

# The BGS magnetic field candidate models for the 11th generation IGRF

Brian Hamilton, Susan Macmillan, and Alan Thomson

British Geological Survey (NERC), West Mains Road, Edinburgh, EH9 3LA, UK

(Received November 30, 2009; Revised May 20, 2010; Accepted May 20, 2010; Online published December 31, 2010)

We describe the British Geological Survey's 11th generation International Geomagnetic Reference Field candidate models. These models are based on a 'parent model' consisting of a degree and order 60 spherical harmonic expansion of selected vector and scalar magnetic field data from satellite and observatory sources within the period 1999.0 to 2010.0. The parent model's internal field time dependence for degrees 1 to 13 is represented by linear spline with knots 400 days apart. The parent model's degree 1 external field time dependence is described by periodic functions for the annual and semi-annual signals, and by dependence on the 20-minute Vector Magnetic Disturbance index. Signals induced by these external fields are also parameterised. Satellite data are weighted according to two noise estimators. Firstly by standard deviation along segments of the satellite track and secondly a larger-scale noise estimator defined in terms of a vector activity measure at the geographically closest magnetic observatories to the sample point.

**Key words:** IGRF, geomagnetic field, geomagnetic secular variation.

## 1. Introduction

The International Geomagnetic Reference Field (IGRF) is a widely used model of the Earth's main magnetic field and is produced every 5 years by a collaboration of modelling teams at the invitation of the International Association of Geomagnetism and Aeronomy (IAGA) Division V committee. The current IGRF (version 10, Macmillan and Maus, 2006) is now due for revision. IGRF-11, the 11th generation model, will include three new sets of coefficients: two main field (MF) models to degree and order 13 at epochs 2005.0 and 2010.0 and an average secular variation (SV) model valid from 2010.0 to 2015.0 to degree and order 8. The production of our BGS candidate models was carried out in three steps: selection of data from Ørsted, CHAMP, and ground observatories; fitting and evaluating the parent model; and finally extraction of the IGRF-11 candidate MF and SV coefficients. Since the production of the BGS IGRF-10 candidate models (Lesur *et al.*, 2005), the most significant changes to our parent model have been to include more detailed parametrisation of external sources and the use of vector data at all latitudes with appropriate noise estimation.

From the time Ørsted was launched in 1999 and CHAMP the following year, a large number of vector and scalar measurements with good geographic coverage have been made. Over the same period, at ground level, a global network of observatories has been collecting vector measurements at fixed geographic locations. Together these data provide a large data set for MF modelling. However, in addition to the core and large scale lithospheric sources that are the focus of the IGRF-11 modelling effort, this large compilation of satellite and observatory data contain fields produced from

other sources. In particular, the day-side ionospheric current system, ring-current, partial ring current, and induced currents in the Earth can be difficult to co-estimate and separate from the desired signals. To avoid contamination, measurements significantly affected by these sources must be removed from the data set through careful selection. We use a combination of local-time, magnetic indices, and measurements of the interplanetary magnetic field (IMF) and solar wind speed in the near-Earth environment to reject data likely to be contaminated by unwanted sources. Unfortunately, the local-time (night-side) data selection tends to result in gaps in the satellite time-series. It is conceivable that this uneven temporal distribution could result in spurious signals appearing in the time parametrisation of the parent model. We seek to address this problem by including observatory data, which helps provide a more continuous time-series.

However, even with rigorous data selection, contamination from unwanted sources remains a problem, particularly at high latitudes. For this reason, it has often been deemed preferable to use only scalar data at high latitudes rather than use vector components that are more sensitive to high latitude current systems. We still use only observatory pseudo-scalar data (vector data projected onto an *a priori* main field model) at higher latitudes, which also has the effect of making the co-estimation of the observatory biases easier by maintaining the linear relationship between data and model coefficients. However, we take a different approach with satellite data and use vector components in addition to scalar data at all latitudes wherever they are available. At the same time, we quantify the increase in noise in the data when we fit our parent model. This is achieved by using a combination of two noise estimators: along-track standard deviations calculated for each component at intervals from short segments of the satellite data,

and a larger scale estimator of the vector disturbance using observatories nearest to the data sample. This allows us to produce a robust parent model without the need for a particularly complex model parametrisation, regularisation, or prior data correction to remove estimates of unmodelled source fields.

