Institute, Technical University of Denmark, Diplomvej, 2800 Kgs. Lyngby, Denmark

<sup>3</sup>Department of Astronomy, University of Maryland, College Park, MD 20742, USA

<sup>4</sup>Institute of Geophysics, ETH Zürich, Sonneggstrasse 5, 8092 Zürich, Switzerland

Accepted 2014 December 22. Received 2014 November 21; in original form 2014 August 13

## SUMMARY

A comprehensive magnetic field model named CM5 has been derived from CHAMP, Ørsted and SAC-C satellite and observatory hourly-means data from 2000 August to 2013 January using the *Swarm* Level-2 Comprehensive Inversion (CI) algorithm. *Swarm* is a recently launched constellation of three satellites to map the Earth's magnetic field. The CI technique includes several interesting features such as the bias mitigation scheme known as Selective Infinite Variance Weighting (SIVW), a new treatment for attitude error in satellite vector measurements, and the inclusion of 3-D conductivity for ionospheric induction. SIVW has allowed for a much improved lithospheric field recovery over CM4 by exploiting CHAMP along-track difference data yielding resolution levels up to spherical harmonic degree 107, and has allowed for the successful extraction of the oceanic  $M_2$  tidal magnetic field from quiet, nightside data. The 3-D induction now captures anomalous Solar-quiet features in coastal observatory daily records. CM5 provides a satisfactory, continuous description of the major magnetic fields in the near-Earth region over this time span, and its lithospheric, ionospheric and oceanic  $M_2$  tidal constituents may be used as validation tools for future *Swarm* Level-2 products coming from the CI algorithm and other dedicated product algorithms.

**Key words:** Geomagnetic induction; Magnetic anomalies: modelling and interpretation; Magnetic field; Satellite magnetics.

## 1 INTRODUCTION

The European Space Agency (ESA) successfully launched the *Swarm* mission (Friis-Christensen *et al.* 2006) on 2013 November 22, a constellation of three satellites to map the Earth's magnetic field to unprecedented accuracy. During its multiyear lifetime, two low-orbiting spacecrafts will act as a cross-track magnetic gradiometer while a third at higher altitude monitors the main and external fields at other local times. ESA has established the *Swarm* 'Satellite Constellation Application and Research Facility' (SCARF) for the purposes of generating derived Level-2 products from the single-satellite Level-1b data. The advanced 'Comprehensive Inversion' (CI) method of Sabaka *et al.* (2013) is a major processing chain of SCARF, producing one version of each of five defined items: core, lithospheric, magnetospheric, and ionospheric spherical harmonic (SH) expansions, time-varying when appropriate, and Euler angles describing the alignment between the vector fluxgate magnetometer frame (VFM) system and that of the star-imager.

Before these Level-2 products are released to the community, they are internally checked by the institutions generating them; in the case of the CI products, these are the NASA Goddard Space Flight Center (GSFC) and the Technical University of Denmark (DTU). They then undergo a series of checks carried out by independent partners of the Level-2 processing facility, which are based upon comparisons with complementary products from other Level-2 chains and other data sets or models developed independently of *Swarm* (Beggan *et al.* 2013). The existence of high-quality satellite and observatory measurements in the recent pre-*Swarm* era and the need for such independent models for *Swarm* product validation has provided the motivation for applying the CI Level-2 algorithm to these earlier data. This paper describes the resulting model, denoted CM5, and represents a natural progression from the paper of Sabaka *et al.* (2013), which established the advanced algorithm and described its application to *Swarm* synthetic test data sets.

The CM5 model is derived from magnetic measurements from the Ørsted, CHAMP and SAC-C satellites and observatory hourly-means (OHM) data over the recent 12-yr time span from 2000 August to 2013 January using a CI method that is slightly modified from that reported in Sabaka *et al.* (2013). These modifications include discriminating between nightside and dayside data, which is then exploited by 'Selective

Infinite Variance Weighting' (SIVW; Sabaka *et al.* 2013) to produce much improved core and crustal field models such that the latter is no longer contaminated by leakage from Equatorial Electrojet (EEJ) signals. To enhance high-degree crustal fidelity, gradient information in along-track vector measurement differences from CHAMP are used to determine the field for SH degrees  $n > 60$  via SIVW resulting in improved high-degree correlation with the MF7 model (Maus 2010). In addition, the magnetic field generated by the oceanic  $M_2$  tide has now been successfully parametrized and extracted within the framework of the CI approach. While these aforementioned fields were all determined with data during quiet magnetic times, a continuous solution over the time span was desired, particularly for the magnetosphere and high-frequency induced fields. Therefore, a final application of SIVW was used to separate the data during disturbed magnetic times in order to achieve this.

The paper is organized as follows: Section 2 provides information on the data sets used to derive the model, Section 3 gives a brief overview of the CI method, which is presented in great detail in Sabaka *et al.* (2013), and describes the modifications previously mentioned, and Section 4 discusses the results.

## 2 DATA SELECTION

As will be seen in Section 3, the use of SIVW to target various parameter sets has motivated the selection of data in two tiers: a coarse selection so as to provide continuous coverage at the hourly level from 2000 August to 2013 January, regardless of geomagnetic activity levels; and a finer selection based upon quiet geomagnetic activity levels over this same time period. Thus, the former allows for a continuous description to be made of rapidly changing, aperiodic fields such as the magnetosphere and its associated induced fields. The latter allows for a more accurate description of the remaining fields. To this end, the Ørsted scalar data were selected between 2000 August and 2013 January, Ørsted vector data between 2000 August and 2004 December, CHAMP vector data between 2000 August and 2010 September and SAC-C scalar data between 2001 January and 2004 December. Although Ørsted data are available since 1999 March, it was decided to take only data after 2000 July, which is the month of the launch of the CHAMP satellite. The CHAMP magnetic field data is Level-3 and consists of vector measurements in the VFM frame and attitude data [between the 'Common Reference Frame' (CRF) of the dual star-imager heads and the J2000 Celestial Reference Frame] provided by the star-imager. The OHM data were selected over the entire time span of the model, 2000 August to 2013 January, and are taken from the corrected data set described in Macmillan & Olsen (2013).

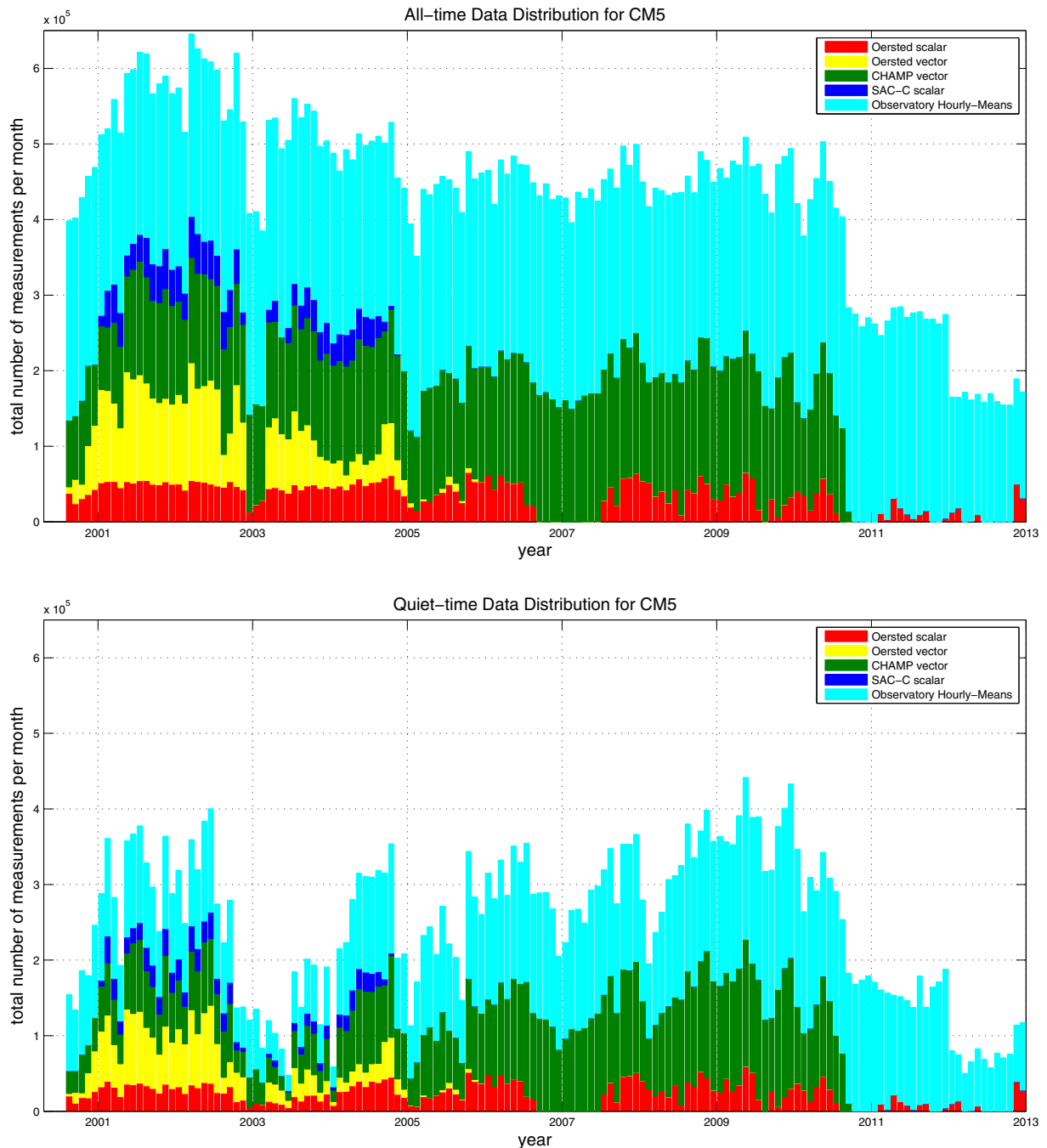
For selecting geomagnetic quiet periods, criteria similar to that used for the CHAOS-4 model (Olsen *et al.* 2014) were employed here: (1) at non-polar latitudes (equatorward of  $\pm 60^\circ$  dipole latitude), only data for which the geomagnetic activity index  $K_p \leq 2^0$  and the  $D_{st}$  index (see Rangarajan 1989), which measures the strength of the magnetospheric ring-current, do not change by more than  $2 \text{ nT hr}^{-1}$ ; (2) for regions poleward of  $60^\circ$  dipole latitude, the merging electric field at the magnetopause,  $E_m$  (see Kan & Lee 1979), must be below  $0.8 \text{ mV m}^{-1}$ . Note that CHAOS-4 used weighted sums of the revised definition of  $E_m$  of Newell *et al.* (2007). All satellite data have been decimated to a 30 s sampling rate. Similar criteria have been used for selecting the OHM data, the requirements being that  $K_p \leq 2^0$  and  $|dD_{st}/dt| \leq 2 \text{ nT hr}^{-1}$  at all latitudes.

Note that although decimation levels and quiet-time definitions used here are similar to those for CHAOS-4, the data sets between the two models are very different, besides the obvious use of data during disturbed geomagnetic conditions in CM5, in three important ways: (1) data at all local times at all latitudes, including polar latitudes, are used here, which is in fact a major departure from almost any other geomagnetic field model; (2) satellite vector data are used in the polar regions and (3) observatory data at the hourly scale are used as opposed to monthly, although hourly-means were used to derive the  $RC$  index which was then used to characterize the ring-current in CHAOS-4. The first and third points are a consequence of wanting to describe fields at all local times, such as the ionospheric, toroidal, magnetospheric, etc., and the second point allows for global descriptions of toroidal fields in the satellite sampling shells.

The total number of measurements used in deriving CM5 are shown as stacked histograms with respect to time in Fig. 1, where the top and bottom plots are for all and quiet geomagnetic times, respectively. Both histograms are at the same scale, which illustrates the relative amount of disturbed to quiet data used. A scalar and vector measurement would count as 1 and 3, respectively, in the histograms. Obviously, periods with fewer quiet time data are due to increased geomagnetic activity levels, but the absence of Ørsted vector data around 2003 is due to the fact that the spacecraft was in full sun (a dawn-dusk orbit), while problems with the attitude stability are the reason for the Ørsted data gaps around 2007 and after 2011 (Olsen *et al.* 2014). In addition, note that there is a significant decrease in observatory data during 2012, which is due to the fact that INTERMAGNET definitive data were not yet available for that year. Finally, the spatial locations of the 186 magnetic observatory stations that provide the OHM data in CM5 are shown as red dots in Fig. 2.

## 3 METHODOLOGY

The CI method is based upon the so-called 'comprehensive approach' (CA) to deriving field models in which the major magnetic field sources are parametrized and then co-estimated via least-squares (LS) in order to achieve optimal separation of the fields (see Sabaka *et al.* 2002, 2004; Olsen *et al.* 2006, 2014; Sabaka & Olsen 2006; Thomson & Lesur 2007; Lesur *et al.* 2008, 2010; Thomson *et al.* 2010). This is in contrast to the 'sequential approach' (e.g. Maus *et al.* 2006, 2007) in which the parameters are separated into groups and sequentially solved via LS and their contributions removed from the data. The latter can be shown to be equivalent to applying a multiblock Gauss-Seidel (MBGS) method (see Demmel 1997) to the full problem, which requires iteration to converge to the solution obtained by the CA. There



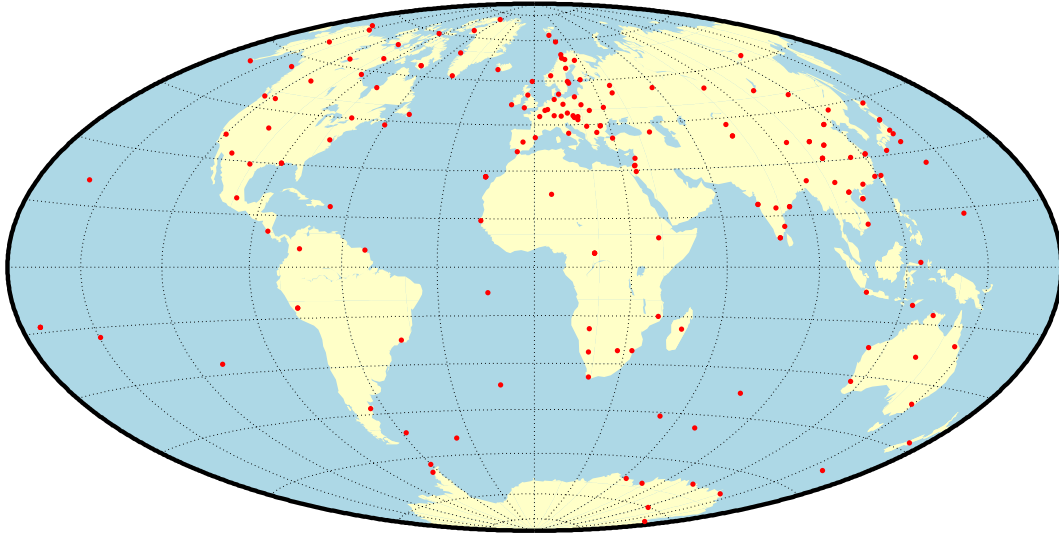
**Figure 1.** Stacked histograms of the number of measurements from all geomagnetic times (top panel) and quiet geomagnetic times (bottom panel) used in deriving the CM5 model as a function of time. Note that both histograms are at the same scale. Measurement counts include all components, that is, a scalar and vector measurement would count as 1 and 3 measurements, respectively.

are, however, at least three problems to this approach as practiced: (1) there is typically only a single iteration performed; (2) even if more iterations were performed, convergence could be very slow and (3) MBGS can easily mask co-linearities between parameter basis functions that lie in separate blocks. Therefore, because of the multitude of sources considered here, the CA is followed in this work.

### 3.1 Model parametrization

#### 3.1.1 Core and lithospheric fields

The parametrizations of the various magnetic field sources included in this model are summarized in Table 1. While the core and lithosphere are of the same general form as in the CM4 model (Sabaka *et al.* 2004), they differ in SH truncation level, B-spline knot spacing and epoch of



**Figure 2.** Locations (red dots) of the 186 magnetic observatory stations that provided the hourly-means measurements used in deriving the CM5 model.

**Table 1.** CM5 parametrization.

Field source/Effect	# Params	Description	
Core/lithospheric fields	26 960	Spatial:	geographic spherical harmonic (SH) $N_{\max} = 120$
		Temporal:	order 4 B-splines, 6 month knot spacing from 2000.5 – 2013.0, epoch 2005.0, up to $N_{\max} = 20$
$M_2$ tidal field	2736	Spatial:	geographic SH $N_{\max} = 36$
		Temporal:	period of 12.42060122 hr, phase fixed with respect to 00:00:00, 1999 January 1 GMT
Ionospheric/induced fields	5520	Spatial:	quasi-dipole (QD) frame, underlying dipole SH $N_{\max} = 60$ , $M_{\max} = 12$
		Temporal:	annual, semi-annual, 24-, 12-, 8- and 6-hr periodicities with $F_{10.7}$ scaling plus induction via <i>a priori</i> 3-D conductivity model ('1-D+oceans') and infinite conductor at depth
Toroidal field	12 240	Spatial:	meridional currents in QD frame, underlying dipole SH $N_{\max} = 60$ , $M_{\max} = 12$ , one for Ørsted centred at 750 km altitude and one for CHAMP centred at 400 km altitude
		Temporal:	annual, semi-annual, 24-, 12-, 8- and 6-hr periodicities
Magnetospheric/induced fields	653 184	Magnetospheric	Spatial: dipole SH $N_{\max} = 1$
			Temporal: discretized in 1-hr bins
		Induced	Spatial: dipole SH $N_{\max} = 1$
			Temporal: discretized in 1-hr bins
OHM biases	558	One vector bias for each station in local spherical system	
VFM-CRF alignment	1110	Three XYZ-type Euler angles every 10 d for CHAMP, one set of three YZY-type Euler angles for Ørsted	
Total	702 308	---	

the secular variation (SV) expansion. The epoch of CM5 was chosen to be 2005.0, which is approximately the midpoint of the CHAMP data envelope, although the temporal basis is defined from 2000.5 to 2013.0. The continuous time sampling of the CHAMP vector data combined with its low altitude towards the end of the mission is the reason for using a 6-month spacing between B-spline knots up to SH maximum degree  $n = 20$  and a static limit of  $N_{\max} = 120$  compared with a 2.5-yr spacing up to  $n = 13$  and a static limit of  $N_{\max} = 65$  for CM4. This

leads to the following expression for the core/lithospheric potential at time  $t$  and position  $\mathbf{r}$ , corresponding to Earth-Centered Earth-Fixed (ECEF) spherical coordinates of radius, colatitude and longitude  $(r, \theta, \phi)$ ,

$$V_{cl}(t, \mathbf{r}) = \Re \left\{ a \sum_{n=1}^{20} \sum_{m=0}^n \sum_{q=0}^{24} \left( \frac{a}{r} \right)^{n+1} \gamma_{nq}^m Y_{nq}^m(t, \theta, \phi) + a \sum_{n=21}^{120} \sum_{m=0}^n \left( \frac{a}{r} \right)^{n+1} \gamma_n^m Y_n^m(\theta, \phi) \right\}, \quad (1)$$

where the  $\Re\{\cdot\}$  operator takes the real part of the expression only and  $Y_n^m$  is the surface SH of degree  $n$  and order  $m$  given by

$$Y_n^m(\theta, \phi) = P_n^m(\cos \theta) \exp i m \phi, \quad (2)$$

where  $a$  is the Earth's mean-radius (6371.2 km),  $P_n^m$  and  $\gamma_n^m$  are the Schmidt semi-normalized associated Legendre function and static complex Gauss coefficient of degree  $n$  and order  $m$ , respectively. The complex Gauss coefficient is related to the usual real Gauss coefficients  $g_n^m$  and  $h_n^m$  according to  $\gamma_n^m = g_n^m - i h_n^m$ . The time-variable Gauss coefficients are expressed as linear combinations of the integrals of 24 cubic B-spline functions (equivalent to quartic or order-5 B-splines) and a static term with multipliers  $\gamma_{nq}^m$ ,  $q = 0, \dots, 24$  and are related to the modified surface SH  $Y_{nq}^m$  as

$$Y_{nq}^m(t, \theta, \phi) = \begin{cases} Y_n^m(\theta, \phi) & \text{for } q = 0, \\ Y_n^m(\theta, \phi) \int_{2005}^t b_q(\tau) d\tau & \text{for } q > 0, \end{cases} \quad (3)$$

where  $b_q$  is the  $q$ th cubic B-spline of the expansion. It should be noted that other models such as GRIMM-2 (Lesur *et al.* 2010) and CHAOS-4 (Olsen *et al.* 2014) use order-6 rather than order-5 B-splines, as in CM5, to represent time-variable Gauss coefficients. While this allows for larger degrees of freedom in their representation, it is not clear that this brings significant advantage, especially given that regularization is always required. Because OHMs are modelled directly, eq. (1) may not be sufficient to capture the local crustal anomaly signal they measure. Therefore, static vector biases are included in order to absorb crustal signal above  $N_{\max} = 120$  and may also capture DC levels in the ionospheric field (see Sabaka *et al.* 2004).

### 3.1.2 Oceanic $M_2$ tidal field

The gravitational fields of the Sun and Moon have components with a time-dependent projection over the surface of the Earth. The spatial gradients in these fields are referred to as tidal forces, the largest-amplitude constituent of which is the so-called  $M_2$  tidal force. The principal lunar semi-diurnal constituent ( $M_2$ ) represents the tidal gravitational field of the Moon as projected onto a circular, equatorial orbit around the Earth. Additional constituents represent improvements to this approximate representation of the lunar tidal force as well as the solar tidal forces and solar–lunar products. These other constituent forces are smaller in amplitude than  $M_2$ .

The work of Tyler *et al.* (2003) showed that tidal magnetic signals, particularly those of the dominant  $M_2$  constituent, can indeed be detected at satellite altitude. This motivated its inclusion in the parametrization of CM5 where it was modelled as a simple internal SH expansion to  $N_{\max} = 36$  having a 12.42060122-hr periodicity with time  $\Delta t$  rendered with respect to an arbitrary reference time, in this case 00:00:00, 1999 January 1 GMT, in order to fix the phase. Note that this phase must be adjusted by an angle of  $35.71931^\circ$  or time-shift of 73.9 min to obtain the conventional Greenwich phase used in ocean tidal literature. This maximum degree,  $N_{\max}$ , reflects the  $5^\circ \times 5^\circ$  degree bins in which Tyler *et al.* (2003) averaged their residuals. Note that there is no information about the geographic distribution of oceans and continents imposed in the basis functions. The  $M_2$  potential at time  $\Delta t$  and position  $\mathbf{r}$  in the ECEF system is then

$$V_{M_2}(\Delta t, \mathbf{r}) = \Re \left\{ a \sum_{n=1}^{36} \sum_{m=-n}^n \left( \frac{a}{r} \right)^{n+1} \tau_n^m Y_{n\omega}^m(\Delta t, \theta, \phi) \right\}, \quad (4)$$

where  $\tau_n^m$  is the complex coefficient and

$$Y_{n\omega}^m(\Delta t, \theta, \phi) = P_n^m(\cos \theta) \exp i(m\phi + \omega_{M_2} \Delta t), \quad (5)$$

with  $\omega_{M_2} = 2\pi/12.42060122$  rads/hr.

