Notes for evaluating BGS magnetic field candidate models for the 10th generation IGRF

Vincent Lesur (V.Lesur@bgs.ac.uk), Susan Macmillan (S.Macmillan@bgs.ac.uk) and Alan Thomson (A.Thomson@bgs.ac.uk)

British Geological Survey, Murchison House, West Mains Road, Edinburgh, EH9 3LA, UK.

Short Version

Data: quiet night-time Ørsted, Champ spanning 1999.2-2004.6 and observatory hourly means spanning 1999.0-2003.3. To improve the secular variation estimates for 2005.0-2010.0, predictions based on application of linear prediction filters to long series of observatory annual means were also used.

Parent model: internal degree 36, quadratic dependence on time up to degree8, linear dependence on time up to degree 12, external degree 2 with linear dependence on time, annual and semi-annual variations, Dst dependence for degree 1 terms. Additionally for the external field non-zonal degree 1 coefficients in the Geocentric Equatorial Inertial reference frame with annual variations and IMF B_y dependence.

Final models: Final models were based on extrapolations of truncated parent model to 2005.0 for the main field and to 2007.5 for the secular variation. This latter set of coefficients were then used to generate a synthetic dataset at Earth's surface and this was augmented with long term predictions at the observatories, to produce the final secular variation model at 2007.5.

Long Version

1. Introduction

The BGS candidates for the IGRF-10 were built in two stages. In the first stage a magnetic field model (the "parent model") was constructed from satellite and observatory data selected between 1999.0 and 2004.58. This model was then extrapolated forward in time to derive a main field model for 2005.0 and a secular variation model for 2007.5. Truncating this main field model provides the BGS main field candidate for IGRF-10. The secular variation component of the model was however only used to synthesize data on a grid on the Earth's surface and this synthetic data set was combined with secular variation data linearly predicted using long time series data from observatories. The BGS secular variation model candidate for 2005.0 to 2010.0 was then estimated by least squares fit of these two data sets.

The following sections describe the data selection, model parameterisation, data weighting and model estimation. The final section describes how the IGRF candidate models were extracted from the “parent model”.

2. Data selection

Satellite data

All available Ørsted and Champ scalar and vector data were utilized. Data were selected 20s apart, using $Kp \leq 1+$, |Bx|, |By| $\leq$ 10 nT and $0 \leq$ Bz ≤ 6 nT, a maximum solar wind speed of 450 km/s and, up to July 2004, *Dst* index values in the range -20nT to 0nT. No real-time *Dst* indices were used at any point in the model. Data with Kp > 1+ in the previous three hour interval before the acquisition time and with 0<*Dst* or *Dst* < -20 nT in the hour before the acquisition time were not used. Contamination by the magnetic field generated in the ionosphere was minimised by choosing only night-side data between 23:00 and 05:00 local time. Only vector data were selected in the geomagnetic latitude range -55° to 55°. Otherwise, at higher latitudes only scalar data were selected using the further constraint of maximum polar cap (PC) index $\leq$ 0.2 (chosen by examining *rms* fit to data as a function of maximum PC). After 2003.0 the maximum solar wind speed limit was allowed to increase to 550 km/s in order to select enough data as a result of the higher wind speed average for that year.

Ørsted and Champ data up to the end of September 2004 and June 2004 respectively were downloaded but the local-time selection criterion limited the span of Ørsted data at 2004.3. The span of Champ data was limited at 2004.35 by the attitude corrections available (Maus, personal communication). We note that the distributions in time of the Ørsted satellite data, and to a lesser extent the Champ data, have gaps in them due to the local-time selection criterion.

Observatory hourly mean data

Hourly mean vector data at 151 geomagnetic observatories were selected. Any measured discontinuities were applied to these data, and any other identifiable discontinuities resulted in the observatory series being split so that separate biases can be solved for in the model. The selection criteria were similar to those for satellite data: Kp $\leq$ 1+, $0 \leq$ Bz, *Dst* index values in the range -15nT to 0nT and local time between 23:00 and 05:00. The distribution in time of observatory data stops at 2004.0 but is more even than for satellite data.