The parent model's internal field time-dependence for degrees 1 to 13 is by linear spline with knots 400 days apart starting from mid-1999. The parent model's degree 1 external field time-dependence is also modelled by the same linear spline arrangement but includes periodic functions for the annual and semi-annual signals, and a dependence on the 20-minute Vector Magnetic Disturbance index (VMD, Thomson and Lesur, 2007). Signals induced by these external fields are also parametrised. Observatory biases are co-estimated when fitting the parent model.

In Section 2 we describe the data used, their selection criteria, and briefly outline the data weighting scheme. In Section 3 we describe the parent model and the process used to fit its parameters, while in Section 4 we evaluate the model coefficients. In Section 5 we extrapolate forwards in time and extract the IGRF-11 MF and SV candidate models. We provide concluding remarks in Section 6.

## 2. Data Selection and Weighting

We use CHAMP (calibration level version 51) scalar and vector data between 2000.6 and 2009.6. Ørsted scalar and vector data are used between 1999.2 and 2009.4. For Ørsted after 2005.9 only scalar data are available. To reduce the quantity of data to manageable levels, the satellite data were sub-sampled at every 60th datum giving a sample frequency of approximately once every 68 seconds.

The following selection criteria were then applied to the satellite data:

- 1) Magnetic indices:  $K_p$  and  $K_p$  for previous 3 hours both  $\leq 2-$ ;  $|dVMD/dt| \leq 5$  nT/hour;  $IE \leq 30$  nT;  $PC \leq 0.2$  mV/m;
- 2) Solar wind data:  $0 \leq IMF B_z \leq +6$  nT;  $-3 \leq IMF B_y \leq +3$  nT;  $-10 \leq IMF B_x \leq +10$  nT; solar wind speed  $\leq 450$  km/s;
- 3) Other:  $22:30 \leq$  local time (hour: min)  $\leq 05:00$ ;  $|B_{obs} - B_a| \leq 100$  nT;  $|F_{Over} - F_{flux}| \leq 2$  nT

where VMD is the vector magnetic disturbance index (see Section 3).  $B_{obs}$  is the observed magnetic field value and  $B_a$  is the corresponding value from an *a priori* main field model developed at BGS.  $F_{Over}$  is the total magnetic field value from the Overhauser magnetometer and  $F_{flux}$  is the corresponding value from the fluxgate magnetometer.

Figure 1(a) and (b) show the temporal distribution of the final selected CHAMP and Ørsted data sets, respectively. From these graphs, it is clear that there are gaps in the time-series, which are due mostly to the rejection of daytime data.

Observatory data were taken from 152 observatories between 1999.0 and 2009.5 held at the World Data Centre for Geomagnetism. Known jumps were corrected for and some poor quality data were eliminated manually. The observatory hourly means were then selected according to the following criteria:

- 1) Magnetic indices:  $K_p \leq 2+$ ,  $|dD_{st}/dt| \leq 5$  nT/hour
- 2) Solar wind data:  $IMF B_z \geq 0$  nT
- 3) Other: night-time (01:00 to 02:00 LT + solar zenith angle test at 110 km altitude above observatory)

For high geomagnetic latitudes ( $>50$  and  $<-50$  degrees) observatory data were projected onto an *a priori* model vector and the resulting pseudo-scalar data were used in the inversion. For other latitudes, vector data were used.

Figure 1(c) shows the time-distribution of the final selected observatory data set. It can be seen that the observatory data provide a more continuous time-series than the satellite data, which is desirable for the robust time parametrisation of the model.

The final data set consists of 186,468 Ørsted data, 654,720 CHAMP data, and 357,175 observatory data giving a total of 1,198,363 individual data points (counting each vector component as one data point). The temporal distribution of the combined data set is shown in Fig. 1(d).