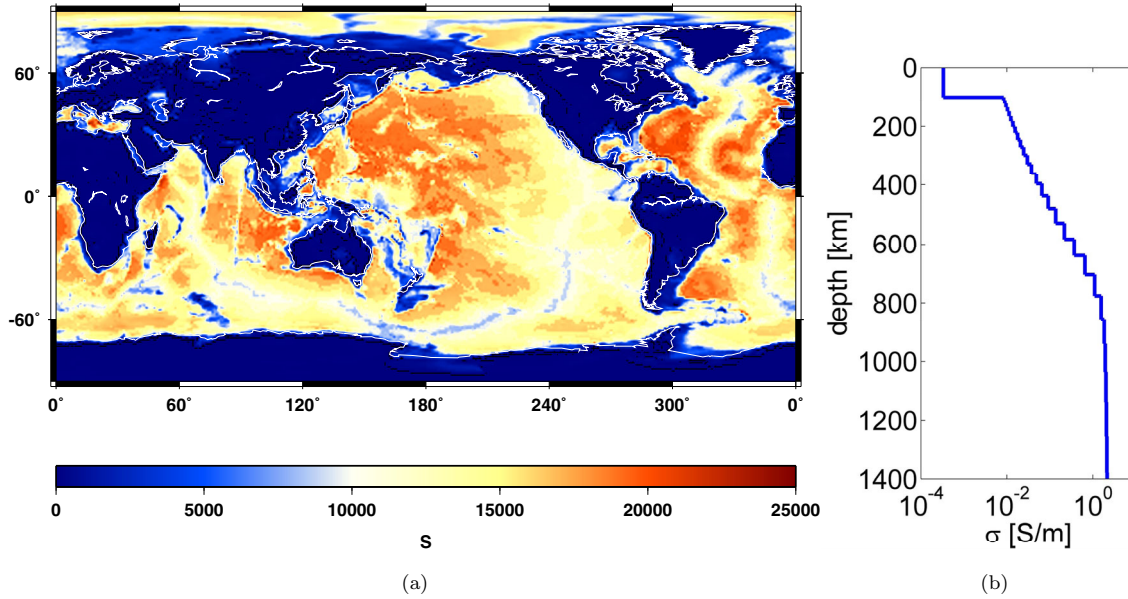
### 3.1.3 Ionospheric and toroidal fields

Although the parametrization is basically the same as that used in CM4 for the ionospheric and toroidal fields, the quasi-dipole (QD) symmetric basis functions used have been updated to epoch 2010. Briefly, the QD coordinates  $(h_q, \theta_q, \phi_q)$  were proposed by Richmond (1995) and are tied to a magnetic field model; the IGRF-11 model (Finlay *et al.* 2010) in this case. The QD longitude,  $\phi_q$ , is the dipole longitude of the apex (highest point above Earth's surface, given by  $h_A$ ) of the magnetic field line passing through the point. The QD latitude,  $\lambda_q$ , is given by

$$\lambda_q = \pm \arccos \left( \frac{a + h_q}{a + h_A} \right)^{1/2}, \quad (6)$$

where  $h_q$  is altitude. Thus,  $\lambda_q$  is defined as the latitude of the footprint of an axial dipole field line through the apex, and goes to zero at the magnetic equator at all altitudes. Hence, QD coordinates are defined by tracing the field line upwards using IGRF-11, and downwards





**Figure 3.** (a) Conductance of the surface shell describing oceans and sediments in the oceans section of the ‘1-D+oceans’ 3-D conductivity structure. Units are in Siemens (S). (b) The 1-D mantle section of the ‘1-D+oceans’ 3-D conductivity structure.

again using an axial dipole (see Emmert *et al.* 2010, for computational aspects of the QD coordinates). This coordinate system is useful in describing the ionospheric and coupling currents that are strongly horizontally organized by the ambient magnetic field.

More importantly, the induced secondary field associated with the ionospheric primary field is now more advanced and reflects a 3-D conductivity structure that consists of an ocean–continent contrast in the upper layer with a 1-D mantle underneath known as ‘1-D+oceans’ (Kuvshinov 2008). In CM4, this field was induced by a 1-D conductivity model. The ‘1-D+oceans’ model consists of a thin spherical shell of variable conductance at the Earth’s surface over a spherically symmetric 1-D mantle. The shell conductance, shown in Fig. 3(a), is obtained by considering contributions both from sea water and from sediments. The conductance of the sea water has been taken from Manoj *et al.* (2006) and accounts for ocean bathymetry (taken from the global  $5' \times 5'$  NOAA ETOPO map of bathymetry/topography), ocean salinity, temperature and pressure, as given by the World Ocean Atlas 2001 ([www.nodc.noaa.gov/OC5/WOA01/pr\\_woa01.html](http://www.nodc.noaa.gov/OC5/WOA01/pr_woa01.html)). Conductance of the sediments (in continental as well as oceanic regions) is based on the global sediment thicknesses given by the map of Laske & Masters (1997) and calculated by a heuristic procedure similar to that described in Everett *et al.* (2003). The resolution of the model is  $1^\circ \times 1^\circ$  in colatitude and longitude. The 1-D mantle section, shown in Fig. 3(b), is a modified global conductivity section that has been inferred from satellite magnetic data by Kuvshinov & Olsen (2006). The modification comes from assigning a realistic low conductivity value to the upper 100 km as these depths are not resolvable by satellite magnetic data.

This information is conveyed through spectral transfer functions or ‘Q matrices’ for the 24-, 12-, 8- and 6-hr periodicities. If  $\epsilon(\omega)$  and  $\iota(\omega)$  are the vectors of SH coefficients for the inducing and induced ionospheric fields, respectively, at frequency  $\omega$ , then the *a priori* coupling via the conductivity model is manifested in the relationship  $\iota(\omega) = \mathbf{Q}(\omega)\epsilon(\omega)$ , where  $\mathbf{Q}(\omega)$  is the coupling matrix at frequency  $\omega$ .

For CM4,  $\mathbf{Q}(\omega)$  is square and diagonal and its elements are only dependent upon SH degree, reflecting the fact that a 1-D conductivity model was used. This means that a particular harmonic can only induce the same harmonic. For CM5, however,  $\mathbf{Q}(\omega)$  is a dense matrix meaning that complicated induced structure can result from relatively smooth inducing structure. Since the QD basis functions are built from ordinary SH expansions with truncation levels of  $N_{\max} = 60$  and  $M_{\max} = 12$ , giving a total of 1368 SH functions, the  $\mathbf{Q}(\omega)$  are then  $1368 \times 1368$  complex matrices. The influence of longer period modulations on induction, such as annual, semi-annual and  $F_{10.7}$ , have been treated via an infinite conductor placed at 1000 km depth. Since these longer periods represent weaker induced fields, this approximation should be adequate. To summarize, the time variability of the primary ionospheric field is represented as 24-, 12-, 8- and 6-hr periodicities modulated by annual and semi-annual periodicities and  $F_{10.7}$  while the secondary field is represented as 24-, 12-, 8- and 6-hr periodicities and an infinite conductor at depth to approximated long-period variations.

Two toroidal field expansions are included, one for Ørsted centred at 750 km and one for CHAMP centred at 400 km. CM4 included the same for Ørsted, but included two other fields for Magsat dawn and dusk data, which are not used in CM5. It did not model the toroidal field in the CHAMP sampling shell since only scalar data were used.

### 3.1.4 Magnetospheric and induced fields

The treatment of the magnetospheric and associated induced fields in CM5 significantly departs from that of CM4. In the latter, the time variability of the magnetosphere was a linear function of the  $D_{st}$  index modulated by seasonal and diurnal oscillations. It is known, however, that the  $D_{st}$  baseline changes with time, making it difficult to use in modelling (Olsen *et al.* 2014). With the advent of *Swarm*, satellite

measurements can be made simultaneously at different local times allowing for the possibility of improved estimates of this time variation over those from observatory data alone, like  $D_{st}$  or the improved  $RC$  index (Olsen *et al.* 2014). The CI algorithm thus parametrizes the magnetosphere and associated induced fields by discretizing time into bins within which the fields are treated as static external and internal SH expansions in dipole coordinates, respectively. Although the local time coverage afforded by *Swarm* is superior to that of the other satellite data sets, nonetheless this treatment has been used in CM5 for very simple geometries by solving for degree  $N_{\max} = 1$  internal and external fields in 1-hr bins over the time envelope of the model.

### 3.1.5 VFM alignment

Along with the geophysical parameters just discussed, CM5 includes a parametrization describing the alignment between the VFM and the CRF systems for both CHAMP and Ørsted. For CHAMP, these rotations consist of three positive counter-clockwise Euler angles of type (XYZ) that rotate from the CRF to the VFM systems such that (Olsen *et al.* 2013; Sabaka *et al.* 2013)

$$\mathbf{R}_{\text{VFM} \leftarrow \text{CRF}} = \mathbf{R}_z(\gamma) \mathbf{R}_y(\beta) \mathbf{R}_x(\alpha), \quad (7)$$

$$= \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}. \quad (8)$$

For Ørsted, these rotations consist of three positive counter-clockwise Euler angles of type (YZY) such that

$$\mathbf{R}_{\text{VFM} \leftarrow \text{CRF}} = \mathbf{R}_y(\gamma) \mathbf{R}_z(\beta) \mathbf{R}_y(\alpha), \quad (9)$$

$$= \begin{pmatrix} \cos \gamma & 0 & \sin \gamma \\ 0 & 1 & 0 \\ -\sin \gamma & 0 & \cos \gamma \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta & 0 \\ \sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{pmatrix}. \quad (10)$$

The (YZY) type is used since  $\beta \approx 90^\circ$ , which makes it more numerically stable than the (XYZ) type.

The observation equations involving vector magnetometer measurements are expressed in the CRF in order to reduce the level of non-linearity in the system (Sabaka *et al.* 2013) and so the data are rotated into the CRF using the transposes of the rotations shown in eqs (7) and (9). For CHAMP VFM measurements, Euler angles are discretized into 10-d bins starting on 2000 August 1 within which they are treated as static. A single set of static Euler angles are used for the Ørsted VFM data.

## 3.2 Estimation procedure

Given the set of model parameters just described, the focus now shifts to how to best estimate them. Both Sabaka & Olsen (2006) and Sabaka *et al.* (2013) present a constrained LS estimation that is realized through the application of the iterative Gauss–Newton (GN) method (Seber & Wild 2003). For CM5, the following LS problem with linear constraints (LSLE) is solved at the  $k$ th GN step (LSLE-GN)

$$\text{LSLE-GN} \left\{ \begin{array}{l} \min_{\Delta \mathbf{x}_k} \quad \left| \mathbf{L}_k^+ (\Delta \mathbf{d}_k - \mathbf{A}_k \Delta \mathbf{x}_k) \right|_2^2 + \sum_{j=1}^{N_q} \lambda_j \left| \mathbf{F}_j^{-1} (\mathbf{x}'_j - \mathbf{x}_k - \Delta \mathbf{x}_k) \right|_2^2 \\ \text{subject to :} \quad \mathbf{G} \Delta \mathbf{x}_k = -\mathbf{G} \mathbf{x}_k, \\ \mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}_k \end{array} \right. , \quad (11)$$

where  $\|\cdot\|_2$  is the  $\ell_2$  norm,  $\Delta \mathbf{d}_k \equiv \Delta \mathbf{d}(\mathbf{x}_k) = \mathbf{d} - \mathbf{a}(\mathbf{x}_k)$  are the residuals of the data vector  $\mathbf{d}$  with respect to the non-linear model vector  $\mathbf{a}(\mathbf{x}_k)$  evaluated at  $\mathbf{x}_k$ ,  $\mathbf{A}_k \equiv \mathbf{A}(\mathbf{x}_k)$  is the Jacobian of the model vector evaluated at  $\mathbf{x}_k$ ,  $\Delta \mathbf{x}_k$  are the adjustments to the current parameter vector  $\mathbf{x}_k$ ,  $\mathbf{L}_k \equiv \mathbf{L}(\mathbf{x}_k)$  is the square-root factor of the data noise covariance matrix  $\mathbf{C}_k = \mathbf{L}_k \mathbf{L}_k^T$  and  $\mathbf{F}_j$  is the square-root factor of the  $j$ th *a priori* covariance matrix  $\mathbf{P}_j^{-1} = \mathbf{F}_j \mathbf{F}_j^T$  that, along with the Lagrange multiplier  $\lambda_j$ , specifies the deviation of the solution from the preferred *a priori* model vector  $\mathbf{x}'_j$ . The matrix  $\mathbf{L}_k^+$  is the pseudo-inverse of  $\mathbf{L}_k$  such that  $\mathbf{L}_k \mathbf{L}_k^+ \mathbf{L}_k = \mathbf{L}_k$  (see Toutenburg 1982) and is relevant when  $\mathbf{C}_k$  assigns infinite variance to certain subspaces. The  $\mathbf{L}_k$  matrix is constructed at each iteration as part of a robust estimation scheme to mitigate the effects of outliers and will be described in greater detail in Section 3.3.3. The system is subject to the linear equality constraints

$$\mathbf{G} \mathbf{x} = \mathbf{0}, \quad (12)$$

where the rank of  $\mathbf{G}$  is less than  $\dim \mathbf{x}$ . The inclusion of such constraints will be discussed in Section 3.4.



The solution to LSLE-GN, denoted  $\widetilde{\Delta \mathbf{x}}_k$ , may be found through Lagrange multiplier theory (Toutenburg 1982; Golub & van Loan 1989; Bertsekas 1999) to be

$$\widetilde{\Delta \mathbf{x}}_k = \overline{\Delta \mathbf{x}}_k - \mathbf{E}_k^{-1} \mathbf{G}^T (\mathbf{G} \mathbf{E}_k^{-1} \mathbf{G}^T)^{-1} \mathbf{G} (\mathbf{x}_k + \overline{\Delta \mathbf{x}}_k), \quad (13)$$

where  $\mathbf{E}_k = \mathbf{A}_k^T \mathbf{C}_k^+ \mathbf{A}_k + \sum_{j=1}^{N_q} \lambda_j \mathbf{P}_j$ , and  $\overline{\Delta \mathbf{x}}_k$  is the unconstrained solution given by

$$\overline{\Delta \mathbf{x}}_k = \mathbf{E}_k^{-1} \left[ \mathbf{A}_k^T \mathbf{C}_k^+ \Delta \mathbf{d}_k + \sum_{j=1}^{N_q} \lambda_j \mathbf{P}_j (\mathbf{x}'_j - \mathbf{x}_k) \right]. \quad (14)$$

In practice, the  $N_q$  quadratic norms in eq. (11) are smoothing terms whose preferred solutions are zero, that is,  $\mathbf{x}'_j = \mathbf{0}$ . These will also be discussed in Section 3.4. From Table 1 it can be seen that  $\dim \mathbf{x} = 702\,308$ . However, since the magnetospheric/induced fields are treated as uncorrelated between bins, a significant part of  $\mathbf{A}_k$  is sparse, which is exploited by the algorithm.

### 3.3 Formal error treatment

A critical element in solving LSLE-GN is a correct specification of  $\mathbf{C}$  or rather  $\mathbf{C}^+$ . Because the near-Earth magnetic field is so complicated in both space and time, the zero-mean noise assumption of LS is often violated. It can be shown that the Mean-Squared-Error (MSE) of an estimate  $\tilde{\mathbf{x}}$  with respect to its true value  $\mathbf{x}$  can be decomposed into

$$\text{MSE}(\tilde{\mathbf{x}}) = E[|\tilde{\mathbf{x}} - \mathbf{x}|_2^2], \quad (15)$$

$$= \text{Tr}[\mathbf{C}_{\tilde{\mathbf{x}}-\tilde{\mathbf{x}}}] + |\mathbf{b}|_2^2, \quad (16)$$

where  $E[\cdot]$  and  $\text{Tr}[\cdot]$  are the expectation and trace operators, respectively,  $\tilde{\mathbf{x}} = E[\tilde{\mathbf{x}}]$  is the expected value of  $\tilde{\mathbf{x}}$ ,  $\mathbf{C}_{\tilde{\mathbf{x}}-\tilde{\mathbf{x}}}$  is the error-covariance matrix and  $\mathbf{b} = \tilde{\mathbf{x}} - \mathbf{x}$  is the bias vector. If the violation is extreme, then often times the second term dominates, leading to severely biased solutions.

#### 3.3.1 Bias mitigation via SIVW

Sabaka *et al.* (2013) introduced a bias mitigation scheme, known as SIVW, for certain classes of noise known as non-zero mean systematic noise  $\boldsymbol{\zeta} = \mathbf{B}\boldsymbol{\xi}$ , where  $\mathbf{B}$  has more rows than columns and  $\boldsymbol{\xi}$  is a Gaussian random vector of mean  $\boldsymbol{\mu}$  and covariance  $\mathbf{C}_\xi$ , denoted as  $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C}_\xi)$ . Thus a shorter non-zero mean random vector  $\boldsymbol{\xi}$  induces a longer non-zero mean random vector  $\boldsymbol{\zeta}$  systematically through  $\mathbf{B}$ . This type of noise is actually quite common and can represent a contaminated set of Gauss coefficients that map into a particular set of measurements through a Jacobian matrix  $\mathbf{B}$ . Thus, SIVW is applied so that data sets with high signal-to-noise (SNR) ratios for certain magnetic sources are exploited to enhance the recovery of the parameters describing these sources. This leads to dense noise covariance matrices  $\mathbf{C}$  that are independent of  $\mathbf{C}_\xi$ , have infinite variance in the column span of  $\mathbf{B}$ , and whose corresponding weight matrices  $\mathbf{C}^+$  are pseudo-inverses.

For example, consider the following set of observation equations

$$\mathbf{d} = \mathbf{A}\mathbf{x} + \mathbf{v}, \quad (17)$$

in which a vector of parameters  $\mathbf{x}$  are linearly related to a vector of measurements  $\mathbf{d}$  through a matrix  $\mathbf{A}$  in the presence of an additive noise term  $\mathbf{v} = \boldsymbol{\zeta} + \boldsymbol{\eta}$ , where  $\boldsymbol{\zeta}$  is the non-zero mean systematic error discussed previously and  $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_\eta)$  is a zero-mean random error independent of  $\boldsymbol{\zeta}$  such that  $\mathbf{v} \sim \mathcal{N}(\mathbf{B}\boldsymbol{\mu}, \mathbf{C}_\eta + \mathbf{B}\mathbf{C}_\xi\mathbf{B}^T)$ . In this case, SIVW specifies the following pseudo-inverse of the data noise covariance matrix

$$\mathbf{C}^+ = \mathbf{C}_\eta^{-1} - \mathbf{C}_\eta^{-1} \mathbf{B} (\mathbf{B}^T \mathbf{C}_\eta^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{C}_\eta^{-1}. \quad (18)$$

Note that since  $\mathbf{C}^+ \mathbf{B} = \mathbf{0}$ , then the LS estimate  $\tilde{\mathbf{x}}$  is unbiased, which would not have been the case had  $\mathbf{C} = \mathbf{C}_\eta + \mathbf{B}\mathbf{C}_\xi\mathbf{B}^T$  been chosen.

SIVW has been applied at three levels in the derivation of CM5. The top level discriminates between measurements collected during quiet and disturbed magnetic times as defined in Section 2. Typically, models are determined exclusively from quiet-time data for obvious reasons, but since CM5 is meant to be a continuous representation of the near-Earth field from 2000 August to 2013 January disturbed times must be included in order to estimate the magnetospheric and induced fields in all 1-hr bins. At the next level, the quiet-time data are divided into ‘dayside’ and ‘nightside’ groups, where the latter is defined as a position during which the subsolar point is more than  $10^\circ$  below the horizon (Olsen *et al.* 2014). This distinction is required in order to best model sources such as the core and lithosphere from nightside data when ionospheric *E*-region activity is minimal. For ionospheric, toroidal, magnetospheric and associated induced sources, data at all local times are needed for accurate modelling. The final application involves the use of CHAMP along-track differences of the measured magnetic field vectors to enhance the recovery of the high degree ( $n > 60$ ) lithospheric field. This difference operation can substantially eliminate transient, broad-scale external field contamination that could otherwise map into the crustal field estimate. However, the complementary data, the along-track vector sums, are also useful, and so SIVW is used in order to retain them. The difference/summation operations are only done

on the observation equations corresponding to CHAMP vector data being no more than 30 s apart. This includes the partial derivatives for the Euler angle series. The operations are carried out by the following orthogonal transformation

$$\begin{pmatrix} \mathbf{v}_- \\ \mathbf{v}_+ \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -\mathbf{R}_{\text{MP} \leftarrow \text{CRF}_A} & \mathbf{R}_{\text{MP} \leftarrow \text{CRF}_B} \\ \mathbf{R}_{\text{MP} \leftarrow \text{CRF}_A} & \mathbf{R}_{\text{MP} \leftarrow \text{CRF}_B} \end{pmatrix} \begin{pmatrix} \mathbf{v}_A \\ \mathbf{v}_B \end{pmatrix}, \quad (19)$$

where ‘A’ and ‘B’ denote the two CHAMP measurement times,  $\mathbf{v}_A$  and  $\mathbf{v}_B$  represent generic vector quantities in the  $\text{CRF}_A$  and  $\text{CRF}_B$  systems, respectively, which are rotated into the midpoint position between ‘A’ and ‘B’ via  $\mathbf{R}_{\text{MP} \leftarrow \text{CRF}_A}$  and  $\mathbf{R}_{\text{MP} \leftarrow \text{CRF}_B}$ , respectively, and  $\mathbf{v}_-$  and  $\mathbf{v}_+$  are the differences and sums at the midpoint, respectively. The noise covariances are also transformed in a consistent manner (see Sabaka *et al.* 2013).

Because  $\mathbf{C}^+$  can be large and dense, manipulating it may be difficult. However, Sabaka *et al.* (2013) have provided an alternative mechanism for applying SIVW through the introduction of ‘nuisance’ parameters that only requires the covariance of the zero-mean part of the noise, which is often of much simpler structure. Returning to the previous example, if eq. (17) is rewritten as

$$\mathbf{d} = (\mathbf{A} \quad \mathbf{B}) \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + \boldsymbol{\eta}, \quad (20)$$

then the LS solution for  $\mathbf{x}$  is the same as that for the system in eq. (17), provided  $\mathbf{C}_\eta$  is the specified noise covariance matrix. That is, the introduction of a set of deterministic parameters  $\mathbf{y}$  and the use of a simple  $\mathbf{C}^+ = \mathbf{C}_\eta^{-1}$  gives the same LS solution  $\hat{\mathbf{x}}$  as that of eq. (17) when using the complicated  $\mathbf{C}^+$  of eq. (18). The parameters in the vector  $\mathbf{y}$  are referred to as ‘nuisance’ parameters and are typically not of interest in themselves, but merely serve as a framework through which  $\mathbf{C}^+$  of eq. (18) may implicitly act.