In an attempt to minimise the noise level due to currents flowing in and above the polar ionosphere, vector data from observatories at geomagnetic latitudes outside the range [-55°; 55°] were projected onto the direction of a predefined magnetic field model. The advantage of using projected data over total intensity data is that a linear relationship is maintained between the observatory offsets (introduced to account for part of the local crustal field) and data.

Observatory secular variation data

Prediction of secular variation to 2010.0 was made using linear predictor filters applied to 159 series of first differences of observatory annual means in *X*, *Y* and *Z* (Macmillan and Quinn, 2000). The time intervals covered by observatory time series are long compared to the time interval of data used in the "parent model" with some data extending back into the 19th century. Linear prediction is successful at extrapolating signals that are smooth and oscillatory, though not necessarily periodic, and tests have shown that when predicting more than about 3 years ahead, this method is better that linear regression applied to recent first differences. The time series from 159 observatories were prepared using the file of annual means maintained by BGS, taking account of any jumps and gaps in the data, and discounting certain early parts of records where the noise levels are particularly high. The data for use in the final secular-variation model were averages of the predictions for 2005.0-2010.0 and were assigned uncertainties that reflected the past success of prediction for the data series in question. Data from 29 observatories which had time series too short for the application of linear prediction filters were also used by computing average secular-variation estimates and assuming that these did not change with time.

3. Model parameterisation

Away from its sources, the magnetic field $\boldsymbol{B}$ is a potential field and therefore can be written as the negative gradient of a potential $\boldsymbol{B}(\theta,\varphi,r,t) = -\nabla V(\theta,\varphi,r,t)$. This potential can be expanded in terms of spherical harmonics:

$$\begin{aligned} V(\theta,\varphi,r,t) = a\left\{\sum_{l=1}^{li}\sum_{m=0}^{l}\left(g_l^m(t)\cos(m\varphi)+h_l^m(t)\sin(m\varphi)\right)\left(\frac{a}{r}\right)^{l+1}\mathrm{P}_l^m(\cos\theta)\right\} \\ + a\left\{\sum_{l=1}^{le}\sum_{m=0}^{l}\left(q_l^m(t)\cos(m\varphi)+s_l^m(t)\sin(m\varphi)\right)\left(\frac{r}{a}\right)^{l}\mathrm{P}_l^m(\cos\theta)\right\} \end{aligned} \tag{1}$$

where a (6371.2 km) is the Earth's reference radius, (θ,φ,r) are spherical coordinates in a geocentric reference frame (GEO), $\mathrm{P}_l^m(\cos\theta)$ are the Schmidt semi-normalized Legendre functions, and $(g_l^m(t), h_l^m(t))$ and $(q_l^m(t), s_l^m(t))$ are the time-dependent Gauss coefficients describing internal and external sources respectively. We have used $li = 36$ and $le = 2$ as the truncation level of the internal and external fields respectively. This maximum internal degree is high and we believe that out estimates of the Gauss coefficients are robust at least up to degree 20.

The internal Gauss coefficients from degree 1 to 8 are assumed to have a quadratic dependence on time:

$$\begin{aligned} g_l^m(t) &= g_l^m + \dot{g}_l^m(t-t_0) + \ddot{g}_l^m(t-t_0)^2 \\ h_l^m(t) &= h_l^m + \dot{h}_l^m(t-t_0) + \ddot{h}_l^m(t-t_0)^2 \end{aligned} \tag{2}$$

where the time is given in decimal year and t_0 is the reference date of the model (2002.0). From degree 9 up to 12 a linear dependence on time of the internal Gauss

coefficients is assumed, and, for higher degrees, the internal Gauss coefficients are constant with time. The external Gauss coefficients have a linear dependence on time.