The selected satellite data were individually weighted using two "noise" estimators (for a detailed description, see Thomson *et al.*, 2010). Firstly, a measure of local magnetic activity using the standard deviation (SD) along short segments (60 samples, approximately 500 km) of satellite track. Secondly, a larger-scale noise estimator derived from activity measured at the geographically closest magnetic observatories to the sample point. This results in the down-weighting of the vector data at high latitudes, particularly in the auroral zones. We find the technique produces a core and lithospheric field model, MEME08, that compares well with other similar models such as xCHAOS (Olsen and Manda, 2008), GRIMM (Lesur *et al.*, 2008), and MF6 (Maus *et al.*, 2008) up to about degree and order 60. We also apply weighting based on instrument accuracy and zenith angle (to account for ionospheric field contamination) up to a maximum of 10 nT, and also one-degree tesseral weighting to account for different spatial data densities, in the same manner as Lesur *et al.* (2005).

For the observatory data weighting, we use a simpler scheme based on a scaling of the instrument accuracy and zenith angle weighting. The application of LAVA and along-track SD has the effect of significantly down-weighting the satellite data. To prevent the observatory data dominating the model post LAVA/SD, we apply a scaling to the observatory weights. This scaling is chosen such that the mean satellite-observatory data weighting post LAVA/SD is the same as that found before LAVA and along-track SD are included.

As a rough estimate of the weight carried by each data source (CHAMP, Ørsted, observatories) when fitting the model, we divide the number of data from each source that meet the selection criteria (defined above) by the mean of the diagonal entries of the data covariance matrix for that source (Ørsted's covariance matrix contains off-diagonal entries to account for the anisotropic error introduced due to the presence of only one star camera, see Lesur *et al.*, 2005). From this we determine the ratio of the weight of (CHAMP: Ørsted: observatories) data in the model to be roughly (2.4:1:2.9). This ratio is provided as a rough guide to the combined effect of data number and weighting but