The practical implementation of SIVW via nuisance parameters can be seen by comparing the original and modified set of observation equations underlying the  $k$ th LSLE-GN step in eq. (11). The original set is given by  $\Delta \mathbf{d} = \mathbf{A} \Delta \mathbf{x} + \mathbf{v}$ , where the subscript ‘ $k$ ’ has been omitted for clarity and where non-zero mean systematic error is implied in  $\mathbf{v}$ . The modified set used in deriving CM5 is

$$\begin{pmatrix} \Delta \mathbf{d}_{qnocd} \\ \Delta \mathbf{d}_{qn} \\ \Delta \mathbf{d}_{qd} \\ \Delta \mathbf{d}_d \end{pmatrix} = \begin{pmatrix} \mathbf{A}_{qnocd}^u & \mathbf{A}_{qnocd}^v & \mathbf{A}_{qnocd}^w & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{qn}^u & \mathbf{A}_{qn}^v & \mathbf{0} & \mathbf{0} & \mathbf{A}_{qn}^w & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_{qd}^u & \mathbf{0} & \mathbf{0} & \mathbf{A}_{qd}^v & \mathbf{A}_{qd}^w & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_d^v & \mathbf{A}_d^z \end{pmatrix} \begin{pmatrix} \Delta \mathbf{u} \\ \Delta \mathbf{v} \\ \Delta \mathbf{w} \\ \Delta \mathbf{v}_n \\ \Delta \mathbf{w}_n \\ \Delta \mathbf{y} \\ \Delta \mathbf{z}_n \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_{qnocd} \\ \boldsymbol{\eta}_{qn} \\ \boldsymbol{\eta}_{qd} \\ \boldsymbol{\eta}_d \end{pmatrix}, \quad (21)$$

where  $\Delta \mathbf{d}$  and  $\boldsymbol{\eta}$  are residual and zero-mean error vectors at the current state, with ‘ $qnocd$ ’, ‘ $qn$ ’, ‘ $qd$ ’ and ‘ $d$ ’ denoting quiet-time, nightside OHMs and CHAMP vector differences; all other quiet-time, nightside data (including CHAMP vector sums); all quiet-time, dayside data; and all disturbed-time data, respectively. For the quiet-time data, the vector  $\Delta \mathbf{v}$  is a set of core/lithospheric parameter adjustments for  $n \leq 60$ , including OHM biases and  $M_2$  tidal parameters;  $\Delta \mathbf{w}$  is a vector of lithospheric parameter adjustments for  $n > 60$ ; and  $\Delta \mathbf{u}$  is a vector of all other parameter adjustments, while ‘ $n$ ’ denotes the nuisance set. For the disturbed-time data, the vector  $\Delta \mathbf{y}$  is the set of magnetospheric and associated induced parameters, and  $\Delta \mathbf{z}_n$  is a vector of all other parameter adjustments, which constitute nuisance fields. The matrices  $\mathbf{A}_i^j$  are the Jacobians at the current state corresponding to the  $i$ th data type and the  $j$ th parameter type. Only the nominal parameter sets ( $\Delta \mathbf{u}$ ,  $\Delta \mathbf{v}$ ,  $\Delta \mathbf{w}$  and  $\Delta \mathbf{y}$ ) constitute the original  $\Delta \mathbf{x}$  in eq. (11) and are therefore used for evaluation at each LSLE-GN step. Because  $\Delta \mathbf{y}$  and  $\Delta \mathbf{z}_n$  are completely decoupled from the other parameter sets, since they are determined by disturbed-time data only, they can be solved for separately at each step. Clearly the matrices and vectors in eq. (11) can be appropriately padded in order to accommodate the nuisance parameter vectors. Note that the number of parameters listed in Table 1 for each source/effect is the nominal value; the system in eq. (21) is substantially larger, but its sparse structure is exploited by the algorithm.

### 3.3.2 Zero-mean error treatment

What remains is to define the covariances associated with each of the zero-mean errors  $\boldsymbol{\eta}$  in eq. (21). Beginning with the simplest case, the  $3 \times 3$  OHM noise covariance matrix is expressed as  $\mathbf{C}_{\text{OHM}} = \sigma_{\text{OHM}}^2 \mathbf{I}$ , where  $\sigma_{\text{OHM}}^2$  is a function of dipole latitude with polar stations having higher variances than lower latitude stations. For CM5, stations poleward of  $\pm 50^\circ$  dipole latitude were assigned  $\sigma_{\text{OHM}} = 15$  nT while those in between were assigned  $\sigma_{\text{OHM}} = 7$  nT. Thus the noise is treated as isotropic and uncorrelated between vector components. Satellite scalar data noise is also treated as uncorrelated with other measurements noise and its variance is denoted as  $\sigma_F^2$ . For both Ørsted and SAC-C,  $\sigma_F = 4$  nT.

For satellite vector measurements, the covariance has two terms, one to account for CRF attitude error and one to account for isotropic noise (Holme 2000) such that at time ‘A’ the  $3 \times 3$  covariance matrix is

$$\mathbf{C}_{AA} = \sigma_F^2 \mathbf{I} + \mathbf{C}_{\text{CRF}_A}(\mathbf{x}_k). \quad (22)$$

Note that  $\sigma_F^2$  is the same as is used for scalar data. For CHAMP this is  $\sigma_F = 3$  nT. The second term, which addresses the attitude error, is constructed in accordance with Sabaka *et al.* (2013) and decomposes the rotation from the inertial frame, say J2000, to the CRF as

$$\mathbf{R}_{\text{CRF} \leftarrow \text{J2000}} = \mathbf{R}_2(\chi) \mathbf{R}_3(\delta) \mathbf{R}_2(\lambda), \quad (23)$$

where  $\delta$  and  $\lambda$  are the colatitude and longitude of the J2000 Z-axis in the CRF, and  $\chi$  is the rotation around the Z-axis. The covariance due to errors in this rotation are determined by propagating the covariance matrix of angular perturbations, which is assumed diagonal. This propagation is a function of the current parameter state  $\mathbf{x}_k$  and so  $\mathbf{C}_{\text{CRFA}}(\mathbf{x}_k)$  changes with each GN iteration. The errors  $(\sigma_\lambda, \sigma_\delta, \sigma_\chi)$  are (10, 10, 10) arcsecs and (10, 10, 40) arcsecs for CHAMP measurements when multiple or single star-imager heads are in use, respectively. For Ørsted this is always (10, 10, 40) arcsecs. The covariances for CHAMP vector differences and sums are propagated from the covariances at two times ‘A’ and ‘B’ in accordance with eq. (19), although as mentioned in Sabaka *et al.* (2013), the block off-diagonal parts of these covariances have been omitted in the current algorithm.

### 3.3.3 Robust estimation

One of the assumptions of classic LS theory is that the errors follow a Gaussian distribution. If this is correct and the true  $\mathbf{C}$  is used, then the solution would provide the maximum-likelihood estimate. However, this is often not the case and the estimate then suffers from the excessive influence of outliers. To combat this, a robust estimation procedure known as ‘Iteratively Reweighted Least-Squares’ (IRLS) with Huber weighting has been employed here (Constable 1988). It has been used successfully in deriving the CM4 model (Sabaka *et al.* 2004) and is described in the context of the CI algorithm in Sabaka *et al.* (2013) where it was applied to synthetic *Swarm* data. IRLS operates on scalar data and assigns Huber weights to the  $i$ th measurement at the  $k$ th GN iteration as a function of its standard deviation  $\sigma_i$  and the current residual  $e_{i,k}$  according to

$$w_{i,k} = \frac{1}{\sigma_i^2} \min \left( \frac{c\sigma_i}{|e_{i,k}|}, 1 \right), \quad (24)$$

where the underlying Huber distribution is defined as having a Gaussian core for  $|e_{i,k}| \leq c\sigma_i$ , and Laplacian tails (Constable 1988). A value of  $c = 1.5$  is used by the CI algorithm.

Both the OHM and satellite scalar data conform to this structure. However, for satellite vector data IRLS is applied to each of the principle components of the covariance matrix  $\mathbf{C}_{\text{AA}}$  in eq. (22). Since the transformation to vector differences and sums is orthogonal in eq. (19), the IRLS can still be applied to the principle components of  $\mathbf{C}_{\text{AA}}$  and  $\mathbf{C}_{\text{BB}}$  corresponding to the two CHAMP measurement times.

## 3.4 Regularization and *a priori* information

While much effort has been given to providing a realistic error assessment and treatment in the estimation procedure, it is also important to assess the observability of the model parameters given the available measurements. In situations where these measurements are not sufficient to discriminate between parameters or when the theory itself poses non-uniqueness, then a regularization procedure must be used. There are two types invoked for CM5, the linear equality constraints and the quadratic constraints. Table 2 lists the Lagrange multiplier, or damping parameter,  $\lambda_j$ , values for each quadratic smoothing norm  $\mathcal{P}_j$  used in the derivation of CM5, where  $j$  indicates what is being smoothed.

### 3.4.1 Linear equality constraints

Recall from Section 3.2 that linear equality constraints in the form of eq. (12) have been considered in the estimation procedure and are manifested in the LSLE-GN step in eq. (11). The reason for this becomes apparent when comparing the parametrizations of the core SV and the induced fields associated with the magnetosphere. Both are represented as internal SH expansions whose time-variability is via B-splines in the former and piecewise constants over 1-hr bins in the latter. It should be clear that as these bin durations decrease, the piecewise constants will be better able to describe the B-spline variations resulting in co-linearity between the basis functions of the two field sources. To alleviate this problem, Sabaka & Olsen (2006) and Sabaka *et al.* (2013) introduced a set of linear equality constraints that force each induced SH time-series to be orthogonal to each core SV temporal basis function, including the constant, through time. This means that the induced field is orthogonal to the core field through time at an ever spatial position. Because the core spline basis is broad-scale in time, the constrained induced field will reflect more high-frequency behaviour in the internal field, which is sufficient for much of the expected induced structure. The details of the  $\mathbf{G}$  matrix may be found in Sabaka & Olsen (2006) and Sabaka *et al.* (2013).

If  $\mathcal{P}(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x}$  is a quadratic smoothing norm such that  $\mathbf{P} = \mathbf{G}^T \mathbf{G}$ , then it can be shown that as the associated damping parameter  $\lambda \rightarrow \infty$ , it transforms into the set of linear equality constraints seen in eq. (12). For this reason it is included in Table 2 as a quadratic norm  $\mathcal{P}(|\mathbf{p}_{\text{ind} \perp \text{core}}|_2^2)$  whose damping parameter value is  $\infty$ . Note that this only makes sense in the case when  $\mathbf{G}$  has a rank less than its column number, otherwise,  $\mathbf{P}$  would be full-rank and the norm would force the estimate  $\hat{\mathbf{x}} = \mathbf{0}$ . As it stands,  $\hat{\mathbf{x}}$  resides in the span of the null-space of  $\mathbf{G}$ .

**Table 2.** CM5 damping parameter values.

Norm	Damping parameter ( $\lambda$ )	
Core/lithosphere		
$\mathcal{P}\left( \ddot{B}_r ^2\right)$	$1.0 \times 10^0$	$(\text{nT} \cdot \text{yr}^{-2})^{-2}$
$\mathcal{P}\left( B_{r,n>60} ^2\right)$	$3.0 \times 10^{-1}$	$(\text{nT})^{-2}$
$\mathcal{P}\left( B_{r,n>60,\text{caps}} ^2\right)$	$1.0 \times 10^2$	$(\text{nT})^{-2}$
Ionosphere		
$\mathcal{P}\left( \mathbf{J}_{\text{eq,MLT:21-05}} _2^2\right)$	$1.0 \times 10^4$	$(\text{A} \cdot \text{km}^{-1})^{-2}$
$\mathcal{P}\left( \nabla_s^2 \mathbf{J}_{\text{eq},p>0,\text{mid-lat}} _2^2\right)$	$1.0 \times 10^{-2}$	$(\text{A} \cdot \text{km}^{-3})^{-2}$
CHAMP toroidal		
$\mathcal{P}\left( \nabla_s^2 J_r ^2\right)$	$1.0 \times 10^{-10}$	$(\text{nA} \cdot \text{m}^{-4})^{-2}$
$\mathcal{P}\left( \nabla_s^2 \mathbf{J}_{h,\text{mid-lat}} _2^2\right)$	$1.0 \times 10^0$	$(\text{nA} \cdot \text{m}^{-4})^{-2}$
Ørsted toroidal		
$\mathcal{P}\left( \nabla_s^2 J_r ^2\right)$	$1.0 \times 10^{-10}$	$(\text{nA} \cdot \text{m}^{-4})^{-2}$
$\mathcal{P}\left( \nabla_s^2 \mathbf{J}_{h,\text{mid-lat}} _2^2\right)$	$1.0 \times 10^0$	$(\text{nA} \cdot \text{m}^{-4})^{-2}$
Magnetosphere/induced		
$\mathcal{P}\left( \mathbf{p}_{\text{mag/ind}} _2^2\right)$	variable	$(\text{nT})^{-2}$
$\mathcal{P}\left( \mathbf{p}_{\text{ind}\perp\text{core}} _2^2\right)$	$\infty$	$(\text{nT})^{-2}$

### 3.4.2 Quadratic constraints

Several quadratic smoothing norms were minimized along with the weighted residual term in eq. (11). The first was used to control the core SV by minimizing the mean-square  $\ddot{B}_r$  component over the core–mantle boundary (CMB) at radius 3480 km from 2000.5 to 2013.0, denoted as  $\mathcal{P}(|\ddot{B}_r|^2)$  in Table 2. This was also used in deriving CM4 (Sabaka *et al.* 2004), but, for instance, the third time derivative of  $B_r$  was damped in the CHAOS4 model (Olsen *et al.* 2014). The high-degree lithosphere ( $n > 60$ ) was smoothed in two ways: first, by minimizing the mean-square  $B_r$  component over the Earth’s mean surface at 6371.2 km ( $\mathcal{P}(|B_{r,n>60}|^2)$  in Table 2), and secondly, the same quantity, but only over the  $2.7^\circ$  northern and southern polar gaps in the CHAMP orbital coverage ( $\mathcal{P}(|B_{r,n>60,\text{caps}}|^2)$  in Table 2). If the polar gaps are not accounted for, the field begins to ring above approximately  $n = 60$ . The regularization matrix associated with this quantity was computed by simply integrating the SH products for  $n > 60$  corresponding to  $B_r^2$  over the two polar caps using Gauss–Legendre quadrature (Atkinson 1989). This type of regularization could also have been applied using Slepian functions (Simons *et al.* 2006) or localized functions (Lesur 2006). Olsen *et al.* (2014) also damps  $B_r$  for  $n > 60$  over the Earth’s surface, but uses synthetic data to control polar gap problems.

Nighttime ionospheric  $E$ -region currents are minimized by a norm, denoted as  $\mathcal{P}(|\mathbf{J}_{\text{eq,MLT:21-05}}|_2^2)$  in Table 2, which measures the mean-square magnitude of the  $E$ -region equivalent currents  $\mathbf{J}_{\text{eq}}$  flowing at 110 km altitude over the nighttime sector, defined as magnetic local time (MLT) 21:00–05:00 hours, through the year. In addition, these currents are further smoothed by a norm, denoted as  $\mathcal{P}(|\nabla_s^2 \mathbf{J}_{\text{eq},p>0,\text{mid-lat}}|_2^2)$ , which measures the mean-square magnitude of the surface Laplacian of the diurnally varying portion of  $\mathbf{J}_{\text{eq}}$  at mid-latitudes at all local times. These two norms were also used in deriving CM4 (Sabaka *et al.* 2004).

For ionospheric  $F$ -region poloidal currents, CM5 follows closely to what was done for CM4. The surface Laplacian of both the radial component,  $J_r$ , and the horizontal components,  $\mathbf{J}_h$ , were damped separately on spheres at 400 and 750 km altitude for CHAMP and Ørsted, respectively, over all local times. Furthermore,  $\mathbf{J}_h$  was damped more heavily at mid-latitudes. These are denoted as  $\mathcal{P}(|\nabla_s^2 J_r|^2)$  and  $\mathcal{P}(|\nabla_s^2 \mathbf{J}_{h,\text{mid-lat}}|_2^2)$  in Table 2. Although  $J_r$  was also damped directly in the Ørsted sampling shell in CM4, this was not done here for Ørsted or CHAMP since the improvement in data coverage has stabilized the solution.

Finally, for the magnetospheric and, by virtue of the  $\mathcal{P}(|\mathbf{p}_{\text{ind}\perp\text{core}}|_2^2)$  norm discussed in the previous section, associated high-frequency induced fields, one additional type of regularization was used. This norm minimizes the Euclidian ( $\ell_2$ ) length of the magnetospheric/induced coefficients in each bin and is denoted as  $\mathcal{P}(|\mathbf{p}_{\text{mag/ind}}|_2^2)$  in Table 2. The determination of the value of the damping parameter is automated and can vary for each bin, as indicated in Table 2. When both OHM and satellite vector data are present in a bin, the procedure first solves

$$\mathbf{E}\mathbf{s} = \mathbf{E}\mathbf{1} \quad (25)$$

for  $\mathbf{s}$ , where  $\mathbf{E}$  is the normal matrix corresponding to the magnetospheric/induced coefficients plus the contribution from  $\mathcal{P}(|\mathbf{p}_{\text{mag/ind}}|_2^2)$ , and  $\mathbf{1}$  is a vector with elements of unity. If there are no numerical precision issues in solving eq. (25), then  $\mathbf{s} = \mathbf{1}$ . However, due to the data type and distribution within the bin and the current level of regularization this can be inaccurate and thus requires increased regularization. Therefore, the procedure checks the following condition:

$$|\mathbf{1} - \mathbf{s}|_\infty < \text{tol}, \quad (26)$$

where  $\|\cdot\|_\infty$  is the  $\ell_\infty$  norm and  $tol$  is a tolerance, which in this case is chosen as  $tol = 10^{-15}$ . It begins by setting  $\lambda = 0$  and then increases it by a factor of 10 and checks until eq. (26) is satisfied. If neither type of vector data are available, then  $\lambda$  is set to  $0.01 \times Tr[\mathbf{E}] / \dim \mathbf{E}$ . This automation is useful because the large number of bins and the highly varying quality and quantity of data found in each make the predetermination of  $\lambda$  intractable.

## 4 RESULTS AND DISCUSSION

The application of the CI algorithm leading to CM5 consisted of four iterations of eq. (11) for the quiet-time data followed by a single iteration for the disturbed-time data. All other quantities involved in LSLE-GN have been defined in previous sections. In order to assess the efficacy of using CHAMP along-track vector differences to exploit gradient information, an additional model was derived, which is called ‘CM5 no-gradients’ (CM5ng), that did not use CHAMP gradient information, but rather treated the vector data separately at each time instant. Similarly, in order to assess the influence of using a 3-D versus 1-D conductivity structure for ionospheric induction, an additional model called ‘CM5 ionospheric-1d’ (CM5i1d) was derived using a 1-D ionospheric conductivity profile.

### 4.1 Residual statistics

Rather than computing the usual arithmetic mean and rms of the residuals, statistics were desired that more closely reflect the objectives of the estimator, especially given that robust estimation in the form of IRLS has been used to mitigate the influence of outliers. The weighted data misfit term in eq. (11) is approximately  $\mathbf{e}^T \mathbf{W} \mathbf{e}$ , where  $\mathbf{e} = \mathbf{d} - \mathbf{a}(\hat{\mathbf{x}})$  is the residual vector with respect to the final model  $\hat{\mathbf{x}}$  and  $\mathbf{W} = \mathbf{C}^{-1}$ . It can be expressed as  $(\text{RSS})^2 = \sum_i w_i e_i^2$ , where RSS is the weighted ‘root-sum-of-Squares’ and  $w_i$  are the eigenvalues of  $\mathbf{W}$  and  $e_i$  is the  $i$ th component of  $\mathbf{e}$  in the principle component system of  $\mathbf{W}$ . In the case of CM5,  $e_i$  and  $w_i$  are the  $i$ th residual and Huber weight for a particular component, respectively. Therefore, the weighted residual mean and rms of that component are defined as

$$\text{Mean} = \sum_i w_i e_i / \sum_i w_i, \quad (27)$$

$$(\text{rms})^2 = \sum_i w_i e_i^2 / \sum_i w_i. \quad (28)$$

Table 3 lists the *a posteriori* weighted residual mean and rms for four statistical groups corresponding to the data subsets defined in Section 3.3.1 in eq. (21) and denoted as ‘qnocd’, ‘qn’, ‘qd’ and ‘d’. All residuals are computed in their analysis frame, that is, the frame in which the observations equations are processed. For satellite vector residuals, these are the principle component directions of their  $3 \times 3$

**Table 3.** CM5 weighted residual statistics. The four categories of statistics, denoted ‘qnocd’, ‘qn’, ‘qd’ and ‘d’, correspond to the data groupings in eq. (21). The satellite vector residuals are computed in the principle component directions of their  $3 \times 3$  data weight (inverse noise covariance) matrices in the CRF system for CHAMP single-vector and all Ørsted measurements, and in the local spherical system at the midpoint position between the two CHAMP measurements used to make along-track differences and sums. The  $B_L$ ,  $B_M$  and  $B_H$  components correspond to the low, middle and high eigenvalues of the weight matrices for the along-track differences and sums. The  $B_L$ ,  $B_M$  and  $B_B$  components are similarly defined for single-vector data, where  $B_B = B_H$  happens to be in the direction of the predicted  $\mathbf{B}_{CRF}$  vector and the other two are in the plane perpendicular to  $\mathbf{B}_{CRF}$ . The  $F$  component is in the direction of the observed  $\mathbf{B}_{CRF}$  vector.