Other authors (Sabaka et al., 2002, Olsen 2002) have shown that both external and internal degree 1 and 2 Gauss coefficients have a seasonal variation. The same seasonal parameterisation is used here. Let $\tilde{g}_1^m(t)$ be the parts of the $g_1^m(t)$ coefficients that account for these seasonal variations, then:

$$\tilde{g}_1^m(t) = \tilde{g}_{1,1c}^m \cos(2\pi(t-t_0)) + \tilde{g}_{1,1s}^m \sin(2\pi(t-t_0)) + \tilde{g}_{1,2c}^m \cos(4\pi(t-t_0)) + \tilde{g}_{1,2s}^m \sin(4\pi(t-t_0)). \quad (3)$$

Similar representations are used for $h_1^m(t)$, $q_1^m(t)$, $s_1^m(t)$ and for degree 2 internal and external Gauss coefficients.

A *Dst*dependence for the degree 1 Gauss coefficients is introduced to represent the variability of the magnetospheric ring current. We used the internal/external *Dst* separation due to Maus (2004), when available, otherwise a fixed ratio of 0.27 (Langel and Estes, 1985) between the external *Dst* dependence and the associated internal induced contribution was assumed.

With the above parameterisation, large-scale external fields perpendicular to the Earth rotation axis cannot be modelled as potential fields due to the rotation of the geocentric reference frame (GEO) relative to the Sun-Earth axis, and also due to the processing technique applied to the observatory data to synthesize the *Dst* index. Evidence of the contribution of these fields to the near-Earth geomagnetic field has been shown in Lesur et al. (2004), where a magnetic signal that correlates well with the Y component of Interplanetary Magnetic Field (B_y) is presented. These contributions can however be modelled in a reference frame that is not rotating with a 24-hour periodicity relative to the Sun-Earth axis. The Geocentric Equatorial Inertial reference frame (GEI) is one such coordinate system. This system is (to the first order) fixed with respect to the distant stars, its Z-axis is along the Earth rotation axis and its X-axis is pointing towards the first point of Aries (Hapgood, 1992). In GEI the Sun-Earth axis rotates with a yearly periodicity. In this system of coordinates we define a potential field for the perpendicular component:

$$V^{gei}(\theta,\varphi,r,t) = r\left(q_1^{1gei}(t)\cos(\varphi) + s_1^{1gei}(t)\sin(\varphi)\right)\mathrm{P}_1^1(\cos\theta)$$

where (θ,φ,r) are the spherical coordinates in GEI. We define the time behaviour of $q_1^{1gei}(t)$ as:

$$q_1^{1gei}(t) = (q_{1,1c}^{1c} + q_{1,1c}^{1i}\mathrm{B_y})\cos(2\pi(t-t_0)) + (q_{1,1s}^{1c} + q_{1,1s}^{1i}\mathrm{B_y})\sin(2\pi(t-t_0))$$

and similarly for $s_1^{1gei}(t)$.

Finally, the data set contains observatory data and we have to introduce offsets at each observatory to take into account the field, mainly generated in the crust, which cannot be described by our model. At an observatory, the magnetic field $\boldsymbol{B}$ is:

$$\boldsymbol{B}(\theta,\varphi,r,t) = -\nabla V(\theta,\varphi,r,t) + \boldsymbol{O}(\theta,\varphi,r) \tag{4}$$

where the offset vector $\boldsymbol{O}(\theta,\varphi,r)$ is constant in time. Therefore, there are three new parameters per observatory in the geomagnetic latitude range -55° to 55°. Outside this range, the data we use are projections of the measured magnetic field onto *a priori* directions. In these cases, only one parameter is introduced per observatory to account for the crustal offset into this same *a priori* direction.