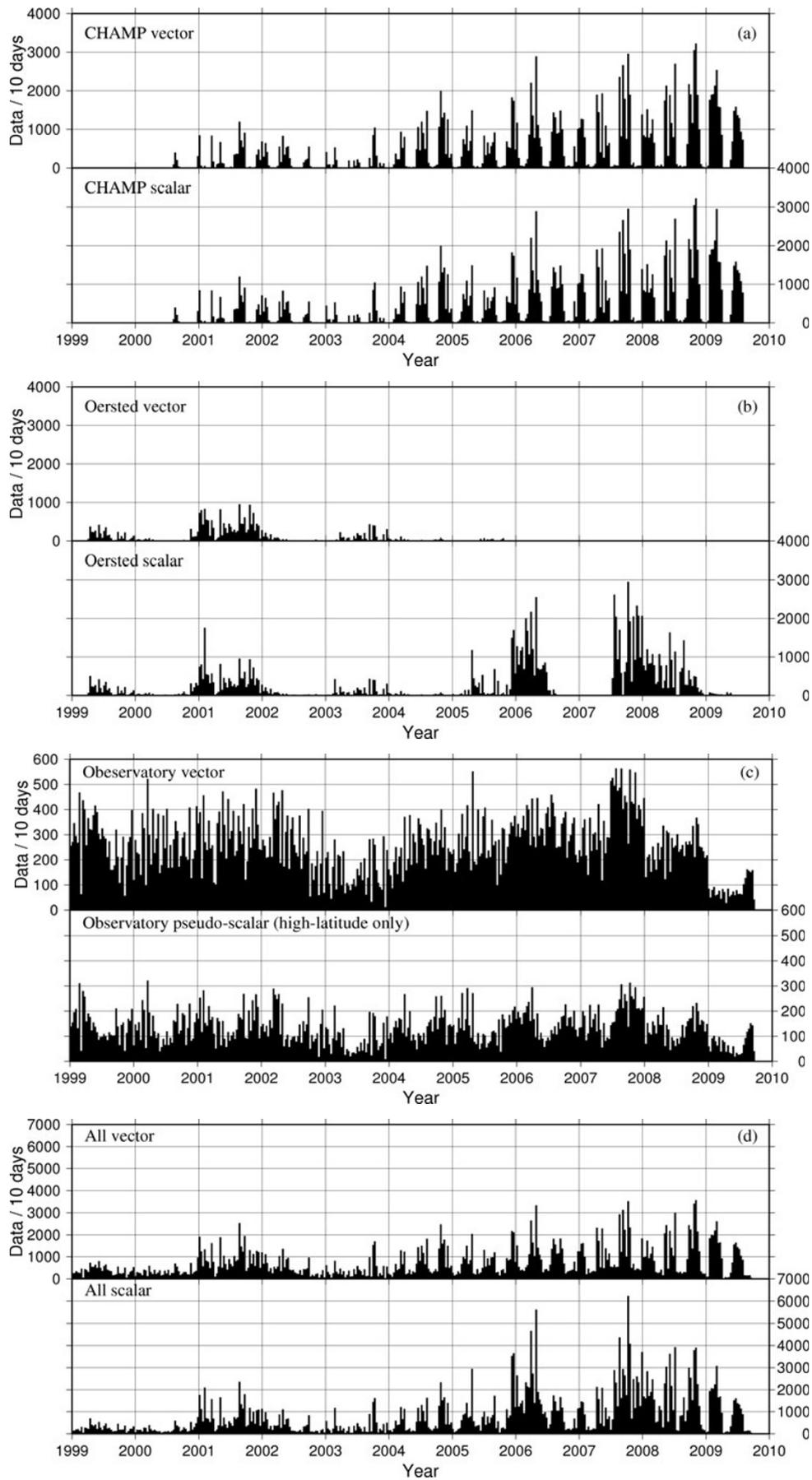


Fig. 1. Temporal distribution of selected data used to fit the parent model from CHAMP (a), Ørsted (b), observatories (c), and the combined data set (d). In these plots, each triplet of vector components is counted as a single measurement.

Table 1. Mean and root mean square misfits between parent model estimate and input data, shown separately for data sources and by geomagnetic latitude. Mid and low geomagnetic latitudes are defined as being between  $\pm 50$  degrees, with high latitudes defined as being outside this range. Satellite  $F$  misfits are calculated with respect to scalar input data. Observatory  $F$  misfits are calculated with respect to vector data projected onto an *a priori* model. The number of data quoted counts each component separately.

Source	No. of data	Mean (nT)				Root mean square (nT)			
		$X$	$Y$	$Z$	$F$	$X$	$Y$	$Z$	$F$
Mid and low latitudes									
CHAMP	438,898	-0.11	0.13	0.00	-0.09	5.37	2.74	5.17	3.63
Ørsted	123,066	-0.69	-0.50	0.10	1.10	7.38	5.86	5.94	3.75
Observatories	304,648	-0.29	0.18	-0.14	—	7.18	5.15	7.09	—
High latitudes									
CHAMP	215,822	1.57	0.52	-1.52	-4.20	18.83	19.09	12.23	11.63
Ørsted	63,402	2.90	1.03	0.47	-0.43	19.30	19.94	7.23	5.59
Observatories	52,527	—	—	—	-0.46	—	—	—	16.30
All latitudes									
CHAMP	654,720	0.44	0.26	-0.50	-1.44	11.67	11.19	8.20	7.31
Ørsted	186,468	0.61	0.06	0.24	0.61	13.03	12.88	6.44	4.43

the precise influence of data from different sources will vary with time and model parameters. With this caveat in mind we can see that of the three sources, CHAMP and observatories contribute most to the model estimation. Overall, satellite data carries more weight than observatory data when fitting the parent model. Observatory data plays an important role in the time-parameterisation of our model and we have given them a significant weighting. However, the optimum balance between satellite and observatory data remains an issue and will be considered further in future work.

### 3. Model Parametrisation

For the parent model we use the same parametrisation as Thomson and Lesur (2007) but with different knot spacings for the piece-wise linear time-dependence. This parametrisation is set out explicitly below and consists of four main terms with a total of 6315 separate parameters:

$$V = V^{\text{int}} + V^{\text{ext}} + V^{\text{vmd}} + V^{\text{ann}}$$

where  $V^{\text{int}}$  and  $V^{\text{ext}}$  are the scalar potentials associated with the internal and external terms, respectively, that have either piecewise-linear time-dependence or are constant.  $V^{\text{vmd}}$  contains terms that are linearly dependent on the VMD index. This index (Thomson and Lesur, 2007) is interpolated at the time of each datum from a degree 1 spherical harmonic model. This model is fitted in Earth-centred Earth-fixed and sun-synchronous coordinate systems with time dependence given by cubic B-splines with knots every hour, produced from observatory data.  $V^{\text{ann}}$  consists of terms that model the annual, semi-annual, and diurnal variations. The various terms are expanded below:

$$V^{\text{int}}(\theta, \phi, r, t) = a \sum_{l=1}^{l=13} \sum_{m=-l}^{+l} g_l^m(t) \left(\frac{a}{r}\right)^{l+1} Y_l^m(\theta, \phi)$$

$$+ a \sum_{l=14}^{l=60} \sum_{m=-l}^{+l} g_l^m \left(\frac{a}{r}\right)^{l+1} Y_l^m(\theta, \phi)$$

$$V^{\text{ext}}(\theta, \phi, r, t) = a \sum_{m=-1}^{+1} q_1^m(t) \left(\frac{r}{a}\right) Y_1^m(\theta, \phi)$$

where  $\theta$ ,  $\phi$ ,  $r$ , and  $t$  are the geocentric colatitude and longitude, radius from centre of Earth, and time, respectively. The reference radius of the Earth is denoted by  $a$ .  $g_l^m(t)$  and  $q_l^m(t)$  are the piecewise-linear time-dependent internal and external Gauss coefficients, respectively and  $g_l^m$  are the static internal coefficients. The knots of the piecewise linear terms start at decimal year 1999.5 and thereafter at intervals of 400 days resulting in 10 knots in total with the final knot at approximately 2009.4.  $Y_l^m(\theta, \phi)$  are the Schmidt semi-normalised spherical harmonic functions of degree  $l$  and order  $m$ . We use the convention that zero and positive  $m$  correspond to  $\cos m\phi$  terms while negative  $m$  correspond to  $\sin m\phi$  terms.

The terms dependent on the VMD index are given below:

$$V^{\text{vmd}}(\theta, \phi, r, t, \text{VMD}) = a \sum_{m=-1}^{+1} \left\{ [b_{1,m} + t b_{2,m} + b_{3,m} \text{VMD}_m^{\text{ext}}(t)] \left(\frac{r}{a}\right) + c_m \text{VMD}_m^{\text{int}}(t) \left(\frac{a}{r}\right)^2 \right\} Y_1^m(\theta, \phi)$$

where  $b_{1,m}$  and  $b_{2,m}$  are the constant and linear external coefficients while  $b_{3,m}$  are the coefficients that are directly scaled by the external VMD indices,  $\text{VMD}_m^{\text{ext}}$ . There are no constant or linearly dependent induced coefficients but  $c_m$  are defined in the same way as  $b_{3,m}$ . The annual, semi-annual, and 24-hour time-dependencies are parametrised by:

$$V^{\text{ann}}(\theta, \phi, r, t) = a \sum_{m=-1}^{+1} \left[ e_m(t) \left(\frac{r}{a}\right) + i_m(t) \left(\frac{a}{r}\right)^2 \right] \cdot Y_1^m(\theta, \phi) + a \sum_{m=-1}^{+1} \epsilon_m(t) \left(\frac{r}{a}\right) Y_1^m(\theta, \tilde{\phi})$$

where  $e_m$  and  $i_m$  are, respectively, the external and internal coefficients consisting of cos and sin annual and semi-annual terms.  $\epsilon_m$  is the external 24-hour term, which is expressed in terms of sun-synchronous longitude,  $\tilde{\phi}$ , rather than geographic longitude (Thomson and Lesur, 2007) and consists of a constant, and cos and sin annual terms.