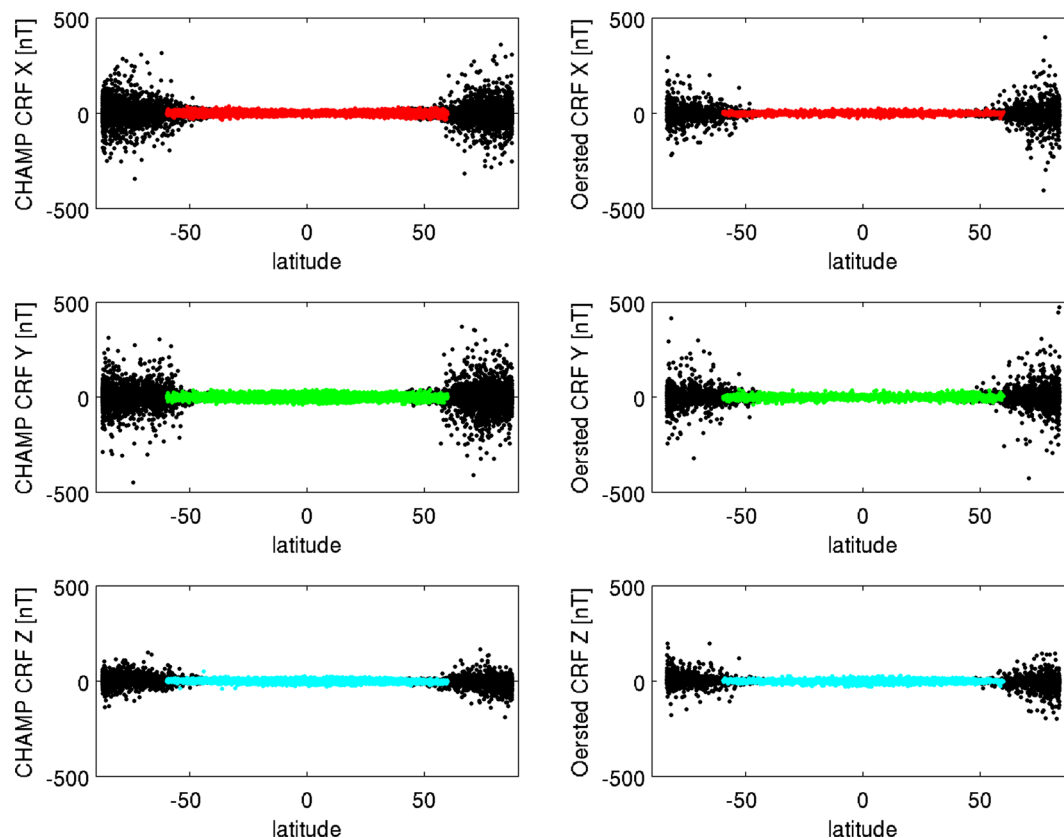
		‘qnocd’			‘qn’			‘qd’			‘d’		
		$N$	Mean	rms	$N$	Mean	rms	$N$	Mean	rms	$N$	Mean	rms
CHAMP	$B_L$	757 230	0.00	5.77	—	—	—	—	—	—	—	—	—
Along-track	$B_M$	757 230	0.00	3.85	—	—	—	—	—	—	—	—	—
Differences	$B_H$	757 230	−0.00	1.65	—	—	—	—	—	—	—	—	—
CHAMP	$B_L$	—	—	—	757 230	−0.04	12.03	—	—	—	—	—	—
Along-track	$B_M$	—	—	—	757 230	−0.02	7.23	—	—	—	—	—	—
Sums	$B_H$	—	—	—	757 230	−0.07	4.26	—	—	—	—	—	—
CHAMP	$B_L$	—	—	—	—	—	—	2 189 882	0.15	12.26	2 212 124	−0.28	39.66
Single	$B_M$	—	—	—	—	—	—	2 189 882	0.23	12.08	2 212 124	0.34	46.74
Vector	$B_B$	—	—	—	—	—	—	2 189 882	−1.31	5.91	2 212 124	−5.34	31.05
Ørsted	$B_L$	—	—	—	323 311	−0.29	12.29	420 455	0.06	17.70	614 052	0.06	31.04
Single	$B_M$	—	—	—	323 311	−0.04	9.41	420 455	0.18	15.64	614 052	0.19	29.12
Vector	$B_B$	—	—	—	323 311	0.45	3.12	420 455	−1.33	5.58	614 052	−5.04	13.47
	$F$	—	—	—	1 307 263	0.57	3.24	1 907 575	−0.46	4.91	1 641 915	−4.20	13.04
SAC-C	$F$	—	—	—	382 501	0.07	3.30	445 123	−0.61	5.44	755 760	−4.62	16.19
OHM	$B_\theta$	2 538 688	0.08	6.55	—	—	—	3 898 998	−0.83	10.54	5 654 876	5.77	22.14
	$B_\phi$	2 538 688	0.02	5.62	—	—	—	3 898 998	−0.31	8.89	5 654 876	0.29	14.72
	$B_r$	2 538 688	−0.04	6.08	—	—	—	3 898 998	0.34	9.45	5 654 876	1.14	17.34

data weight (inverse noise covariance) matrices in the CRF system for CHAMP single-vector and all Ørsted measurements, and in the local spherical system at the midpoint position between the two CHAMP measurements used to make along-track differences and sums. For OHM residuals, the local spherical system ( $B_r, B_\theta, B_\phi$ ) is the analysis frame. For satellite vector residuals these components are labelled  $B_L, B_M$  and  $B_H$  and correspond to the low, middle and high eigenvalues of the weight matrices. For single-vector data these are labelled  $B_L, B_M$  and  $B_B$ , where  $B_B = B_H$  is in the direction of the predicted  $\mathbf{B}_{CRF}$  vector, while the other two components lie in the plane perpendicular to  $\mathbf{B}_{CRF}$ . The  $F$  component is in the direction of the observed  $\mathbf{B}_{CRF}$  vector.

As expected, the rms increases for all data subset components when going from the left to right columns as this indicates a progression in the field complexity from night, to day, to disturbed times. One can also see a general decrease in rms from  $B_L$ , to  $B_M$ , to  $B_H$  or  $B_B$  with the exception of CHAMP single-vector data during disturbed times where  $B_L$  is actually smaller than  $B_M$ , but are both larger than  $B_B$ . This makes sense given that the eigenvalue weighting increases from  $B_L$ , to  $B_M$ , to  $B_H$  or  $B_B$  by definition. The behaviour of the Ørsted  $B_B$  and  $F$  components are also very similar due to the fact that they represent similar directions. It can also be seen that the residuals are smaller for the CHAMP along-track differences than for the sums. This is likely because the sums are more susceptible to broad-scale, highly variable field contamination and is the reason why the entire nominal crust is directly influenced by the differences as opposed to only up to degree 60 for the sums in eq. (21). A peculiarity is seen when comparing residual means of satellite data in that those of the CHAMP along-track sums and differences are very low, but those of the single-vector CHAMP and Ørsted  $B_B$  and Ørsted and SAC-C  $F$  components are relatively large. Even though  $B_B$  and  $F$  residuals exhibit the smallest rms, they exhibit the largest means, and these increase with field complexity. However, the same type of behaviour is seen in the vector residual means of the CHAOS-4I model (same as CHAOS-4, but without the lithospheric field for  $n > 24$ ) of Olsen *et al.* (2014). Finally, the OHM residuals exhibit higher mean and rms values for the  $B_\theta$  and  $B_r$  components, which is clearly a result of magnetospheric influence.

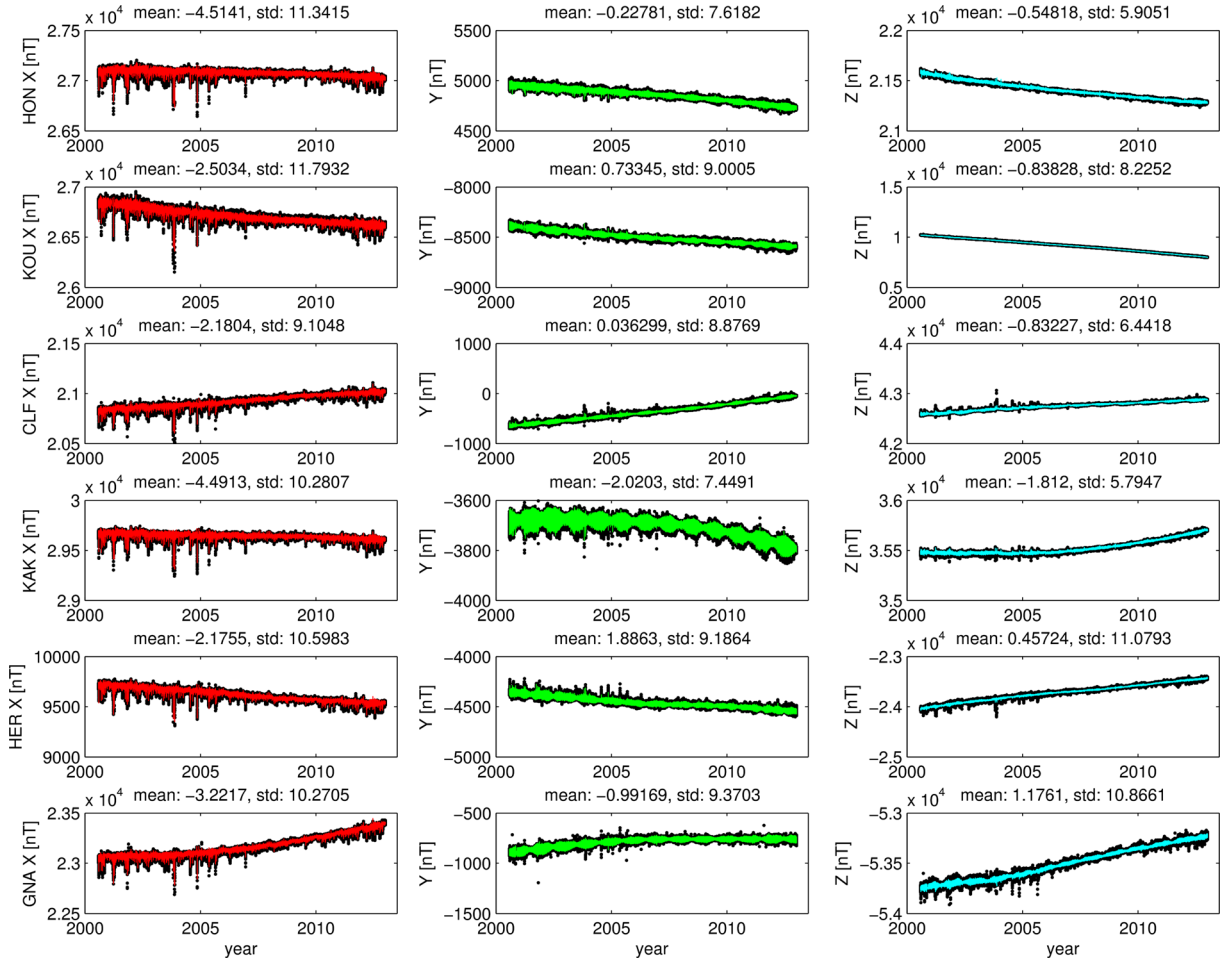
The spatial distributions of satellite vector residuals with respect to geographic latitude are shown for CHAMP and Ørsted in Fig. 4 in the CRF system during quiet-times. These show a clear dichotomy between polar (black) and non-polar (colour) residual behaviour with the former being much larger, as expected, from the presence of high-latitude disturbances from field-aligned currents (FACs) and polar-electrojets (PEJs). The polar CHAMP residuals also appear larger in at least two components which is likely due to the presence of larger disturbances at its lower altitude.

The quality of fits to vector OHMs are shown in Fig. 5 for a selection of low- and mid-latitude stations in the local spherical NED (XYZ) system through time at all magnetic disturbance levels. Both the observed (black) and predicted (colour) field components are shown in each panel with the unweighted residual mean and std listed above each panel. The predictions are generally in very good agreement with the data,



**Figure 4.** Plots of quiet-time CHAMP (left-hand side) and Ørsted (right-hand side) vector residuals with respect to CM5 as a function of latitude in the CRF system for  $X$  (top panels),  $Y$  (middle panels) and  $Z$  (bottom panels) components in nT. The black/coloured symbols indicate positions poleward/equatorward of dipole latitudes  $\pm 50^\circ$ , respectively. Data have been decimated to every 100th point.





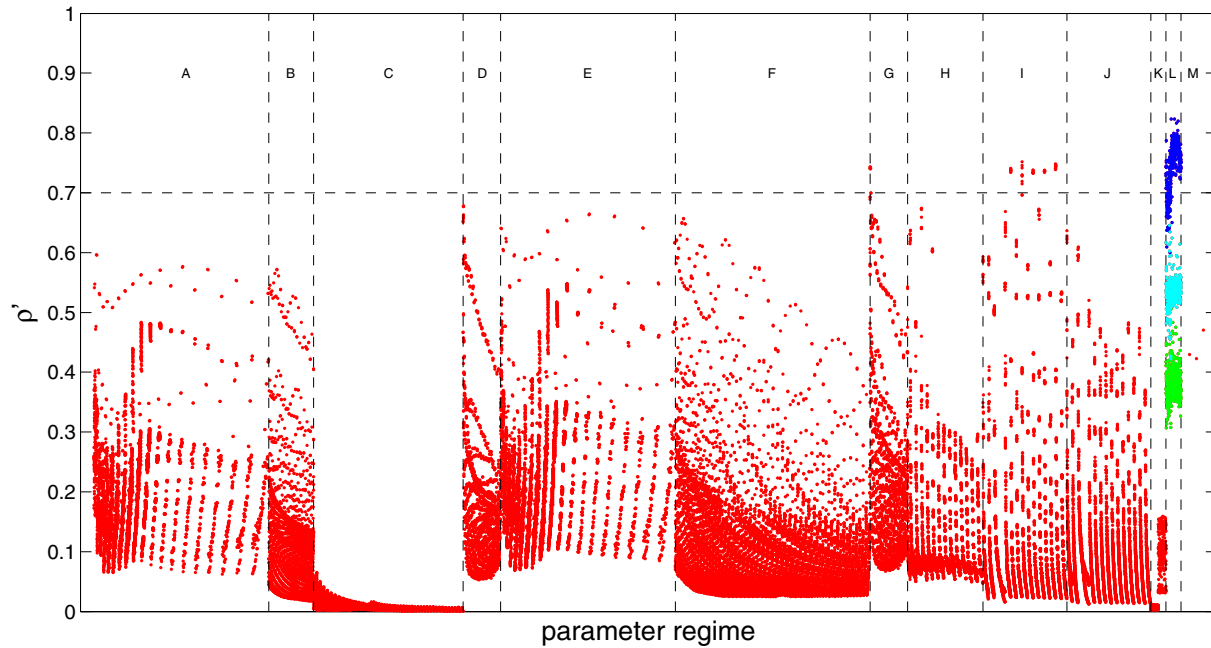
**Figure 5.** Plots of selected OHM vector measurements (black) and predictions from CM5 (colours) in the local spherical NED or XYZ system for  $X$  (left-hand panels),  $Y$  (middle panels) and  $Z$  (right-hand panels) components in nT for all times. The mean and standard deviation (std) of the implied residuals are listed above each panel in nT. From the top, the stations are Honolulu, Hawaii, USA (HON), Kourou, French Guiana (KOU), Chambon la Foret, France (CLF), Kakioka, Japan (KAK), Hermanus, South Africa (HER) and Gngangara, Australia (GNA).

even during disturbed times. However, stations at Kourou, French Guiana (KOU), Hermanus, South Africa (HER) and Gngangara, Australia (GNA) exhibit variance in their  $Z$  components due to coastal induction effects that are not well tracked by CM5. This is not surprising since the degree-one induced field associated with the magnetosphere is likely not of sufficient resolution to capture the signal.

## 4.2 Parameter separability

As mentioned in Section 3.4, of critical importance to any inverse modelling effort is the ability of the estimator to observe and separate the parameters of interest given the measurements and *a priori* information. CM5 has over 700 000 parameters and one should be aware of any separability issues between the numerous field sources that could impact its use. Two separability metrics will be employed, the classic parameter correlation coefficient  $\rho$  and the parameter subspace correlation coefficient  $\rho'$ , both discussed in detail in Appendix A.

Recall that the magnetospheric and associated induced fields are estimated in hourly bins. Allowing for such rapid variations can generate co-linearity with other sources like the lithosphere (Thébault *et al.* 2012; Lesur *et al.* 2013). However, classic correlations involving these parameters are difficult to obtain because, due to their large number, the full error-covariance matrix is never explicitly computed. Instead, their effects are accounted for by creating an equivalent system in block-row echelon form (Anton 1984), which is later backsolved to obtain their updated values. This is exactly the case where  $\rho'$  can provide information about this type of co-linearity. Therefore, the formal variances were extracted from the CM5 error-covariance matrix  $\mathbf{C}$  as well as those from a matrix in which magnetospheric/induced parameters were not considered. The  $\rho'$  coefficients were computed for each non-magnetospheric/induced parameter and are shown in Fig. 6. A value of  $\rho' = 0.7$ , shown as a horizontal dashed line, indicates an intersection angle of about  $45^\circ$  between the basis vector and the magnetospheric/induced subspace and is considered a ‘half-way’ point above which the co-linearity is assumed significant. Interestingly, most parameters, including the nominal lithospheric field, have values much below 0.7, especially the lithosphere for  $n > 60$ . While there are some values above 0.7 in the nuisance  $M_2$  tidal and the CHAMP toroidal field coefficients, the largest and most numerous are found in the CHAMP Euler angle regime, particularly in the  $Y$ -axis rotation shown in blue. Even so, these values are below 0.83 and no obvious degradation is seen in the Euler angle



**Figure 6.** The  $\rho'$  coefficients for each non-magnetospheric/induced parameter in the case when magnetospheric and associated induced parameters are either considered or not. The horizontal dashed line at  $\rho' = 0.7$  indicates the geometric ‘half-way point’ where two vectors intersect at approximately a  $45^\circ$  angle and which is used as a heuristic for deciding that two vectors are more co-linear than not. The vertical dashed lines delineate different parameter regimes: (A) nominal core/SV, (B) nominal crust for  $n \leq 60$ , (C) nominal crust for  $n > 60$ , (D) nominal  $M_2$  tide, (E) nuisance core/SV, (F) nuisance crust, (G) nuisance  $M_2$  tide, (H) ionosphere, (I) CHAMP toroidal, (J) Ørsted toroidal, (K) OHM biases, (L) CHAMP Euler angles and (M) Ørsted Euler angles. The horizontal scale has been stretched inside regime (M) in order to better view the three Euler angles. The CHAMP Euler angles have been colour-coded such that the X-, Y- and Z-axis rotations are shown in green, blue and cyan, respectively.

series. The conclusion is that the coestimation of magnetospheric/induced fields is not significantly degrading the estimation of the other parameters. This is in keeping with Thébault *et al.* (2012) who find that, from a theoretical viewpoint, the CI method is in principle superior to any other method at reducing contamination.

Because the CM5 error-covariance matrix  $\mathbf{C}$  is available for the non-magnetospheric/induced parameters, the classic correlation coefficient will be computed for these parameters. As previously explained, only those correlations with  $|\rho| \geq 0.7$  were considered and they will be discussed in terms of dominant groupings. The first group is between spline parameters within the temporal basis of a given SH coefficient in both the nominal and nuisance core SV models. These are always positive and can be extremely high, in some cases  $\rho \approx 1$ . They become more numerous as SH degree  $n$  increases and are due to the influence of the  $\mathcal{P}(|\ddot{B}_r|^2)$  regularizing term, which has a null-space of rank one within each SH expansion. This is exactly the scenario presented in Appendix A where it is shown that correlations induced by rank-deficient regularization norms do not cause instability in the solution. In fact, the eigenvectors of  $\mathbf{C}$  that span this null-space are associated with the smallest eigenvalues and are thus the best determined. As such, an inspection of data fits, particularly those of the observatory time-series, does not show any degradation due to these correlations, as expected.

The second and third groups are actually related through a series of eigenvectors corresponding to the largest eigenvalues of  $\mathbf{C}$  and thus have the potential to degrade the solution. The second group is between nominal crustal coefficient pairs ( $g_n^m, g_{n+2}^m$ ) and ( $h_n^m, h_{n+2}^m$ ) for  $n > 60$  where they are always positive such that  $\rho < 0.910$ . The third group is between the same OHM bias, either X, Y or Z, and either nominal or nuisance, in two different time segments of the same station or in two stations in close proximity. These are also positive and can be extremely high such that  $\rho \approx 1$ , particularly when two time segments are involved. These correlations are not surprising when considering that high-degree lithospheric fields can mimic delta functions at the OHM station positions and so the biases can look like this lithospheric pattern to the data, particularly the CHAMP along-track differences data. These difference data are more insensitive to these patterns and this is why the second group is only seen in the nominal coefficients. It will be shown in Section 4.3 that the use of along-track differences actually eliminates contamination in the lithospheric field, and so, there appears to be a trade-off between decreasing bias, but increasing variance in the model with its use. Despite this link between the second and third groups, inspecting of the lithospheric field does not show any obvious signs of a spurious delta-function pattern. With the level of correlation seen in the third group, one might expect the residual statistics of these OHM components to exhibit large biases. However, this is not the case as most of these biases are on the order of a few nT. This may be due to the elimination of error in the system due to the lithospheric field. The OHM bias correlation mechanism is described in Appendix B.

The final group of significant correlations is between the constant dipole term of the Ørsted toroidal field expansion and the Ørsted Euler angles  $\alpha$  ( $\rho = 0.983$ ) and  $\beta$  ( $\rho = 0.950$ ). These constitute the rotations  $\mathbf{R}_\phi(\alpha)$  and  $\mathbf{R}_\theta(\beta)$  in eq. (9). These two correlations are not surprising since the partial derivatives of toroidal terms have the same form as differential rotations, and in this case both the Euler angles and the dipole term are constant over the time domain of the model. Interestingly, a third correlation is also seen between these same angles with

$\rho = 0.946$  that is apparently induced by this toroidal field parameter. Indeed, when the dipole term is held fixed, the correlation between  $\alpha$  and  $\beta$  disappears, thus corroborating this interpretation. Furthermore, the final  $\alpha$  and  $\beta$  values change by only 2.7 and 1.5 arcsecs, respectively, between the cases where the dipole term is free or fixed. Therefore, this indicates again that these high correlations are likely not degrading the performance of the model.

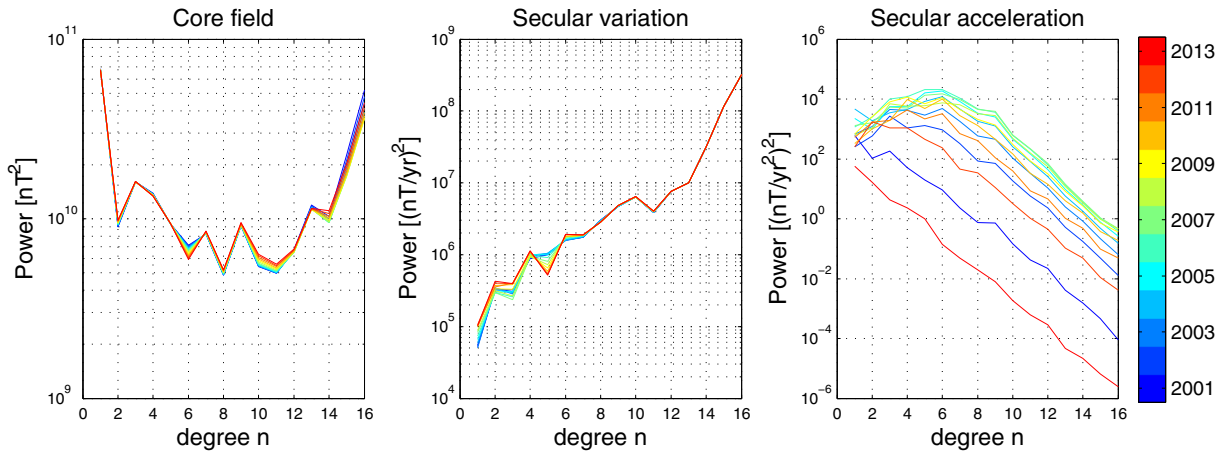
Minor positive correlations of  $\rho < 0.735$  are found between the constant and linear terms of the radial expansion of the zonal CHAMP toroidal coefficients for purely seasonal and diurnally modulated seasonal modes. An additional four minor positive correlations of  $\rho < 0.718$  were found between adjacent 10-d bins for the CHAMP Euler angle  $\gamma$  that constitutes the rotation  $\mathbf{R}_2(\gamma)$  in eq. (7). Specifically, these are between bins 21 and 22, 340 and 341, 341 and 342 and 342 and 343 out of 369 total bins. Once again, none of these appear to be degrading the solution.