4. Data weighting, attitude errors and covariance matrix

The standard deviations (SDs) associated with both satellite and observatory data were defined as:

$$\sigma = \sigma_0 + d_z(1+\cos(za)) \tag{5}$$

where za is the zenith angle of the sun, the factor $d_z = 3$ nT for satellite scalar data and 20 nT for observatory projected data. Otherwise d_z is set to 0 nT and σ_0 to 2 nT for all vector data. The dependence relative to the zenith angle was introduced to account for the noise level due to the increased conductivity of the sun-lit ionosphere.

As in Holme (2000) and Olsen (2002), we account for the anisotropy in the attitude accuracy of Ørsted vector data by estimating off-diagonal elements of the data covariance matrix.

No further selection or decimation of the data was introduced to deal with the high data density at high latitudes. Instead, the data values were multiplied by weights. The spherical surface was divided in roughly equal-area cells whose size at the equator was 5° in latitude and longitude. For a data point in a given cell, a weight was calculated as the ratio of the average number of data in non-empty cells to the number of data in that cell. These weights were computed independently for observatory and satellite data. To introduce these weights into our parameter estimation scheme, their inverses were put in diagonal weight matrices that left and right multiplied the covariance matrix.

Let $\boldsymbol{C}_d$ be the data covariance matrix multiplied by the weight matrices above. This matrix is real and symmetric and therefore it can be reduced to a real diagonal matrix by orthogonal rotation:

$$\boldsymbol{C}_d = \boldsymbol{A}^t(\boldsymbol{S}^t\boldsymbol{S})^{-1}\boldsymbol{A} \tag{6}$$

where the superscript ^{t} denotes the transpose and the superscript $^{-1}$ the inverse, $\boldsymbol{S}$ is a real diagonal matrix whose elements are the inverse of the weighted SDs and $\boldsymbol{A}$ is the inverse of the rotation matrix. (The notation used is consistent with that of Olsen (2002)).

5. Model estimation

We estimated the model parameters by fitting the data using a classic least-squares approach. For this application the iterative process can be written as:

$$\begin{aligned} \boldsymbol{d}_i &= \mathrm{G}(\boldsymbol{p}_i) \\ \boldsymbol{\delta d}_i &= \boldsymbol{d}_{obs} - \boldsymbol{d}_i \\ \boldsymbol{\delta p}_i &= \left[\boldsymbol{G}_i^t\, \boldsymbol{A}^t\, \boldsymbol{S}^t\, \boldsymbol{S}\, \boldsymbol{A}\, \boldsymbol{G}_i\right]^{-1} \boldsymbol{G}_i^t\, \boldsymbol{A}^t\, \boldsymbol{S}^t\, \boldsymbol{S}\, \boldsymbol{A}\, \boldsymbol{\delta d}_i \\ \boldsymbol{p}_{i+1} &= \boldsymbol{p}_i + \boldsymbol{\delta p}_i \end{aligned} \tag{7}$$

where the superscript ^{i} denotes the ith iterate, $\boldsymbol{p}$ is the model vector (i.e. the model parameters), $\mathrm{G}(\boldsymbol{p})$ is the forward non-linear function used to calculate the predicted values $\boldsymbol{d}$ from a model $\boldsymbol{p}$, $\boldsymbol{d}_{obs}$ is the data vector and $\boldsymbol{G}$ is the $n \times p$ matrix associated with the equations of condition (n: number of data values; p number of model parameters):

$$\boldsymbol{G} = \left[\frac{\partial \mathrm{G}_{\mathrm{i}}(\boldsymbol{p})}{\partial \boldsymbol{p}_j}\right]_{\substack{i=1,n \\ j=1,p}} \tag{8}$$

The linear system is solved using eigenvalue/eigenvector decomposition. The problem can be regularised by removing the very small eigenvalues and their associated eigenvectors, but that proved unnecessary in this case.

The number of data, mean and the *rms* misfits to the data for the resulting model are given in Table 1.