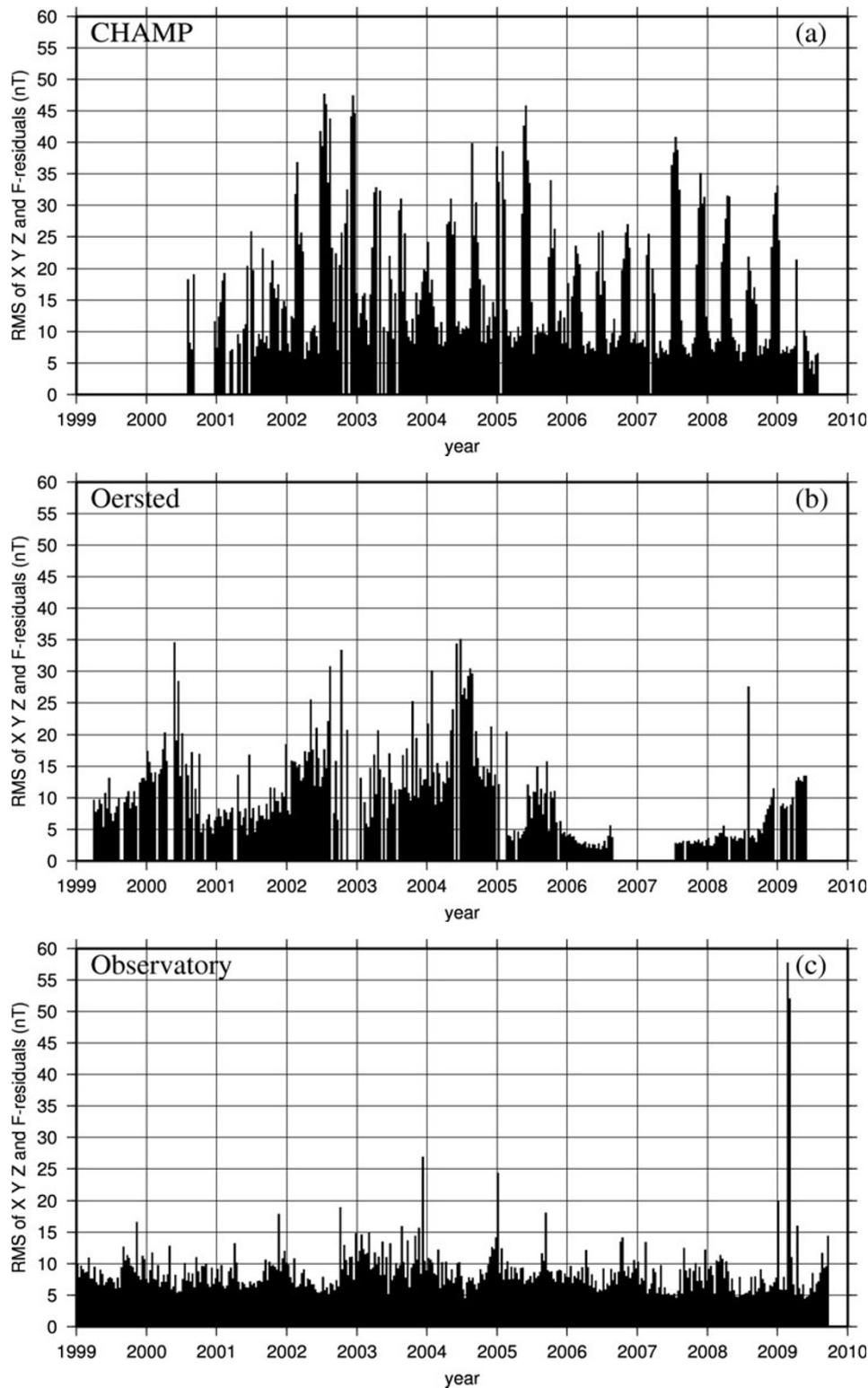


Fig. 2. Root-mean-square of residuals between model and input data from (a) CHAMP, (b) Ørsted, and (c) observatories. The RMS values are calculated over 10 day periods using all components ( $X$ ,  $Y$ ,  $Z$ , and  $F$ ). To improve robustness of statistics, ten day periods with fewer than 10 data are not plotted.

#### 4. Model Estimation

The model described in Section 3 is fitted to the data using an iterative reweighted least-squares technique. A model vector with the only non-zero main-field coefficient being  $g_1^0 = -10000$  nT is used as the starting point of the initial L2-norm fit. The results of this fit are then used as the starting point for 3 iterations of an L1-norm fit to produce

the final model vector (previous experience suggests three iterations are sufficient for our models). The (unweighted) mean and root mean square (RMS) of the misfits are summarised in Table 1.