### 4.3 Core and lithospheric fields

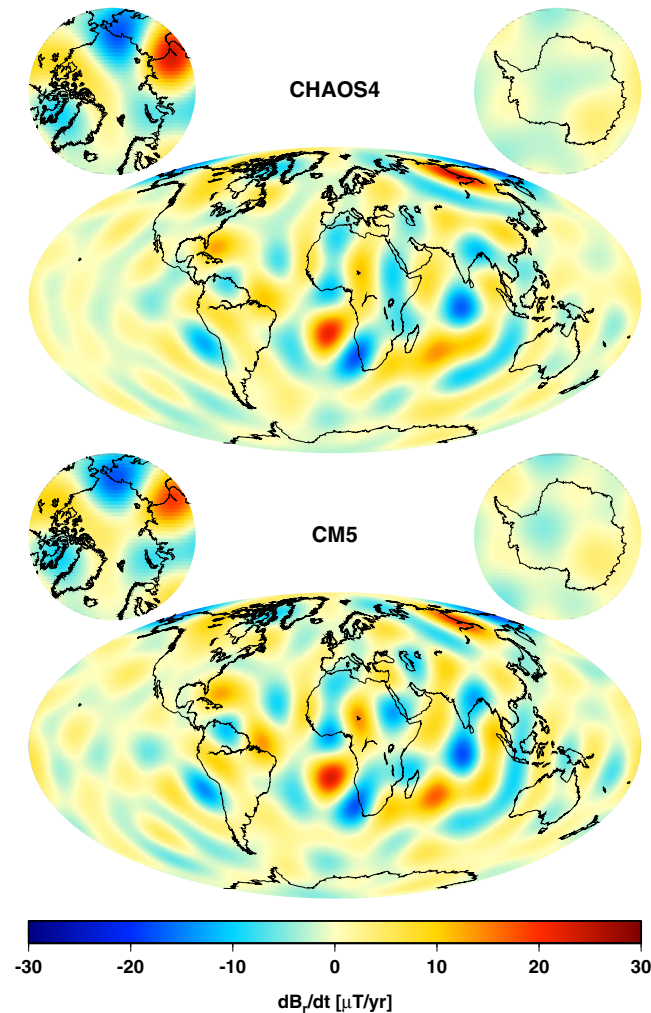
The successful extraction of the core field signal from geomagnetic measurements is of utmost importance as it is one of the few ways of remotely sensing the deep Earth structure. One of the fundamental properties of this field and its time variation is the power per SH degree. This is provided by the Mauersberger–Lowes spectrum (Lowes 1966), which is defined as the mean-square field magnitude per degree over a given sphere. When the spectrum reflects the field itself it is referred to as the  $R_n$  spectrum. Following a similar figure for the CHAOS-4 model in Olsen *et al.* (2014), Fig. 7 shows the spectra for the core field (left-hand panel), its SV (middle panel) and its secular acceleration (SA; right-hand panel) at the CMB for  $n = 1–16$  from CM5 at 1-yr intervals from 2001 to 2013, as indicated by the line colour. The core field does show some variation over this 12-yr span and is generally parabolic, showing less power at the mid-degrees. The SV spectrum shows a generally increasing trend with degree and is in rough agreement in magnitude with CHAOS-4 up to about  $n = 13$ . However, the CM5 spectrum increases faster for  $n > 13$ , being about an order of magnitude larger than CHAOS-4 at  $n = 16$ . There is also much less variation with time for  $n > 8$ . The SA spectrum is generally parabolic for the mid-years, peaking at about  $n = 6$ , and linearly decreasing for the early and late years. The CHAOS-4 spectrum is parabolic for all years, peaking at  $n = 12$ , and is about an order of magnitude larger than CM5 at its peak. This behaviour is likely related to the fact that CM5 minimizes a measure of  $d^2 B_r / dt^2$  at the CMB as opposed to  $d^3 B_r / dt^3$  in CHAOS-4 (although  $d^2 B_r / dt^2$  is minimized at the CMB near the endpoints of the time interval in CHAOS-4), and is thus more stringent.

In order to gain some measure of the spatial quality of the CM5 SV, a comparison is made with the map of the radial component of SV,  $dB_r/dt$ , at the CMB for epoch 2008 from the CHAOS-4 model presented in Olsen *et al.* (2014). The original map reflects an SH expansion of  $n = 1–15$ , but after inspection of the CM5 SV power spectrum (see Fig. 7, middle panel), it was decided that degrees 14 and 15 were beginning to show excessive strength, and so a similar map reflecting an SH expansion of  $n = 1–13$  is shown in Fig. 8 from CHAOS-4 (top panel) and CM5 (bottom panel). It can be seen that most of the major features do indeed agree between the two maps, which yields a 0.936 global correlation coefficient. This includes such features as a generally weak SV under the Pacific and strong SV under the Atlantic hemispheres, and extremely weak SV under the southern polar region (Holme *et al.* 2011). The maps also show the same strong pair of SV patches in the northern polar region, which have been attributed to westward motion of a large flux patch under Siberia (Blokhin & Gubbins 1985). As far as can be inferred from this comparison, the CM5 SV appears to be in reasonable agreement with that of CHAOS-4, if not quite at the same level of resolution.

Moving now to the static part of the internal field, the top panel of Fig. 9 shows the  $R_n$  spectra for the MF7, CHAOS4, CM5ng and CM5 models, along with the spectra of the differences of CM5 with the other models, for  $n > 20$ . All models show approximately the same power in their overlap regions, but with CM5ng decaying after about  $n = 110$ . The differences between CM5 and CHAOS4 (black) are of higher



**Figure 7.** Mauersberger–Lowes spectra (Lowes 1966) of the core field (left-hand panel), secular variation (middle panel) and secular acceleration (right panel) at the CMB for  $n = 1–16$  from the CM5 model at 1-yr intervals from 2001 to 2013. These years are indicated by the colour of the lines.



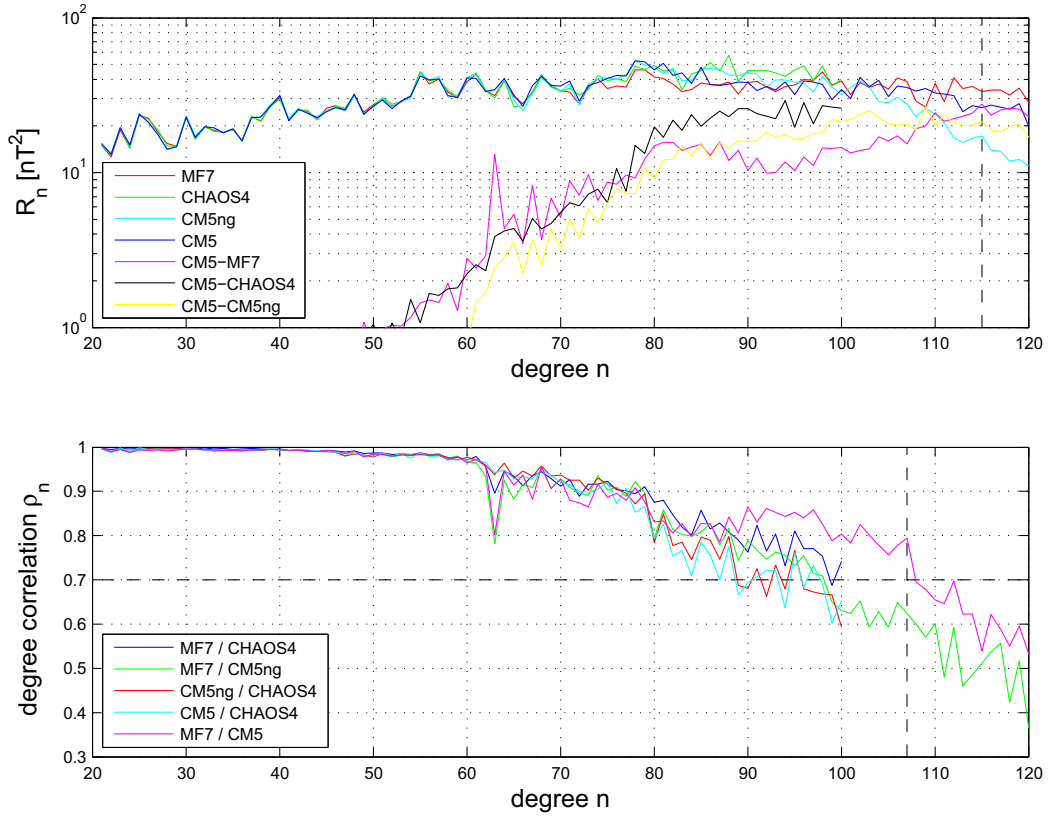
**Figure 8.** Secular variation of the radial component of the main magnetic field,  $dB_r/dt$ , at epoch 2008 on the CMB for  $n = 1–13$  from the CHAOS-4 model of Olsen *et al.* (2014) (top) and the CM5 model (bottom).

power than those between CM5 and MF7 (magenta) for  $n > 77$ . The power in the differences between CM5 and MF7 surpasses the power in CM5 itself (blue) at  $n = 115$ , as shown by the dashed vertical line, and suggests a possible trust region for the CM5 lithospheric field.

While the  $R_n$  spectra give an indication of the magnitude of the vector of field coefficients at a particular degree, the degree correlation  $\rho_n$  gives information about the angle between the coefficient vectors of two different models at a given degree. The bottom panel of Fig. 9 shows  $\rho_n$  for all pairs of models, except for CM5 and CM5ng. For  $n = 21–62$  all degree correlations appear to be in very good agreement. For  $n = 63–79$  correlations with MF7 are slightly better for CHAOS4, followed by CM5ng, and then CM5. There is an anomalous dip in correlations at  $n = 63$  between all models and MF7. The reason for this is unknown, but may be linked to behaviour in the transition zones between polar and non-polar areas. For  $n = 80–100$ , the correlations between MF7 and CM5 begin to separate from the others as they decrease. For  $n > 100$  the MF7 and CM5 correlations are significantly higher than between MF7 and CM5ng. In fact, the correlations between MF7 and CM5 (magenta) remain above those of MF7 and CM5ng (green) for  $n > 86$ . This is taken as an indication of the usefulness of the CHAMP along-track differences in eliminating noise and increasing the SNR of the high-degree lithospheric field. Since  $\rho_n = \cos \alpha$ , where  $\alpha$  is the angle between the vectors of coefficients at degree  $n$  between two models, a value of  $\rho_n = 0.7$  (indicated by the dashed horizontal line) then corresponds approximately to  $\alpha = 45^\circ$ , which is the geometric ‘half-way point’ between parallel and orthogonal. It is used here as a heuristic to indicate whether the vectors are more or less ‘similar’. The MF7 and CM5 correlations remain above  $\rho_n = 0.7$  up to  $n = 107$  (indicated by the vertical dashed line), and so this is taken as the nominal resolution level of the CM5 lithospheric field. It should be emphasized, however, that the MF7 model is only a surrogate for the true lithospheric field and commonality of parametrization, regularization, etc., between CM5 and MF7 could be responsible for part of the agreement seen, which could be different from the true field.

The CM5 lithospheric field is compared with those of MF7 and CHAOS4 in physical space in Fig. 10 using the  $B_r$  component at the Earth’s surface at epoch 2010. The top four panels show CM5 and the differences among the three models for  $n = 16–90$ . The global correlation coefficients are 0.909 between CM5 and CHAOS4, 0.926 between CM5 and MF7 and 0.931 between CHAOS4 and MF7, thus, both CM5 and CHAOS4 are closer to MF7 than to each other. The most striking features of the difference maps are the oscillating north–south striping, which follow CHAMP tracks and are due to varying magnetic activity not accounted for in the models (see Thébault *et al.* 2012;





**Figure 9.** Top panel shows the  $R_n$  spectra (Lowes 1966) of the MF7, CHAOS4, CM5ng and CM5 models for  $n > 20$  at the Earth’s surface, and also for the model differences with CM5. The vertical dashed line at  $n = 115$  indicates where the power in the differences between CM5 and MF7 surpasses that of CM5 itself. The bottom panel shows the degree correlation  $\rho_n$  between various combinations of the MF7, CHAOS4, CM5ng and CM5 models. The horizontal dashed line at  $\rho_n = 0.7$  indicates the geometric ‘half-way point’ where two vectors intersect at approximately a  $45^\circ$  angle and which is used as a heuristic for deciding that two vectors are more ‘similar’ than not. The vertical dashed line at  $n = 107$  indicates the cross-over point above which  $\rho_n < 0.7$  between the MF7 and CM5 models.

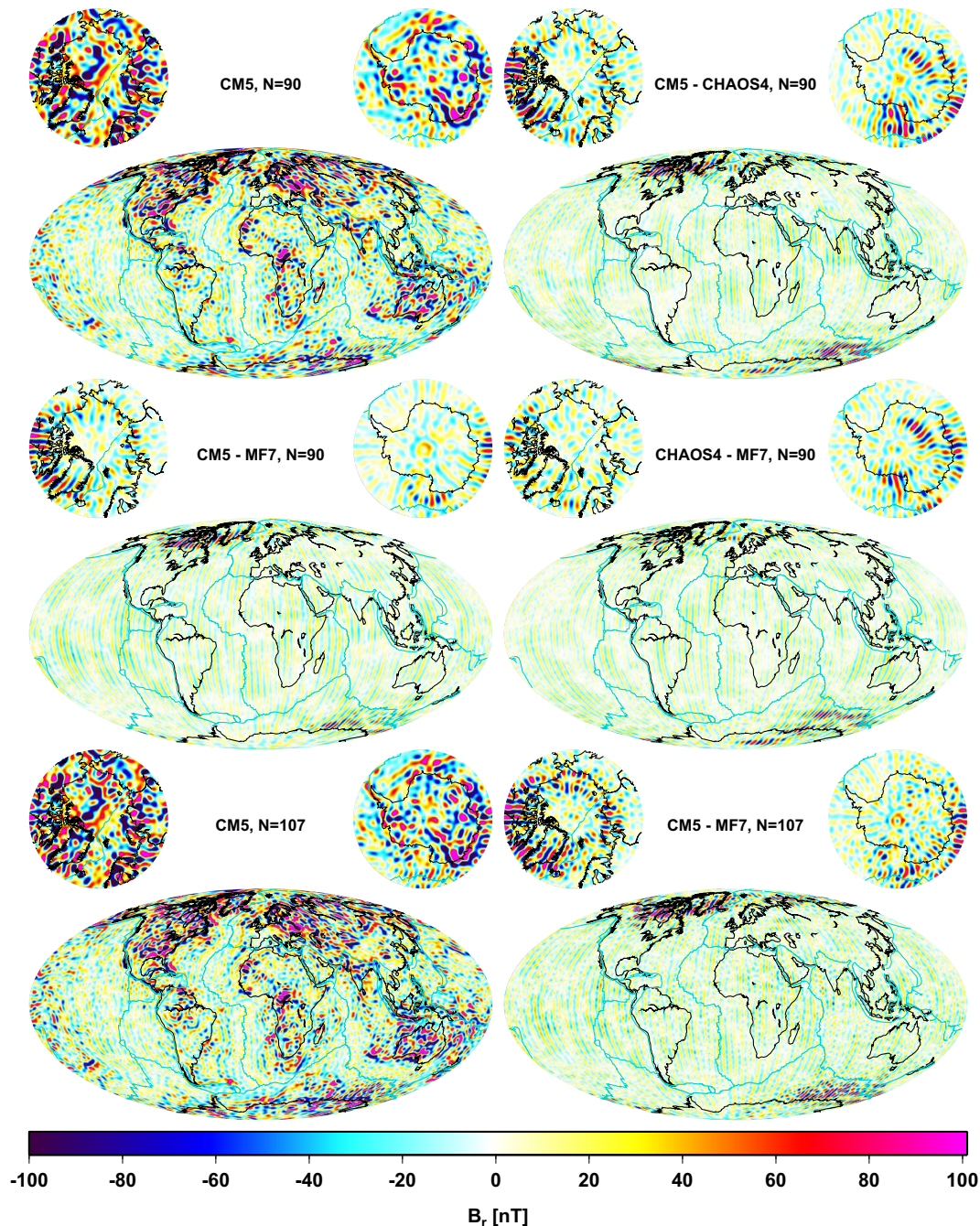
Lesur *et al.* 2013), and the large deviations in the auroral zones around the magnetic poles, which are due largely to the presence of PEJs. The auroral zone differences are less/more in the south/north for CM5 versus CHAOS4 with respect to MF7. To a great extent the PEJs were selected out of the data from which the MF7 model, and its predecessor MF6 (Maus *et al.* 2008), were derived, but not so for CM5 and CHAOS4. In addition, the striping in the CHAOS4 and MF7 differences is somewhat subdued compared to those between CM5 and MF7, which is likely due to the more stringent quality criteria used in the selection of the CHAOS4 data. The bottom two panels show CM5 and its difference with MF7 for  $n = 16$ – $107$ , the nominal resolution level of CM5. The global correlation coefficient between them is 0.902. The difference map shows the same basic patterns as before except now the auroral zones are much better defined, as expected, since the smaller-scale effects of the PEJs are now mapped into the lithospheric field. Although the nominal  $n = 107$  resolution level of the CM5 lithospheric field can really only be trusted outside of the auroral zones, this marks a milestone in that CA methods are becoming increasingly more competitive with targeted sequential methods in the realm of lithospheric field modelling.

#### 4.4 Oceanic $M_2$ tidal field

In this section, the  $M_2$  tidal magnetic fields as predicted by two numerical models, and as extracted from CM5, shall be presented and discussed. A preliminary matter is to describe and compare the models used to create the predictions, and this is provided in Appendix C. As these two models have been previously introduced in Tyler *et al.* (2003) and Kuvshinov (2008), they will herein be referred to as the ‘Tyler’ and ‘Kuvshinov’ models, respectively.

A first, fundamental comparison involves the power spectra of the  $M_2$  magnetic fields from each model. While the CM5  $M_2$  tidal field is modelled with an SH expansion up to degree and order 36, it may not be stable up to this degree. In order to determine an appropriate truncation level up to which the model can be physically interpreted, an  $R_n$  spectrum is again employed as with the lithospheric field. However, since this is a time varying field, the  $R_n$  spectrum will now be the mean-square field magnitude per degree over a sphere of radius  $r$  and over the  $M_2$  period  $T_{M_2} = 12.42060122$  hr. This follows from eq. (4) and is found to be

$$\frac{1}{4\pi r^2 T_{M_2}} \int_0^{T_{M_2}} \int_{\Omega} |\mathbf{B}_{M_2}(t, \mathbf{r})|_2^2 d\Omega dt = \sum_{n=1}^{36} R_n(r), \quad (29)$$

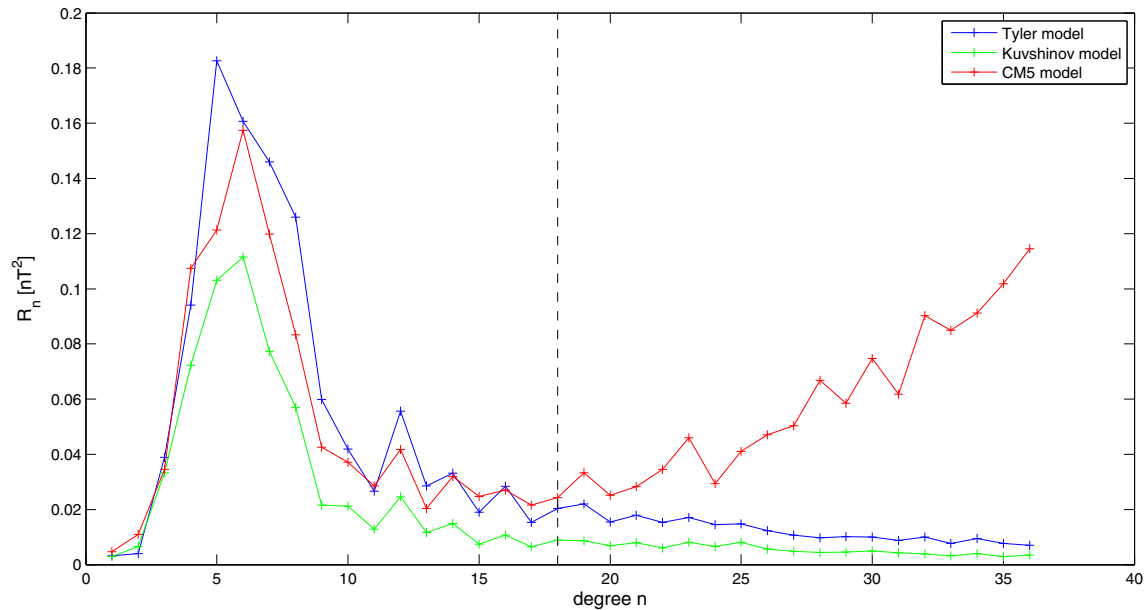


**Figure 10.** The top four panels show maps of the  $B_r$  component of the lithospheric field at the Earth's surface at epoch 2010 from  $n = 16$ – $90$  for the CM5 model and its differences with the MF7 and CHAOS4 models, and the difference between the MF7 and CHAOS4 models in nT. The bottom pair of panels show maps of  $B_r$  at epoch 2010 from  $n = 16$ – $107$  for the CM5 model and its difference with the MF7 model. Note that  $n = 107$  is considered the maximum degree of resolution for the CM5 model. The polar projections show  $30^\circ$  caps. The colour scale saturates at  $\pm 100$  nT.

$$R_n(r) = (n+1) \left( \frac{a}{r} \right)^{2n+4} \left\{ \frac{1}{2} |\tau_n^0|^2 + \sum_{m=1}^n \left[ |\tau_n^m|^2 + |\tau_n^{-m}|^2 \right] \right\}. \quad (30)$$

The  $R_n$  spectra have been calculated at the Earth's surface ( $r = 6371.2$  km) for the theoretical forward models of Tyler (blue), and Kuvshinov (green) and the estimate from CM5 (red) and are shown in Fig. 11. The Kuvshinov and CM5 spectra both have a distinct peak at  $n = 6$ , while the Tyler model peaks at  $n = 5$ , after which the behaviour of CM5 is approximately parabolic. The Kuvshinov theoretical model shows consistently less power than CM5 at all degrees while the Tyler model shows less power for  $n > 16$ , but generally more power for lower degrees. However, the profiles of the three spectra agree quite well up to about  $n = 18$ , and so this is taken as the truncation level of the interpretable part of the CM5 model. The fact that the CM5 power lies between the two theoretical models at the lower degrees makes sense in light of the assumptions made about upper-mantle conductivity in these models and is discussed in more detail in Appendix C.





**Figure 11.** The  $R_n$  spectra (Lowes 1966) of the time-averaged oceanic  $M_2$  tidal magnetic field at the Earth's surface from  $n = 1$ –36 for the theoretical forward models of Tyler (blue), Kuvshinov (green) and the estimate from CM5 (red). The vertical dashed line at  $n = 18$  shows the approximate point beyond which the spectral power begins to increase in the CM5 estimate and the profile begins to diverge from the theoretical predictions.

The spatial features of the predicted and extracted  $M_2$  fields shall now be discussed. Because the tidal signal is temporally periodic and the wave shape (sinusoid) known, there are only two numbers (amplitude/phase or, alternatively, the real/imaginary components) required to completely specify the signal at a given location. Here the convention in ocean tidal literature is followed where amplitude/phase are used and referenced to zero (for the amplitude) and Greenwich phase (for the phase). The sense of propagation of the features is towards increasing values in the phase contours. The propagation circles around the so-called ‘amphidromes’, where the phase values become undefined.