Table 1: Mean and rms misfit to the data set for the parent model (nT)

Component	N	mean	rms
X satellite	60,931	0.49	6.45
Y satellite	60,931	0.27	5.80
Z satellite	60,931	-0.27	5.00
F Satellite	14,004	-2.25	6.75
X observatory	75,652	0.00	5.39
Y observatory	75,652	0.00	4.65
Z observatory	75,652	0.00	3.54
F observatory	35,653	0.57	18.71

where the *X*, *Y* and *Z* components are oriented North, East and down respectively and N is the number of data values.

6. Extrapolation to 2005.0 and 2007.5

The model resulting from the above process is valid only from 1999.0 to 2004.35 and therefore has to be extrapolated forward in time. The coefficients for the main field model at 2005.0 and secular variation model at 2007.5 were calculated using Equations 2. However for this extrapolation, the $\dot{g}_l^m$ and $\dot{h}_l^m$ coefficients for degree 11 and above as well as the $\ddot{g}_l^m$ and $\ddot{h}_l^m$ for degree 7 and above, were set to zero.

The resulting main field model was then truncated to degree 13 to produce the main field candidate for the IGRF-10. The secular variation model was used to produce a synthetic data set on a 6371.2 km radius sphere, composed of vector data located at the centre of equal-area cells whose size at the equator was 5° in latitude and longitude. This data set, comprising 1654 vector values, was then augmented by the secular variation predictions from 188 observatories. Each data was weighted by the inverse of its estimated uncertainty. The synthetic data were down-weighted because the formal standard deviations of the underlying model Gauss coefficients are known to be too small (Lowes and Olsen, 2004) but they were on average given higher weights in the final model than the observatory secular variation values. A spherical harmonic model of the secular variation with maximum degree 10 was then fitted to these data and truncated to degree 8 to produce the BGS secular variation candidate model for the IGRF-10.

Acknowledgements

The staff of magnetic observatories around the world, the Ørsted and Champ data centres and the World Data Centres in Copenhagen, Kyoto and Edinburgh are thanked for providing the data on which this study depends. The IMF data are from the Advance Composition Explorer (ACE) spacecraft and made available through OMNIWeb (courtesy NASA).

References

Hapgood, M.A., 1992, Space physics coordinate transformation: A user guide, *Planetary and Space Science*, **40**, 711-717.

Holme, R., 2000. Modelling of attitude error in vector magnetic data: application to Ørsted data, *Earth Planets and Space*, **52**, 1187-1197.

Langel, R.A. and Estes, R.H., 1985. Large-scale, Near-Earth magnetic fields from external sources and the corresponding induced internal field, *J. Geophys. Res.*, **90**, 2487-2494.

Lesur, V., Macmillan, S. and Thomson, A., 2004, Magnetic field model with daily variations of the magnetospheric field and its induced counterpart in 2001. *Geophys. J. Int*., accepted for publication.

Lowes, F. J. and Olsen, N., 2004. A more realistic estimate of the variances and systematic errors in spherical harmonic geomagnetic field models. *Geophys. J. Int.*, **157**, 1027-1044.

Macmillan, S. and Quinn, J. M., 2000. The 2000 revision of the joint UK/US geomagnetic field models and an IGRF2000 candidate model. *Earth, Planets and Space*, **52**, 1149-1162.

Maus, S. and Weidelt, P., 2004, Separating magnetospheric disturbance magnetic field into external and transient internal contributions using 1D conductivity model of the Earth. *GRL*, **31**, n° 12, L12614, 10.1029/2004GL020232.

Olsen, N., 2002. A model of the geomagnetic field and its secular variation for epoch 2000 estimated from Ørsted data, *Geophys. J. Int*., **149**, 454-462.

Sabaka, T.J., Olsen, N. and Langel, R.A., 2002. A comprehensive model of the quiet-time, near-Earth magnetic field: phase 3, *Geophys. J. Int*., **151**, 32-68.