Direct comparison between CHAMP, Ørsted and observatory misfits is not straightforward. Although vector and scalar data are used at all latitudes for satellite data, scalar

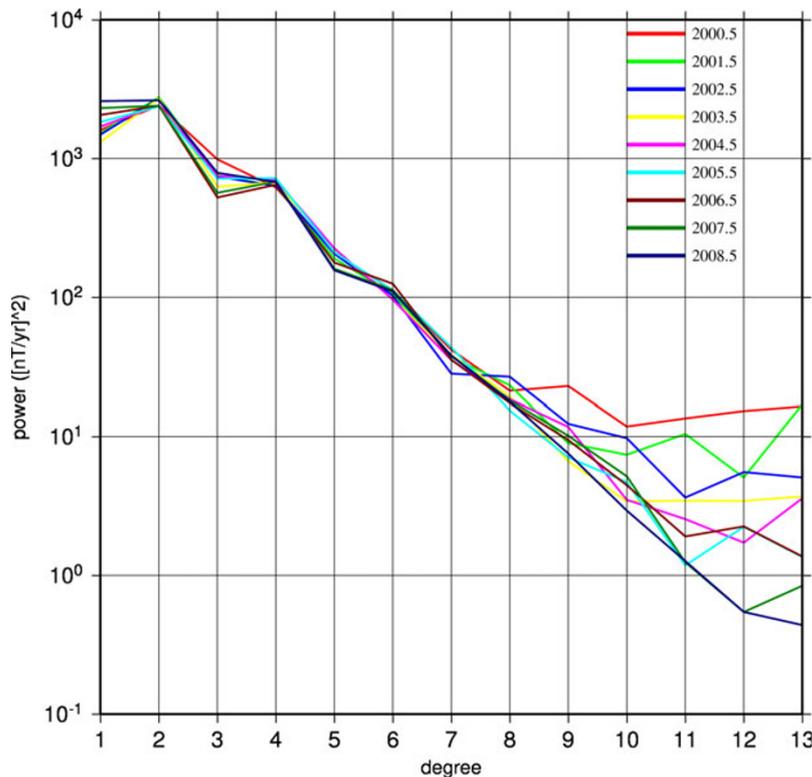


Fig. 3. Spectra of secular variation from parent model, extracted at yearly intervals.

data constitute a larger proportion of the total after Ørsted vector data becomes unavailable (see Fig. 1). In contrast, CHAMP and observatory scalar and vector data are available for all time but unlike satellite data, observatory scalar and vector data are confined to high and mid/low latitudes respectively. However, it is worth commenting on some patterns in the misfits. As expected, the mean and RMS satellite misfits are larger at high geomagnetic latitudes ( $>50$  or  $<-50$  degrees) than at mid and low latitudes ( $\geq -50$  and  $\leq 50$  degrees). Contamination from un-modelled auroral current systems will increase the high-latitude misfits but so too will the downweighting described in Section 2. At high latitudes the satellite  $Z$  misfits are significantly lower than those of the  $X$  or  $Y$  components due to the effects of field-aligned currents. Although observatory biases are co-estimated in the parent model this does not mean that we should expect zero mean observatory misfits, since an L1-norm is used when fitting the model.

The RMS misfits versus time for CHAMP, Ørsted and observatory data, taken over all magnetic components, latitudes, and times, are shown in Fig. 2. The CHAMP misfits do not show any long-term evolution but there are clearly periodic signals in the misfits that closely correlate with that satellite's 130 day local-time cycle. This is almost certainly due to increasing contamination from un-modelled sources as the satellite's orbit precesses from a day-night to a dawn-dusk configuration. A similar pattern is seen in the Ørsted residuals prior to 2006 that is close to that satellite's 800 day local-time cycle. After 2006, when only scalar data are available, the pattern becomes less clear. While the peaks in the misfits reach 10s of nT, it should be noted that they represent relatively few data as would be expected as the

satellites approach the limit of the data selection window. This is confirmed in Table 1, which show the RMS misfits of all CHAMP and Ørsted data are below 14 nT. The observatory misfits show greater consistency through time compared with the satellite misfits but there are still some peaks. The largest peak, in 2009, is the result of spurious data from a single observatory that was not filtered out in the data selection process. This is not expected to have any significant effect on the model but it is still undesirable and we will implement more stringent data checking in future models. In addition, there are smaller peaks spread throughout the time-series. These are not as regular as the satellite peaks and are very localised in time. The cause of these features is not yet clear.