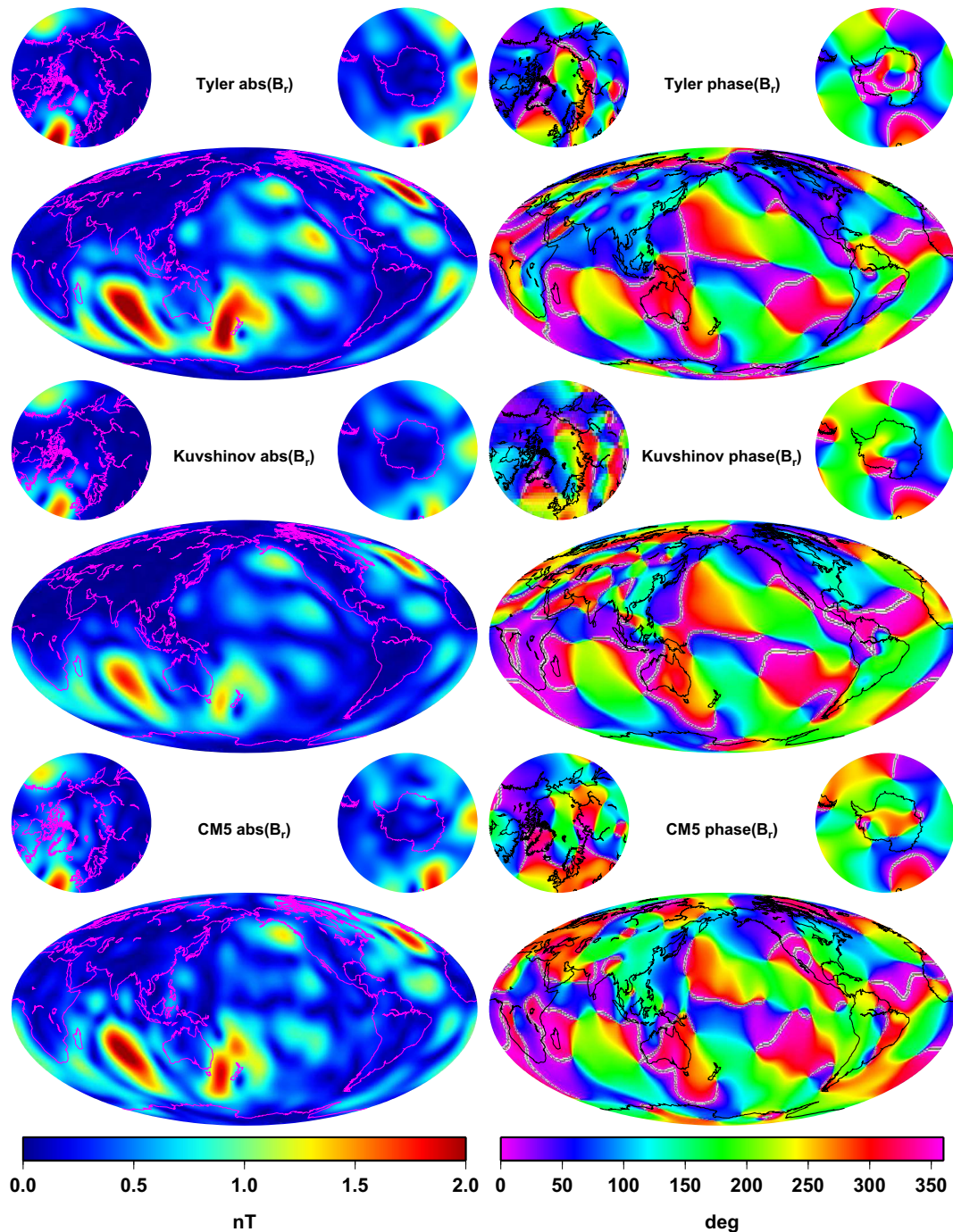
In Fig. 12 are shown the  $M_2$  radial magnetic fields from each of the two prediction models, and the CM5 extraction. In keeping with the spectral results, all fields shown have been truncated at degree  $n = 18$ . Furthermore, the fields are presented for a spherical surface at altitude 430 km. This represents the approximate mean-altitude of CHAMP over the first 4 yr of the mission. The satellite data, because they cover the oceans, are expected to be primarily responsible for the  $M_2$  resolving power in CM5 and so the choice of presenting the comparison at 430 km is aimed at collocating predicted and observed fields. Note, however, that at this altitude the smaller-scale features are further reduced due to geometric attenuation.

One immediately sees that the amplitude distributions (Fig. 12, left-hand panels) agree very well to within the coarse comparison attempted here. The global correlation coefficients between the maps of Tyler and CM5, Kuvshinov and CM5 and Tyler and Kuvshinov are 0.879, 0.921 and 0.929, respectively. That is, the two forward models, despite their different formulations and representations of the conductive mantle, produce similar  $M_2$  predictions; these predictions also independently agree with observations as extracted through CM5. This provides validation for the CM5 recovery as well as mutual validation between the forward models. Recall that the  $M_2$  tidal basis functions in CM5 have no information about continental versus oceanic regions, and the weak strength of the signal over the continents is a strong indication that the extracted signal is indeed of oceanic origin. More specifically, these basis functions presently include only temporal information and are given no *a priori* spatial information. The correlation in the spatial distributions then surely confirms the ocean tides as the source. The maximum amplitude of the radial component of the CM5  $M_2$  at 430 km is about 2.11 nT, but near the end of the mission CHAMP had descended to an altitude of about 250 km at which point the maximum value is about 2.58 nT. This is a very promising result given that much of the recovered signal is on the order of 1–2 nT in strength over this satellite altitude range and represents the next largest magnetic field source after the lithosphere.

The Greenwich phase distributions (Fig. 12, right-hand panels) also show coarse agreement once some caveats are observed: note that the concept of phase becomes undefined as the associated amplitude vanishes. Hence, it is not surprising that reasonable agreement between predicted and CM5 fields is found only where the amplitudes are large. Also note that in the continental regions, agreement in phase is not expected between the two forward models as the Tyler model employs a thin-shell induction approximation, which may become inaccurate as the assumed thin conductor vanishes.

#### 4.5 Ionospheric field

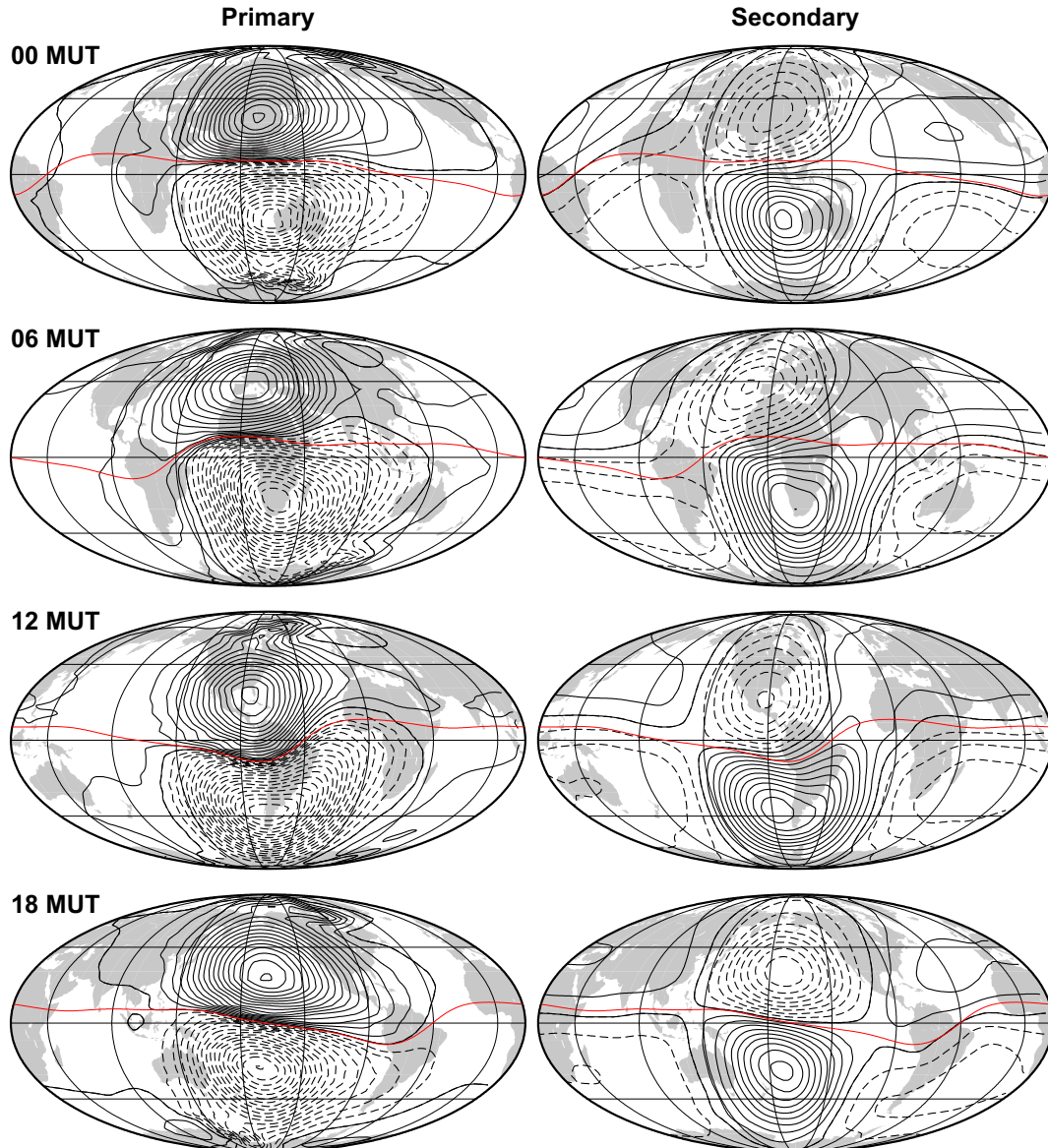
The ionospheric  $E$ -region current system is modelled as a sheet current at an altitude of 110 km (Sabaka *et al.* 2002) and the corresponding equivalent current (stream) function for the primary (left) and secondary (right) systems are shown in Fig. 13 during vernal equinox centred on noon MLT, but at different magnetic universal times. The primary function shows the distinctive counter-clockwise/clockwise flow in the



**Figure 12.** The amplitude (left-hand side) and phase (right-hand side) of the  $B_r$  component of the oceanic  $M_2$  tidal magnetic field at 430 km altitude computed from the theoretical forward models of Tyler *et al.* (2003; top) and Kuvshinov (2008; middle) and the estimate from CM5 (bottom) for  $n = 1-18$ .

northern/southern Solar Quiet ( $Sq$ ) focus and the opposite for the secondary function. There is a notable lack of current flow on the limbs of the primary maps, which is due to the influence of the  $\mathcal{P}(|\mathbf{J}_{\text{eq,MLT:21-05}}|^2)$  smoothing norm. However, current does flow on the limbs of the secondary maps, as it should since the conductive medium for the induced field exists on the nightside. There is approximately 1/2 to 1/3 of the current flowing in the secondary system compared to the primary. The QD-symmetric basis functions conform to the main field contour lines as can be seen by the progression through time with respect to the dip-equator, shown as a red contour line.

Close inspection of the primary and secondary current functions reveals the possibly counter-intuitive finding that the secondary induced system actually leads the primary system by approximately 20 min, as opposed to lagging. This, however, is expected on the physical grounds that the inducing field is not proportional to the inducing field  $\mathbf{B}$ , but rather its first time derivative  $\dot{\mathbf{B}}$  (Schmucker 1985). Numerically this is evidenced by the fact that for realistic 1-D conductivity profiles the Q-response, the complex ratio between internal (induced) and external (inducing) SH expansion coefficients (assuming a time harmonic dependence  $\exp i\omega t$ ) has a positive imaginary part (Olsen *et al.* 2005).

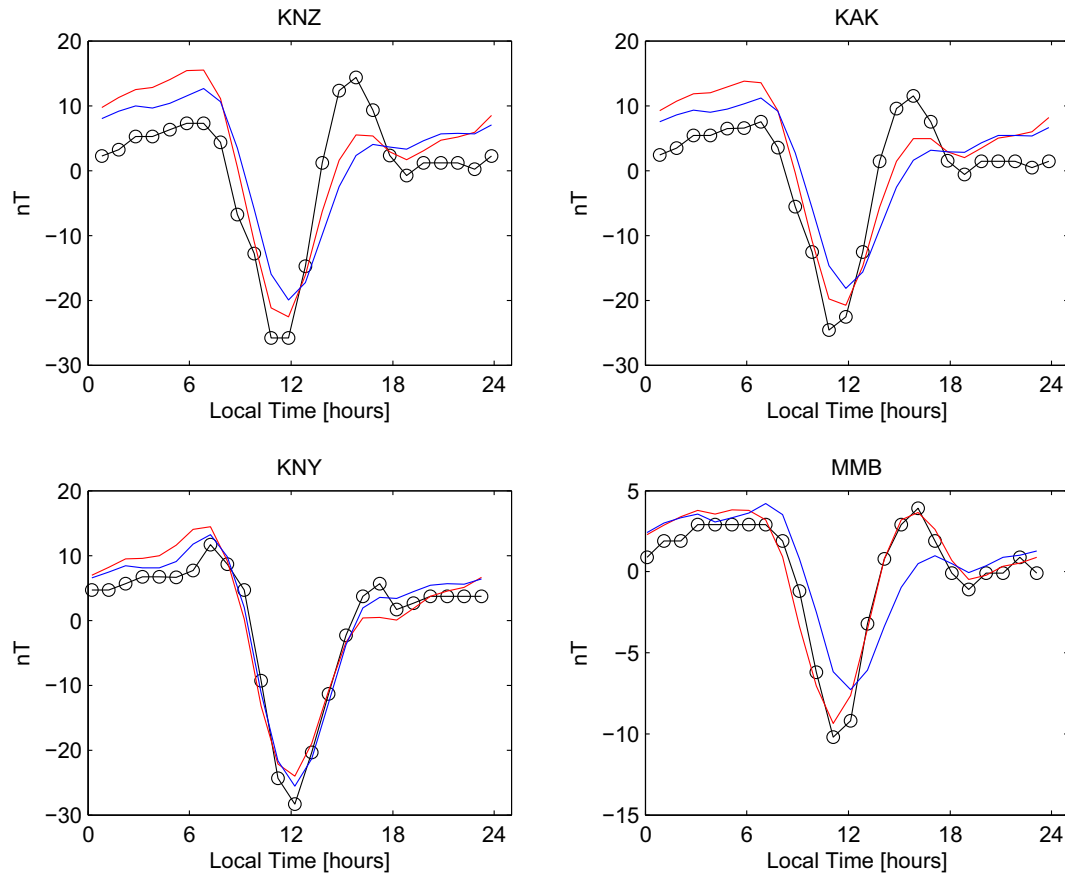


**Figure 13.** Ionospheric *E*-region primary (left-hand side) and secondary (right-hand side) equivalent current functions during vernal equinox centred on noon magnetic local time, but at various magnetic universal times (MUT). Solid/dashed contour lines indicate counter-clockwise/clockwise current flow. A 10-kA current flows between contours. The dip-equator is shown as a red contour. A value of  $F_{10,7} = 126.0 \times 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$  was used for evaluation.

The largest differences between the ionospheric 1-D induction used in CM4 and the ‘1-D+oceans’ used here are manifested along the coasts and continental margins. These coastal effects have been seen as anomalous daily magnetic variations in the vertical *Z* components at observatory stations, for instance, in Japan by Rikitake *et al.* (1956) and were explained as such by Kuvshinov & Utada (2010). In particular, these were seen as a roughly 2-hr shift of the local noontime peak towards morning hours. After scaling the local *Sq* signal for that day from CM4, Kuvshinov & Utada (2010) showed plots of the predictions when considering 1-D versus 3-D conductivity at several observatories. This scaling was necessary since CM4 does not track, nor was it meant to track, day-to-day ionospheric variations.

This exercise is repeated here for 2011 March 18 at the four Japanese observatories of Kanozan (KNZ), Kakioka (KAK), Kanoya (KNY) and Memambetsu (MMB). The predictions in *Z* from CM5 and CM5i1d were scaled and shifted to fit the observed data, whose daily mean had been removed, between 06 LT (local time) and 18 LT. The results are shown in Fig. 14 where the observed measurements are shown with black symbols and the CM5 and CM5i1d predictions are shown in red and blue, respectively. There are appreciable differences between the 1-D and 3-D predictions at KNZ, KAK and MMB, but not so for KNY, which was also seen by Kuvshinov & Utada (2010). While the morning-side peak shifts are predicted at KNZ and KAK from CM5, the side-lobe signature is not. However, MMB shows not only good agreement in peak shift between observed and CM5, but also at the side-lobes. This appears to be opposite of the findings by Kuvshinov & Utada (2010) who do not fit MMB well. It should be noted, however, that there exists much day-to-day variation at each station and this represents only a sample of this variation. Overall, the 3-D treatment of ionospheric induction is benefitting the performance of CM5.





**Figure 14.** Observed (black symbols) versus predicted vertical Z component on 2011 March 18 at the four Japanese observatories of Kanozan (KNZ), Kakioka (KAK), Kanoya (KNY) and Memambetsu (MMB). Scaled and shifted predictions for this day from CM5 are in red and from CM5i1d are in blue. The daily mean has been removed from the observed data.

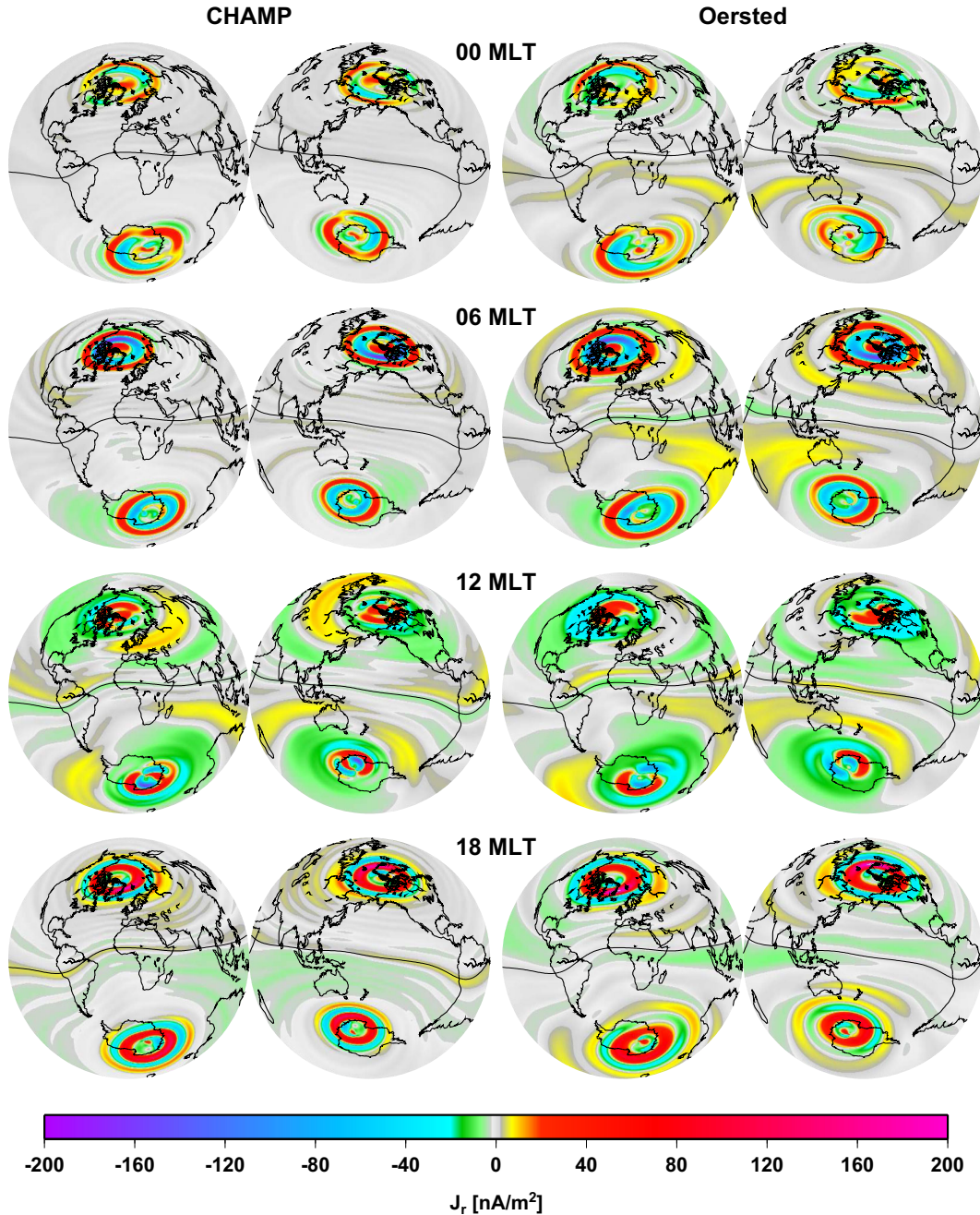
#### 4.6 Toroidal field

Although the toroidal magnetic field is not an official *Swarm* Level-2 product, it is still of interest since it indeed has a contribution to what is measured by the satellites. Estimates of the radial component  $J_r$  of the ionospheric  $F$ -region current system through which CHAMP and Ørsted fly are shown in Fig. 15 at different MLTs. Note that each map shows a constant MLT at each spatial location. One can see that within the CHAMP sampling shell,  $J_r$  is mostly absent at low and mid-latitudes at local midnight (00 MLT) and dawn (06 MLT), but then builds through noon (12 MLT) and into dusk (18 MLT) along the dip-equator (black contour). This upwelling of current along the dip-equator with downwelling along the side-lobes is the meridional current system associated with the EEJ (Maeda *et al.* 1982). In fact,  $J_r$  changes directions from negative to positive along the dip-equator from noon to dusk. Comparison with estimates from CM4 shows a slight decrease in peak strength of the meridional current at dusk in CM5. This may be due to the fact that vector data over the entire CHAMP mission are used in CM5, while CM4 used only scalar data from the first 2 yr of the mission. The polar latitudes of the maps are dominated by the region-1 and 2 FACs (see Burrows *et al.* 1984; Friis-Christensen 1984; Iijima & Potemra 1976; Zanetti *et al.* 1983, 1984).

For the Ørsted sampling shell, the current pattern is much more complicated, but shows the general build-up of currents from midnight to dawn to noon at low and mid-latitudes. However, the strength of the meridional currents appears to be slightly stronger at noon than at dusk MLT and may be due to the fact that the Ørsted shell is sampling the currents at higher altitudes.

#### 4.7 Magnetospheric and induced fields

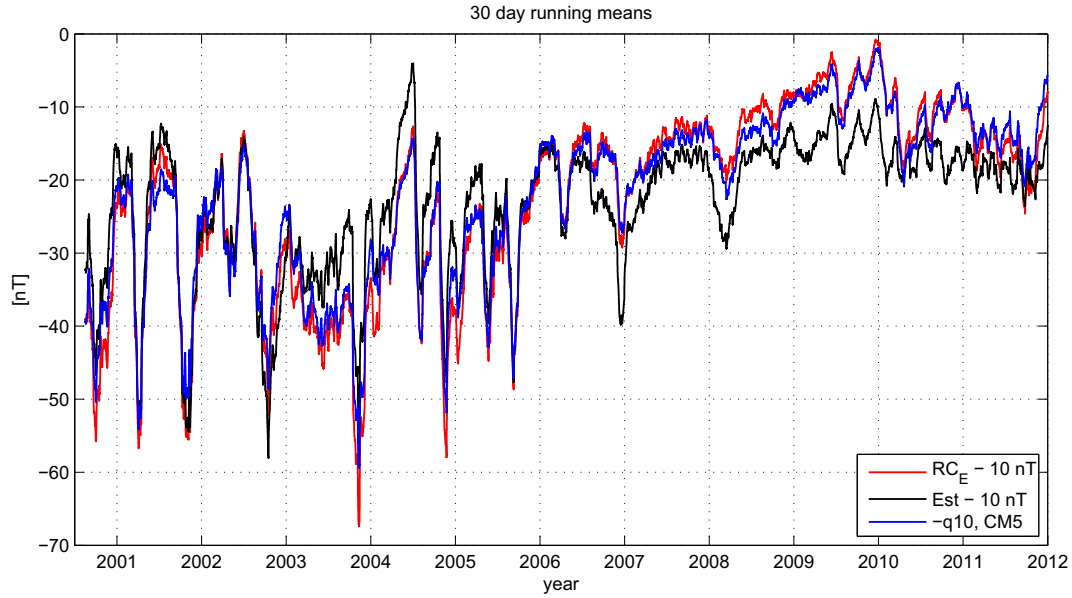
Recall that independent magnetospheric and associated induced fields are estimated in CM5 over 1-hr bins covering the time domain of the model. Focusing first on the magnetospheric field recovery, and in particular the axial dipole term  $q_1^0$ , a comparison is made with two other indices in Fig. 16, which shows the 30-d running means of  $E_{st}$  (the external, magnetospheric part of  $D_{st}$ , shown in black),  $RC_E$  (the external part of  $RC$ , shown in red) and of  $-q_1^0$  (blue) as determined by CM5 (the minus sign is introduced in order to make this measure of the magnetospheric ring-current strength compatible to the definition of  $E_{st}$  and  $RC_E$  as the southward component of the magnetic field at the dipole equator). All three quantities measure the strength of the magnetospheric ring-current and should therefore be the same, apart from an offset for  $E_{st}$  and  $RC_E$ . The reason for this offset is the fact that the absolute value of the ring-current strength can only be determined from satellite data and thus is undetermined for  $E_{st}$  and  $RC_E$ , which are solely derived from ground observatory data.



**Figure 15.** Ionospheric  $F$ -region radial current component  $J_r$  during vernal equinox for various magnetic local times (MLT) within the CHAMP (left-hand pairs) and Ørsted (right-hand pairs) sampling shells. Each pair consists of two azimuthal equidistant projections centred at  $0^\circ$  (left member) and  $180^\circ$  (right member) longitude. Note that each map shows  $J_r$  for the same MLT at all spatial locations. The dip-equator is shown as a black contour.

Comparison with  $-q_1^0$  allows for the determination of this offset, and the excellent agreement between  $-q_1^0$  and  $RC_E - 10$  nT indicates that the strength of the magnetospheric ring-current as determined from observatory data alone (i.e.  $RC_E$ ) and by combining observatory and satellite data (i.e.  $-q_1^0$ ) yields very similar results. In contrast, although the variations of  $E_{st}$  follow roughly those of  $-q_1^0$  and  $RC_E$ , the fact that the black curve is above the red and blue curves before 2006, but below after 2006 reveals the baseline instabilities of the  $D_{st}$  index.

Linear regression between  $RC_E$  and  $-q_1^0$  yields a regression coefficient of 0.88 which means that the time changes of the magnetospheric field as seen by the satellites combined with ground data are only 88 percent of that predicted by  $RC$  as derived from ground data only. Interestingly, the corresponding value, 0.91 as determined by Olsen *et al.* (2014), is also lower than one, which may indicate that some of the signal described by  $D_{st}$  and  $RC$  may be due to a source between surface and satellite altitude.



**Figure 16.** The 30-d running means of  $E_{st}$  (the external, magnetospheric part of  $D_{st}$ , shown in black),  $RC_E$  (the external part of  $RC$ , shown in red) and of  $-q_1^0$  (blue) as determined by CM5 (the minus sign is introduced in order to make this measure of the magnetospheric ring-current strength compatible to the definition of  $E_{st}$  and  $RC_E$  as the southward component of the magnetic field at the dipole equator). As the absolute levels of  $E_{st}$  and  $RC_E$  cannot be determined, the  $-10$  nT offset brings them to the level of  $-q_1^0$ .

Now, in order to demonstrate the quality and consistency of the combined internal/external field recoveries, the 1-D electromagnetic (EM) response functions are next estimated as the frequency-dependent transfer function  $Q_1^0(\omega)$  between the Fourier transformed time-series of the external and internal axial dipole coefficients  $q_1^0$  and  $\iota_1^0$  of the magnetospheric and induced fields of CM5, respectively, such that

$$\iota_1^0(\omega) = Q_1^0(\omega)q_1^0(\omega), \quad (31)$$

where  $\omega = \frac{2\pi}{T}$  is angular frequency, and  $T$  is a period. The responses are estimated in a period range between 1.3 d and 200 d. The section averaging approach (Olsen 1998) was used to estimate  $Q_1^0(\omega)$  while the errors  $\delta Q_1^0(\omega)$  were estimated using a standard statistical approach (see Jenkins & Watts 1968) with a chosen error probability of  $\beta = 0.05$ . The  $Q_1^0(\omega)$  is finally transformed to the  $C$ -response through the relationship (Weidelt 1972)

$$C(\omega) = \frac{a}{2} \frac{1 - 2Q_1^0(\omega)}{1 + Q_1^0(\omega)}, \quad (32)$$

with corresponding errors (Schmucker 1999)

$$\delta C(\omega) = \frac{3a}{2} \frac{1}{|1 + Q_1^0(\omega)|^2} \delta Q_1^0(\omega), \quad (33)$$

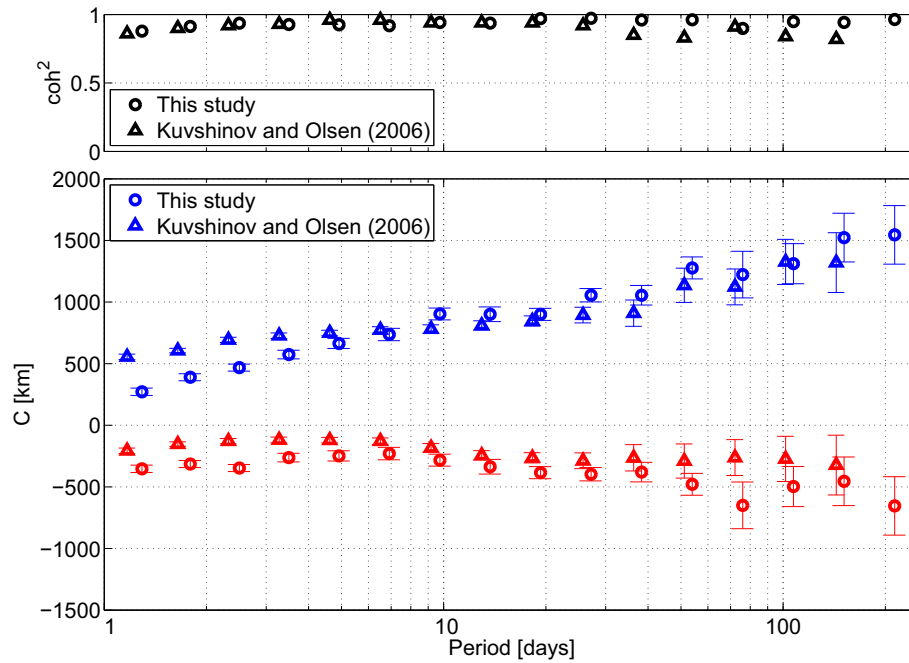
where again  $a$  is the Earth's mean-radius (6371.2 km).

Fig. 17 shows the estimated  $C$ -responses in the bottom plot as the real,  $\Re\{C\}$  (blue), and imaginary,  $\Im\{C\}$  (red), parts of  $C$  from CM5 (circles) and the uncorrected responses obtained by Kuvshinov & Olsen (2006; triangles) with error bars of one standard deviation,  $\delta C$ . Note that results of Kuvshinov & Olsen (2006) are based on 5 yr of CHAMP, Ørsted and SAC-C magnetic data. The top plot shows the squared-coherence,  $\text{coh}^2$ , between input and output time-series from CM5 (circles) and from Kuvshinov & Olsen (2006; triangles). From this figure it may be concluded that the CM5 results are respectable in that the coherencies are very high, even at the longest periods, and the responses themselves are physically very reasonable and smooth having small-to-moderate uncertainties. Interpretation of these responses is out of scope of this paper but will be a subject of a separate publication.

## 5 CONCLUSIONS

The *Swarm* Level-2 CI algorithm has been applied to over 12 yr of CHAMP, Ørsted, SAC-C and observatory data to produce CM5, a pre-*Swarm* comprehensive magnetic field model. This model represents a marked departure from the previous CM4 model in several ways including the use of the bias mitigation technique known as SIVW, which allows for the exploitation of high SNRs for certain parameter and data subset combinations. This has allowed for a much cleaner recovery of the lithospheric field, free from contamination by, for instance, the EEJ. It has increased the lithospheric resolution level to arguably  $n = 107$  by allowing gradient information from CHAMP along-track differences to be exploited. It has also made possible a very impressive extraction of the oceanic  $M_2$  tidal field. Inclusion of a 3-D conductivity model for ionospheric induction has greatly improved the prediction of the  $Sq$  signal at coastal observatories. In addition, the independent





**Figure 17.**  $C$ -responses,  $C$ , derived from the time-series of external and internal axial dipole coefficients of the magnetospheric and associated induced fields of CM5, respectively. The bottom plot shows the real,  $\Re\{C\}$  (blue), and imaginary,  $\Im\{C\}$  (red), parts of  $C$  with error bars of one standard deviation with respect to period from CM5 (circles) and from the uncorrected responses obtained by Kuvshinov & Olsen (2006; triangles). The top plot shows the squared-coherence,  $\text{coh}^2$ , between input and output time-series with respect to period from CM5 (circles) and Kuvshinov & Olsen (2006; triangles). Both plots are computed following Kuvshinov & Olsen (2006).

estimation of magnetospheric and associated high-frequency–induced fields continually through time has allowed for a very respectable determination of  $C$ -responses, which yield information about the 1-D conductivity structure in the mantle.

The results presented here indicate that the CI is performing satisfactorily and that the various CM5 source descriptions may be considered as validation reference fields for the *Swarm* Level-2 products. In particular, the lithospheric, ionospheric and oceanic  $M_2$  tidal fields are all expected to remain consistent during the *Swarm* mission duration such that the first two may be used to validate the CI as well as dedicated inversion techniques. While the  $M_2$  tidal field is not an official Level-2 product, it should be considered a target field for assessing model sensitivity in the CI approach by virtue of its relatively weak, but predictable signal.

It is hoped that CM5 will serve the community by providing reference fields of its various constituents. To this end, a forward code will be made available soon, as well as sets of SH coefficients, etc., describing various magnetic and electric current fields. Those interested should contact the authors directly.

## ACKNOWLEDGEMENTS

We thank Vincent Lesur and an anonymous reviewer for valuable comments that improved the quality of the paper. The NASA Center for Climate Simulation at Goddard Space Flight Center provided computational resources. TJS and RHT are supported by the NASA Earth Surface and Interior program.

## REFERENCES

- Anton, H., 1984. *Elementary Linear Algebra*, John Wiley and Sons.
- Atkinson, K.E., 1989. *An Introduction to Numerical Analysis*, John Wiley and Sons.
- Beggan, C.D., Macmillan, S., Hamilton, B. & Thomson, A.W.P., 2013. Independent validation of Swarm Level2 magnetic field products and ‘Quick Look’ for Level1b data, *Earth Planets Space*, **65**, 1345–1353.
- Bertsekas, D., 1999. *Nonlinear Programming*, Athena Scientific.
- Bloxham, J. & Gubbins, D., 1985. The secular variation of Earth’s magnetic field, *Nature*, **317**, 777–781.
- Burrows, J.R., Hughes, T.J. & Wilson, M.D., 1984. A study of high latitude current systems during quiet geomagnetic conditions using MAGSAT data, in *Magnetospheric Currents*, Vol. 28, pp. 104–114, ed. Potemra, T.R., American Geophysical Union.
- Constable, C.G., 1988. Parameter estimation in non-Gaussian noise, *Geophys. J.*, **94**, 131–142.
- Demmel, J., 1997. *Applied Numerical Linear Algebra*, Society for Industrial and Applied Mathematics.
- Emmert, J.T., Richmond, A.D. & Drob, D.P., 2010. A computationally compact representation of magnetic-apex and quasi-dipole coordinates with smooth base vectors, *J. geophys. Res.*, **115**, A08322, doi:10.1029/2010JA015326.
- Everett, M., Constable, S. & Constable, C., 2003. Effects of near-surface conductance on global satellite induction responses, *Geophys. J. Int.*, **153**, 277–286.
- Finlay, C.C., Maus, S., Beggan, C.D., Hamoudi, M., Lowes, F.J., Olsen, N. & Thébaud, E., 2010. Evaluation of candidate geomagnetic field models for IGRF-11, *Earth Planets Space*, **62**, 719–727.

- Friis-Christensen, E., 1984. Polar cap current systems, in *Magnetospheric Currents*, Vol. 28, pp. 86–95, ed. Potemra, T.R., American Geophysical Union.
- Friis-Christensen, E., Lühr, H. & Hulot, G., 2006. *Swarm*: a constellation to study the Earth's magnetic field, *Earth Planets Space*, **58**, 351–358.
- Golub, G. & van Loan, C., 1989. *Matrix Computations*, The John Hopkins Univ. Press.
- Holme, R., 2000. Modelling of attitude error in vector magnetic data: application to Ørsted data, *Earth Planets Space*, **52**, 1187–1197.
- Holme, R., Olsen, N. & Bairstow, F., 2011. Mapping geomagnetic secular variation at the core–mantle boundary, *Geophys. J. Int.*, **186**, 521–528.
- Iijima, T. & Potemra, T.A., 1976. The amplitude distribution of field-aligned currents at northern high latitudes observed by TRIAD, *J. geophys. Res.*, **81**, 2165–2174.
- Jenkins, G. & Watts, P., 1968. *Spectral Analysis and Its Application*, Holden Day.
- Kan, J.R. & Lee, L.C., 1979. Energy coupling function and solar wind-magnetosphere dynamo, *Geophys. Res. Lett.*, **6**, 577–580.
- Kuvshinov, A., 2008. 3-D global induction in the oceans and solid earth: recent progress in modeling magnetic and electric fields from sources of magnetospheric, ionospheric and oceanic origin, *Surv. Geophys.*, **29**(2), 139–186.
- Kuvshinov, A.V. & Olsen, N., 2006. A global model of mantle conductivity derived from 5 years of CHAMP, Ørsted, and SAC-C magnetic data, *Geophys. Res. Lett.*, **33**, L18301, doi:10.1029/2006GL027083.
- Kuvshinov, A. & Utada, H., 2010. Anomaly of the geomagnetic sq variation in Japan: effect from 3-D subterranean structure or the ocean effect? *Geophys. J. Int.*, **183**, 1239–1247.
- Laske, G. & Masters, G., 1997. A global digital map of sediment thickness, *EOS, Trans. Am. geophys. Un.*, **78**, F483.
- Lesur, V., 2006. Introducing localized constraints in global geomagnetic field modelling, *Earth Planets Space*, **58**, 477–483.
- Lesur, V., Wardinski, I., Rother, M. & Manda, M., 2008. GRIMM: the GFZ Reference Internal Magnetic Model based on vector satellite and observatory data, *Geophys. J. Int.*, **173**, 382–394.
- Lesur, V., Wardinski, I., Hamoudi, M. & Rother, M., 2010. The second generation of the GFZ reference internal magnetic model: GRIMM-2, *Earth Planets Space*, **62**, 765–773.
- Lesur, V., Rother, M., Vervelidou, F., Hamoudi, M. & Thebault, E., 2013. Post-processing scheme for modeling the lithospheric magnetic field, *Solid Earth*, **4**, 105–118.
- Lowes, F.J., 1966. Mean-square values on sphere of spherical harmonic vector fields, *J. geophys. Res.*, **71**, 2179, doi:10.1029/JZ071i008p02179.
- Macmillan, S. & Olsen, N., 2013. Observatory data and the Swarm mission, *Earth Planets Space*, **65**, 1355–1362.
- Maeda, H., Iyemori, T., Araki, T. & Kamei, T., 1982. New evidence of a meridional current system in the equatorial ionosphere, *Geophys. Res. Lett.*, **9**, 337–340.
- Manoj, C., Kuvshinov, A.V., Maus, S. & Lühr, H., 2006. Ocean circulation generated magnetic signals, *Earth Planets Space*, **58**, 429–437.
- Maus, S., 2010. 'Magnetic field model MF7', Available at: [www.geomag.us/models/MF7.html](http://www.geomag.us/models/MF7.html), last accessed 15 January 2015.
- Maus, S., Rother, M., Hemant, K., Stolle, C., Lühr, H., Kuvshinov, A.V. & Olsen, N., 2006. Earth's crustal magnetic field determined to spherical harmonic degree 90 from CHAMP satellite measurements, *Geophys. J. Int.*, **164**, 319–330.
- Maus, S., Lühr, H., Rother, M., Hemant, K., Balasis, G., Ritter, P. & Stolle, C., 2007. Fifth generation lithospheric magnetic field model from CHAMP satellite measurements, *Geochem. Geophys. Geosyst.*, **8**(5), Q05013, doi:10.1029/2006GC001521.
- Maus, S. et al., 2008. Resolution of direction of oceanic magnetic lineations by the sixth-generation lithospheric magnetic field model from CHAMP satellite magnetic measurements, *Geochem. Geophys. Geosyst.*, **9**(7), Q07021, doi:10.1029/2008GC001949.
- Myers, R., 1990. *Classical and Modern Regression with Applications*, Duxbury Thomson Learning.
- Newell, P.T., Sotirelis, T., Liou, K., Meng, C.-I. & Rich, F.J., 2007. A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables, *J. geophys. Res.*, **112**(A1), doi:10.1029/2006JA012015.
- Olsen, N., 1998. Estimation of C-responses (3 h to 720 h) and the electrical conductivity of the mantle beneath Europe, *Geophys. J. Int.*, **133**, 298–308.
- Olsen, N., Sabaka, T.J. & Lowes, F., 2005. New parameterization of external and induced fields in geomagnetic field modeling, and a candidate model for IGRF 2005, *Earth Planets Space*, **57**, 1141–1149.
- Olsen, N., Lühr, H., Sabaka, T.J., Manda, M., Rother, M., Toffner-Clausen, L. & Choi, S., 2006. CHAOS: a model of Earth's magnetic field derived from CHAMP, Ørsted, and SAC-C magnetic satellite data, *Geophys. J. Int.*, **166**, 67–75.
- Olsen, N. et al., 2013. The Swarm Satellite Constellation Application and Research Facility (SCARF) and Swarm data products, *Earth Planets Space*, **65**, 1189–1200.
- Olsen, N., Lühr, H., Finlay, C.C., Sabaka, T.J., Michaelis, I., Rauberg, J. & Toffner-Clausen, L., 2014. The CHAOS-4 geomagnetic field model, *Geophys. J. Int.*, **197**, 815–827.
- Rangarajan, G.K., 1989. Indices of geomagnetic activity, in *Geomagnetism*, Vol. 3, pp. 323–384, ed. Jacobs, J.A., Academic Press.
- Richmond, A.D., 1995. Ionospheric electrodynamics using magnetic Apex coordinates, *J. Geomagn. Geoelectr.*, **47**, 191–212.
- Rikitake, T., Yokoyama, T. & Sato, S., 1956. Anomaly of the geomagnetic sq variation in Japan and its relation to the subterranean structure, *Bull. earthq. Res. Inst., Univ. Tokyo*, **34**, 197–234.
- Sabaka, T.J. & Olsen, N., 2006. Enhancing comprehensive inversions using the *Swarm* constellation, *Earth Planets Space*, **58**, 371–395.
- Sabaka, T.J., Olsen, N. & Langel, R.A., 2002. A comprehensive model of the quiet-time near-Earth magnetic field: phase 3, *Geophys. J. Int.*, **151**, 32–68.
- Sabaka, T.J., Olsen, N. & Purucker, M.E., 2004. Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data, *Geophys. J. Int.*, **159**, 521–547.
- Sabaka, T.J., Toffner-Clausen, L. & Olsen, N., 2013. Use of the comprehensive inversion method for Swarm satellite data analysis, *Earth Planets Space*, **65**, 1201–1222.
- Schmucker, U., 1985. Sources of the geomagnetic field, in *Landolt-Börnstein, New-Series*, 5/2b, pp. 31–73, eds Fuchs, K. & Soffel, H., Springer-Verlag.
- Schmucker, U., 1999. A spherical harmonic analysis of solar daily variations in the years 1964–1965: response estimates and source fields for global induction-II. Results, *Geophys. J. Int.*, **136**, 455–476.
- Seber, G.A.F. & Wild, C.J., 2003. *Nonlinear Regression*, Wiley-Interscience.
- Simons, F.J., Dahlen, F.A. & Wiczeorek, M.A., 2006. Spatiospectral concentration on a sphere, *SIAM Rev.*, **48**(3), 504–536.
- Thébault, E., Vervelidou, F., Lesur, V. & Hamoudi, M., 2012. The satellite along-track analysis in planetary magnetism, *Geophys. J. Int.*, **188**, 891–907.
- Thomson, A.W.P. & Lesur, V., 2007. An improved geomagnetic data selection algorithm for global geomagnetic field modelling, *Geophys. J. Int.*, **169**(3), 951–963.
- Thomson, A., Hamilton, B., Macmillan, S. & Reay, S., 2010. A novel weighting method for satellite magnetic data and a new global magnetic field model, *Geophys. J. Int.*, **181**, 250–260.
- Toutenburg, H., 1982. *Prior Information in Linear Models*, John Wiley and Sons.
- Tyler, R.H., Maus, S. & Lühr, H., 2003. Satellite observations of magnetic fields due to ocean tidal flow, *Science*, **299**, 239–241.
- Weidelt, P., 1972. The inverse problem of geomagnetic induction, *Z. Geophys.*, **38**, 257–289.
- Zanetti, L.J., Baumjohann, W. & Potemra, T.A., 1983. Ionospheric and Birkeland current distributions inferred from the MAGSAT magnetometer data, *J. geophys. Res.*, **88**, 4875–4884.
- Zanetti, L.J., Baumjohann, W., Potemra, T.A. & Bythrow, P.F., 1984. Three-dimensional Birkeland-ionospheric-current system, determined from MAGSAT, in *Magnetospheric Currents*, Vol. 28, pp. 123–130, ed. Potemra, T.R., American Geophysical Union.