## 5. Extraction of IGRF Candidate

The main field spectra derived at yearly intervals from our parent model exhibit a strict monotonic decline with increasing degree for all yearly samples. Relative changes between the samples over the eight year span are small and mostly at the higher degrees. The secular variation, extracted at the same dates and plotted in Fig. 3 show a fairly steady decline in power up to degree eight. Above this the spectra from most years begin to level out and even increase in power suggesting that noise is beginning to dominate in that part of the model. For degree eight and below, there is some year-to-year variation in the spectra but this is relatively low and shows no clear trend. For degrees nine and above there is a large variation in spectral power, which seems to decline with time. Given the evidence of noise at these degrees, this trend most likely represents a decline in this noise rather than an evolution in the secular

variation. Why this occurs is not obvious but may be related to the quieter geomagnetic conditions and overall increase in selected data in later years that can be seen in Fig. 1.

On the evidence presented above, the main field seems robust to degree and order 13 and we extract a set of MF coefficients at 2005.0 as a candidate model. To derive the 2010 main field candidate model, we project the behaviour of the parent model between the final two spline knots forward to 2010.0, up to degree and order 13. To do this we extract a degree and order 13 main field model at 2009.0 and use the SV estimate from the same period up to degree and order 13. We believe that using the full degree 13 SV estimate is justified given the apparent lower noise in the most recent SV behaviour seen in Fig. 3 and the relatively short time step. Given the longer time span of the average SV from 2010.0 to 2015.0, we did not rely on a single extraction of SV from the parent model. Instead we used an average of the parent model's SV up to degree and order 8 between 2005.0 to 2009.0. The period over which the SV average was calculated, 2005 to 2009, was chosen because it represented the most recent part of the parent model covering a time-range similar to the 5 year life-time of the SV prediction.

## 6. Summary and Conclusions

We have calculated MF and SV coefficients for use in the derivation of IGRF-11. These models are based on a large data set selected from both satellite and observatory data that have been carefully selected to remove unwanted sources. Unlike previous models, we include weighted vector satellite data at all latitudes by utilising two noise estimators. We believe our IGRF candidate models are a robust contribution to IGRF-11.

**Acknowledgments.** We would like to acknowledge the Ørsted and CHAMP science data centres, INTERMAGNET (at <http://www.intermagnet.org>), our colleagues in BGS who run the World Data Centre for Geomagnetism, and the many individual institutes operating observatories. We would also like to thank colleagues within the other IGRF-11 teams for their comments on our work and thank the referees for comments and suggestions that have improved the paper. This paper is published with the permission of the Executive Director, BGS (NERC) and the work was supported by NERC grant NER/O/S/2003/00677 (GEOSPACE).

## References

- Lesur, V., S. Macmillan, and A. Thomson, The BGS magnetic field candidate models for the 10th generation IGRF, *Earth Planets Space*, **57**, 1157–1163, 2005.
- Lesur, V., I. Wardinski, M. Rother, and M. Manda, GRIMM: the GFZ Reference Internal Magnetic Model based on vector satellite and observatory data, *Geophys. J. Int.*, **173**, 382–394, 2008.
- Macmillan, S. and S. Maus, International Geomagnetic Reference Field—the tenth generation, *Earth Planets Space*, **57**, 1135–1140, 2006.
- Maus, S., F. Yin, H. Lüher, C. Manoj, M. Rother, J. Rauberg, I. Michaelis, C. Stolle, and R. D. Müller, 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), 2008.
- Olsen, N. and M. Manda, Rapidly changing flows in the Earth's core, *Nature Geosci.*, **1**, 390–394, 2008.
- Thomson, A. W. P. and V. Lesur, An improved geomagnetic data selection algorithm for global geomagnetic field modelling, *Geophys. J. Int.*, **169**, 951–963, 2007.
- Thomson, A. W. P., B. Hamilton, S. Macmillan, and S. Reay, A novel weighting method for satellite magnetic data and a new global magnetic field model, *Geophys. J. Int.*, **181**, 250–260, 2010.

---

B. Hamilton (e-mail: [bham@bgs.ac.uk](mailto:bham@bgs.ac.uk)), S. Macmillan, and A. Thomson