## APPENDIX A: PARAMETER SEPARABILITY METRICS

### A1 Classic parameter correlation coefficient

The classic parameter correlation coefficient (Myers 1990) is introduced by considering a partitioned system similar to that in eq. (20)

$$\mathbf{d} = (\mathbf{A} \quad \mathbf{B}) \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + \boldsymbol{\eta}, \quad (\text{A1})$$

where  $\mathbf{x}$  and  $\mathbf{y}$  are generic parameter vectors,  $\mathbf{A}$  and  $\mathbf{B}$  are matrices whose columns are generic basis vectors, and  $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . If  $\mathbf{B}$  has two columns, then the separability between parameters  $y_1$  and  $y_2$  depends upon the degree of co-linearity between the two columns of

$$\mathbf{B}' = \mathbf{P}_{A\perp} \mathbf{B}, \quad (\text{A2})$$

where

$$\mathbf{P}_{A\perp} = (\mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T) \quad (\text{A3})$$

is the projection matrix onto the null-space of the columns of  $\mathbf{A}$ . In a sense, multiplication by  $\mathbf{P}_{A\perp}$  eliminates information in the columns of  $\mathbf{B}$  that could be attributable to those in  $\mathbf{A}$ . If the columns of  $\mathbf{B}'$  are denoted as  $\mathbf{b}'_1$  and  $\mathbf{b}'_2$  and their corresponding unit vectors as  $\hat{\mathbf{b}}'_1$  and  $\hat{\mathbf{b}}'_2$ , respectively, then this co-linearity is quantified by the direction cosine  $\varrho_{12} = \hat{\mathbf{b}}'_1 \cdot \hat{\mathbf{b}}'_2$ . As  $|\varrho_{12}| \rightarrow 1$ , the co-linearity increases as does the potential for instability due to data perturbations. This can be seen by inspecting the differences between the true values  $y_1$  and  $y_2$  and their LS estimates  $\tilde{y}_1$  and  $\tilde{y}_2$ , respectively, in the above system

$$\begin{pmatrix} \tilde{y}_1 - y_1 \\ \tilde{y}_2 - y_2 \end{pmatrix} = \frac{1}{1 - \varrho_{12}^2} \begin{pmatrix} 1/m_1 & -\varrho_{12}/m_1 \\ -\varrho_{12}/m_2 & 1/m_2 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{b}}'_1 \cdot \boldsymbol{\eta}' \\ \hat{\mathbf{b}}'_2 \cdot \boldsymbol{\eta}' \end{pmatrix}, \quad (\text{A4})$$

where  $m_1 = |\mathbf{b}'_1|_2$ ,  $m_2 = |\mathbf{b}'_2|_2$ , and  $\boldsymbol{\eta}' = \mathbf{P}_{A\perp} \boldsymbol{\eta}$ . Clearly, the errors in the estimates are amplified as  $|\varrho_{12}| \rightarrow 1$ . However, if the components of  $\boldsymbol{\eta}'$  in the directions of  $\hat{\mathbf{b}}'_1$  and  $\hat{\mathbf{b}}'_2$  are small enough, then this effect can be somewhat mitigated and the quality of the estimates will not greatly suffer.

It is important to realize that while high correlations have the potential to lead to contaminated parameter estimates, this is dependent upon how efficiently the errors in the system are accounted for. Indeed, eq. (A4) shows that when high correlations are due to issues with parameter observability from inadequate data distributions, then this can lead to solution instability due to contamination by finite error. However, there is another mechanism that leads to high correlations, but does not cause solution instability. Consider adding smoothing constraints to the system in eq. (A1) such that

$$\begin{pmatrix} \mathbf{d} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{G}_A & \mathbf{G}_B \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta} \\ \mathbf{v} \end{pmatrix}, \quad (\text{A5})$$

where  $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \lambda^{-1} \mathbf{I})$ . If the constraints have a null-space of rank one, then there exists a vector  $\mathbf{s}$  such that

$$\mathbf{G}\mathbf{s} = \mathbf{0}, \quad (\text{A6})$$

where  $\mathbf{G} = (\mathbf{G}_A \quad \mathbf{G}_B)$ . As  $\lambda \rightarrow \infty$ , then eq. (A6) is true and only the bottom partition of eq. (A5) remains. Thus, if  $\mathbf{P}_{G_{A\perp}}$  is defined as in eq. (A3), then following eq. (A2) gives  $\mathbf{g}'_{B_1} = \mathbf{P}_{G_{A\perp}} \mathbf{g}_{B_1}$  and  $\mathbf{g}'_{B_2} = \mathbf{P}_{G_{A\perp}} \mathbf{g}_{B_2}$ , where  $\mathbf{g}_{B_1}$  and  $\mathbf{g}_{B_2}$  are the first and second columns of  $\mathbf{G}_B$ , respectively. However, left-multiplying eq. (A6) by  $\mathbf{P}_{G_{A\perp}}$  leads to  $\mathbf{g}'_{B_1} s_1 = -\mathbf{g}'_{B_2} s_2$ , where  $s_1$  and  $s_2$  are the corresponding elements of  $\mathbf{s}$ . Clearly,  $\mathbf{g}'_{B_1}$  and  $\mathbf{g}'_{B_2}$  are perfectly co-linear. However, the error in the directions of these vectors has vanished. Unlike the contamination in eq. (A4), which is driven by finite error in the direction of the co-linearity, here the co-linearity caused by rank-deficient regularization increases as  $\lambda \rightarrow \infty$ , but the error decreases, allowing for a stable solution.

Typically, co-linearity is deduced by inspection of the correlation coefficient between the  $i$ th and  $j$ th parameters,  $\rho_{ij}$ , which is derived from the corresponding element,  $C_{ij}$ , of the parameter error-covariance matrix,  $\mathbf{C}$ , given by

$$C_{ij} = \sigma_i \sigma_j \rho_{ij}, \quad (\text{A7})$$

where  $\sigma_i = \sqrt{C_{ii}}$  and  $\sigma_j = \sqrt{C_{jj}}$  are the formal errors in the  $i$ th and  $j$ th parameters, respectively. The relationship between  $\varrho_{ij}$  and  $\rho_{ij}$  can be determined by first rearranging the basis vectors in eq. (A1) so that the  $i$ th and  $j$ th vectors comprise  $\mathbf{B}$  and correspond to  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , respectively. The resulting error-covariance matrix corresponding to the  $i$ th and  $j$ th parameters is then given by

$$\mathbf{C}' = (\mathbf{B}'^T \mathbf{B}')^{-1}, \quad (\text{A8})$$

$$= \frac{1}{1 - \varrho_{ij}^2} \begin{pmatrix} 1/m_i^2 & -\varrho_{ij}/(m_i m_j) \\ -\varrho_{ij}/(m_i m_j) & 1/m_j^2 \end{pmatrix}, \quad (\text{A9})$$

where it can be seen that

$$\sigma_i = \frac{1}{m_i \sqrt{1 - \rho_{ij}^2}}, \quad \sigma_j = \frac{1}{m_j \sqrt{1 - \rho_{ij}^2}}, \quad \rho_{ij} = -\rho_{ji}. \quad (\text{A10})$$

One of the most effective ways of detecting parameter insensitivity due to inadequate data distribution is to examine the eigen-decomposition of  $\mathbf{C}$  given by

$$\mathbf{C} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T, \quad (\text{A11})$$

where  $\mathbf{\Lambda}$  is a diagonal matrix whose  $i$ th diagonal  $\lambda_i \geq 0$  is an eigenvalue and  $\mathbf{Q}$  is an orthogonal matrix whose  $i$ th column  $\mathbf{Q}_i$  is the corresponding eigenvector. The most poorly (well) observed linear combination of parameters is given by the eigenvector elements corresponding to the largest (smallest) eigenvalue of  $\mathbf{C}$ . The correlation coefficient  $\rho_{ij}$  is then related to this decomposition by

$$\rho_{ij} = \frac{\sum_k \lambda_k (\mathbf{Q}_i)_k (\mathbf{Q}_j)_k}{\sqrt{(\sum_k \lambda_k (\mathbf{Q}_i)_k^2) (\sum_k \lambda_k (\mathbf{Q}_j)_k^2)}}. \quad (\text{A12})$$

If the  $K$ th eigenvalue/eigenvector pair are dominant in the summations, then eq. (A12) reduces to

$$\rho_{ij} = \frac{(\mathbf{Q}_i)_K}{|(\mathbf{Q}_i)_K|} \cdot \frac{(\mathbf{Q}_j)_K}{|(\mathbf{Q}_j)_K|}, \quad (\text{A13})$$

which leads to correlations of  $\rho_{ij} = \pm 1$ .

## A2 Parameter subspace correlation coefficient

An alternative to the classic pair-wise parameter correlation coefficient is one that corresponds to the direction cosine between a parameter basis vector and a subspace. This can provide some information on separability and is useful in the case where full error-covariance matrices are not available. It may be derived by rewriting the system in eq. (A1) as

$$\mathbf{d} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{b} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ y \end{pmatrix} + \boldsymbol{\eta}, \quad (\text{A14})$$

where  $\mathbf{x}_1$ ,  $\mathbf{x}_2$  and  $y$  are generic parameters, the columns of  $\mathbf{A}_1$ ,  $\mathbf{A}_2$  and  $\mathbf{b}$  are generic basis vectors, and  $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . Define  $\mathbf{P}_{\mathbf{A}_{1\perp}}$  as in eq. (A3) and let  $\mathbf{A}'_2 = \mathbf{P}_{\mathbf{A}_{1\perp}} \mathbf{A}_2$  and  $\mathbf{b}' = \mathbf{P}_{\mathbf{A}_{1\perp}} \mathbf{b}$ . If  $\mathbf{P}_{\mathbf{A}'_{2\perp}}$  is defined as in eq. (A3) and  $\mathbf{b}'' = \mathbf{P}_{\mathbf{A}'_{2\perp}} \mathbf{b}'$ , then the formal error in  $y$  is  $\sigma_y = 1/m''$ , where  $m'' = |\mathbf{b}''|_2$ . If  $\mathbf{x}_2$  is excluded from consideration in the above system, then the formal error in  $y$  is  $\sigma_{y|\mathbf{x}_2} = 1/m'$ , where  $m' = |\mathbf{b}'|_2$ . What is of interest is the ratio of the sigmas in  $y$  when  $\mathbf{x}_2$  is excluded to when it is not, which gives

$$\frac{\sigma_{y|\mathbf{x}_2}}{\sigma_y} = \sqrt{\frac{\mathbf{b}'^T \left( \mathbf{I} - \mathbf{A}'_2 (\mathbf{A}'_2{}^T \mathbf{A}'_2)^{-1} \mathbf{A}'_2{}^T \right) \mathbf{b}'}{\mathbf{b}'^T \mathbf{b}'}}, \quad (\text{A15})$$

$$= \sin \theta, \quad (\text{A16})$$

where  $\theta$  is the angle between  $\mathbf{b}'$  and its projection onto the column space of  $\mathbf{A}'_2$ . The corresponding direction cosine or correlation coefficient, denoted  $\rho'_{y|\mathbf{x}_2}$ , is then given by

$$\rho'_{y|\mathbf{x}_2} = \sqrt{1 - \frac{\sigma_{y|\mathbf{x}_2}^2}{\sigma_y^2}}. \quad (\text{A17})$$

If  $\rho'_{y|\mathbf{x}_2} = 1$ , then  $\mathbf{b}'$  is completely spanned by the columns of  $\mathbf{A}'_2$ , which is that part of  $\mathbf{A}_2$  unique from  $\mathbf{A}_1$ , and can be thought of as completely co-linear with this parameter subspace. Conversely, if  $\rho'_{y|\mathbf{x}_2} = 0$ , then  $\mathbf{b}'$  is orthogonal to this subspace, that is, has a co-linearity of zero. It can be seen that  $\rho'_{y|\mathbf{x}_2} \rightarrow 1$  when the formal error in  $y$  greatly increases due to the consideration of the  $\mathbf{x}_2$  parameters.

## APPENDIX B: OHM BIAS CORRELATION MECHANISM

Consider the OHM biases for the same component in two time segments of the same station or two stations in close proximity. Their basis vectors are orthogonal, yet after projecting out the other basis vector information they become highly positively correlated. Though perhaps non-intuitive, the reason for this is the commonality of the lithospheric field. This can be seen by again considering the system in eq. (A1). Let

$\ell$  denote the total number of station components, including segments, and  $K_i$  be the number of measurements for the  $i$ th station component. Without loss of generality, if the last two components are those of interest, then define

$$\mathbf{A} = \begin{pmatrix} \mathbf{S} \\ \mathbf{1}_{K_1} \mathbf{g}_1^T \\ \vdots \\ \mathbf{1}_{K_{\ell-1}} \mathbf{g}_{\ell-1}^T \\ \mathbf{1}_{K_\ell} \mathbf{g}_\ell^T \end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{1}_{K_{\ell-1}} \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{1}_{K_\ell} \end{pmatrix}, \quad (\text{B1})$$

where  $\mathbf{A}$ ,  $\mathbf{b}_1$  and  $\mathbf{b}_2$  represent the basis vectors for the lithosphere and the biases of interest, respectively, and  $\mathbf{1}_{K_{\ell-1}}$  and  $\mathbf{1}_{K_\ell}$  are vectors of ones of length  $K_{\ell-1}$  and  $K_\ell$ , respectively. The  $\mathbf{g}_i$  vectors and the rows of  $\mathbf{S}$  correspond to the SH lithospheric expansions for the various station components and the satellite data, respectively.

Towards the application of eq. (A2), let  $\mathbf{G}$  be a matrix whose columns are the set of unique  $\mathbf{g}_i$ , let  $\mathbf{\Lambda}$  be a diagonal matrix whose  $i$ th diagonal element corresponds to the cumulative number of measurements associated with the unique  $\mathbf{g}_i$ , let  $\mathbf{E} = \mathbf{S}^T \mathbf{S}$ , and define

$$\mathbf{Y} = (\mathbf{E} + \mathbf{G} \mathbf{\Lambda} \mathbf{G}^T)^{-1} \mathbf{G}, \quad (\text{B2})$$

$$\mathbf{H} = \mathbf{G}^T \mathbf{Y}, \quad (\text{B3})$$

$$= \mathbf{G}^T \mathbf{E}^{-1} \mathbf{G} (\mathbf{I} + \mathbf{\Lambda} \mathbf{G}^T \mathbf{E}^{-1} \mathbf{G})^{-1}, \quad (\text{B4})$$

$$\approx \mathbf{\Lambda}^{-1}, \quad (\text{B5})$$

where eq. (B5) becomes more precise as the diagonals of  $\mathbf{\Lambda}$  become large. Similarly, define

$$\mathbf{J} = \mathbf{Y}^T \mathbf{E} \mathbf{Y}, \quad (\text{B6})$$

$$= (\mathbf{I} + \mathbf{\Lambda} \mathbf{G}^T \mathbf{E}^{-1} \mathbf{G})^{-T} \mathbf{H}, \quad (\text{B7})$$

$$\approx \mathbf{\Lambda}^{-1} (\mathbf{G}^T \mathbf{E}^{-1} \mathbf{G})^{-1} \mathbf{\Lambda}^{-1}, \quad (\text{B8})$$

where the superscript ‘-T’ denotes the transpose of the inverse and eq. (B8) becomes more precise as the diagonals of  $\mathbf{\Lambda}$  become large. Applying eq. (A2) to eq. (B1) leads to

$$\mathbf{b}'_1 = \begin{pmatrix} -K_{\ell-1} \mathbf{S} \mathbf{y}_{\mu(\ell-1)} \\ -K_{\ell-1} H_{\mu(1)\mu(\ell-1)} \mathbf{1}_{K_1} \\ \vdots \\ 1 - K_{\ell-1} H_{\mu(\ell-1)\mu(\ell-1)} \mathbf{1}_{K_{\ell-1}} \\ -K_{\ell-1} H_{\mu(\ell)\mu(\ell-1)} \mathbf{1}_{K_\ell} \end{pmatrix}, \quad \mathbf{b}'_2 = \begin{pmatrix} -K_\ell \mathbf{S} \mathbf{y}_{\mu(\ell)} \\ -K_\ell H_{\mu(1)\mu(\ell)} \mathbf{1}_{K_1} \\ \vdots \\ -K_\ell H_{\mu(\ell-1)\mu(\ell)} \mathbf{1}_{K_{\ell-1}} \\ 1 - K_\ell H_{\mu(\ell)\mu(\ell)} \mathbf{1}_{K_\ell} \end{pmatrix}, \quad (\text{B9})$$

where  $\mathbf{y}_i$  is the  $i$ th column of  $\mathbf{Y}$ , and  $\mu(i)$  maps the  $i$ th station component to the set of unique  $\mathbf{g}_i$ .

For the case of a particular component at a station in two different time segments, then  $\mu(\ell-1) = \mu(\ell)$ . Under the assumption that eq. (B5) holds, then eq. (B9) becomes

$$\mathbf{b}'_1 = \begin{pmatrix} -K_{\ell-1} \mathbf{S} \mathbf{y}_{\mu(\ell)} \\ \mathbf{0} \\ \vdots \\ \left( \frac{K_\ell}{K_{\ell-1} + K_\ell} \right) \mathbf{1}_{K_{\ell-1}} \\ - \left( \frac{K_{\ell-1}}{K_{\ell-1} + K_\ell} \right) \mathbf{1}_{K_\ell} \end{pmatrix}, \quad \mathbf{b}'_2 = \begin{pmatrix} -K_\ell \mathbf{S} \mathbf{y}_{\mu(\ell)} \\ \mathbf{0} \\ \vdots \\ - \left( \frac{K_\ell}{K_{\ell-1} + K_\ell} \right) \mathbf{1}_{K_{\ell-1}} \\ \left( \frac{K_{\ell-1}}{K_{\ell-1} + K_\ell} \right) \mathbf{1}_{K_\ell} \end{pmatrix}. \quad (\text{B10})$$

The  $\ell_2$  norms of the top partitions of  $\mathbf{b}'_1$  and  $\mathbf{b}'_2$  are given by  $K_{\ell-1} \sqrt{J_{\mu(\ell)\mu(\ell)}}$  and  $K_\ell \sqrt{J_{\mu(\ell)\mu(\ell)}}$ , respectively. Under the assumption that eq. (B8) holds, then these both become very small compared to the last two partitions. Thus,  $\mathbf{b}'_1$  and  $\mathbf{b}'_2$  become nearly antiparallel leading to

$\rho_{ij} \approx 1$ . The case of a particular component at two nearby stations is closely approximated by this same development, thus resulting in similar correlations.

## APPENDIX C: FORWARD MODELLING METHODS FOR PREDICTING THE OCEANIC $M_2$ MAGNETIC FIELD

The  $M_2$  field extracted from CM5 is validated through comparison with forward-model results from two numerical models that have been previously described in Tyler *et al.* (2003) and Kuvshinov (2008). For the study here, both models have been reconfigured with updated parameters: both models use the tidal flow of the data-assimilated TPXO7.2 global tidal model (<http://volkov.oce.orst.edu/tides/>), and both use the inhomogeneous surface conductance described in Section 3.1.3; the Tyler model uses the main magnetic field of CM5 evaluated at 2007.0, while the Kuvshinov model uses the main field evaluated at 2007.0 from the WMM magnetic model (<http://www.ngdc.noaa.gov/geomag/WMM/DoDWMM.shtml>). At the level of precision required for the comparisons in this study, the difference in the two model results due to the different input parameters is expected to be small. A larger difference between the two model results might have been anticipated from the quite different numerical approaches used in each model and the associated approximations involved. This would of course reveal inaccuracies that are purely due to the different approximate methods used. In this study, an important cross-validation in the predicted fields has been obtained because at the level of comparison provided there has been no discernible difference in the two predicted fields that can be attributed to the different methodologies. Rather, the two fields are remarkably similar with an amplitude difference that is fully expected due to the different assumptions for the conductivity in the upper mantle. These methods and assumptions are described in more detail below.

The method of Tyler formally implements a 2-D thin-shell version of the EM induction equation within the inhomogeneous layer comprising the ocean and conductive sediments. This 2-D equation serves as the boundary equation for the magnetic variables in the adjacent 3-D domains, which are solved through SH expansion. The 2-D equation is discretized on a flux-conserving finite difference grid involving longitude/latitude spherical coordinates. In this study, the full 1/4 degree resolution of the TPXO7.2 tidal flow data was used in the finite-difference equation. The resolution is ultimately controlled by the truncation degree of the SH expansion in the 3-D component, which was chosen to be degree  $n = 300$  for consistent comparison with the Kuvshinov results resolved on a  $1^\circ \times 1^\circ$  spatial grid. The higher resolution of the finite-difference component, even with reduced resolution through a lower truncation degree, avoids interpolation errors as the sources are provided directly on the native grid and the same flux-conserving scheme for both flow and electric charge is rigorously obeyed.

The region above the conductive surface layer is treated as an electrical insulator and as such the magnetic field can be derived from a potential which can be upward continued through this region either using finite differences, as in Tyler *et al.* (2003), or SH expansion, as in this study. While it is possible to add coupled elements allowing for conductive layers below the surface, the purpose of the Tyler model has primarily been to produce solutions describing clear limiting-case scenarios. That is, while the conductivity of the ocean is known rather precisely, the conductivity of the mantle is not. As such, model solutions become convolved with uncertainties in these assumptions and the Tyler model is run in a configuration to provide the limiting case of an insulating upper mantle. It is assumed that the effect of the conductive mantle will be to reduce the amplitudes of the  $M_2$  fields at the surface and therefore the insulating case should provide an appropriate end-member. The CM5  $M_2$  amplitudes indeed appear to fall below this end-member for degrees  $n < 18$ , as expected. Although the Tyler model has an insulating upper mantle, examination of Fig. 3(b) shows that the conductivity of the lower mantle is quite high and it is reasonable to assume a perfectly reflecting boundary condition, just as in Section 3.1.3. While the depth of the reflector is taken to be at 1000 km depth for the lower frequencies in Section 3.1.3, the depth chosen for the higher  $M_2$  frequency in the Tyler model is 800 km.

The model of Kuvshinov (2008) solves an integral equation governing EM induction in a 3-D domain. This model estimates both the electric and magnetic fields generated by prescribed electric-field sources within a 3-D distribution of electrical conductivity. In the application here, this distribution consists of a surface spherical shell of laterally varying conductance underlain by a radially symmetric conductor (see more details about this model in Section 3.1.3), and the extraneous source is described by the sheet current density, which is calculated as

$$\mathbf{J}_\tau^{\text{ext}} = \sigma_w (\mathbf{U} \times \mathbf{B}^m), \quad (\text{C1})$$

where  $\sigma_w = 3.2 \text{ S m}^{-1}$  is the mean conductivity of sea water,  $\mathbf{U}$  is the depth-integrated velocity due to  $M_2$  tidal oscillations taken from the TPXO7.2 global tidal model and  $\mathbf{B}^m$  is the main magnetic field evaluated at 2007.0 from the WMM magnetic model. The simulations are made on a mesh with a spatial resolution of  $1^\circ \times 1^\circ$ .