Evaluation of IGRF-11 candidate models

Chris Finlay, Institüt für Geophysik, Dept. Erdwissenshaften, ETH Zürich.

15/11/2009

Summary

Tests have been carried out to compare submitted candidates for the various components of IGRF-11: main field models DGRF-2005 and IGRF-2010 (to spherical harmonic degree 13) as well as predictive secular variation models SV-2010-2015 (to degree 8). The tests performed involve statistical comparisons between the models focusing on differences between the candidates themselves and between candidates and mean models in both the spectral domain and in physical space. Regarding DGRF-2005 candidates it is found that candidate D is worryingly different to the other candidates, particularly at degrees 10 and above; candidates C2, E2 and F also show appreciable differences at high degrees apparently stemming from difficulties in the polar regions. In contrast candidates A, B and G are found to agree rather well. It is therefore recommended that candidate D be rejected while C2, E2 and F be down-weighted in the determination of DGRF-2005. Turning to the IGRF models for epoch 2010 it was found that candidates D and E are most different to the other models, while candidates G and B are very similar and closest to the mean model, despite the differences in their data selection and modelling approaches. It is thus suggested that D and E be down-weighted relative to C2, A and F while G and B should receive the highest weighting in the determination of IGRF-2010. Regarding the average predictive SV for epoch 2010-2015 it was found that models E and G were rather anomalous with large differences from other models in certain spherical harmonic coefficients. It is suggested that candidates E and G be down-weighted in the determination of SV-2010-2015 while the other SV candidates are equally weighted because it is difficult to rigorously distinguish which models should be preferred.

1 Introduction

Seven candidate main field candidate models were submitted to the 11th generation IGRF for the epochs 2005 (retrospective) and 2010 (predictive) while eight candidates were submitted proposing average secular variation (rate of change of field with time) estimates for the interval 2010-2105. These candidate models consist of spherical harmonic coefficients to degree 13 (main field) and degree 8 (secular variation). Tables 1, 2 and 3 below list the teams that submitted candidates, document the names used to denote the various models in this evaluation, collect the major data sources used in the construction of the candidates and contain very brief comments concerning the various modelling approaches adopted.

The purpose of this document is to report quantitative comparisons between the candidate models and mean models, presenting the results in an accessible graphical manner when possible. Such information is needed by the IGRF-11 task force for them to make informed decisions concerning weighting of candidate models during the compilation of the final IGRF-11 models and also to help identify whether any of the candidates possess undesirable characteristics. Model evaluations would ideally be based upon not only statistical analysis but also on comparisons with independent data that accurately measure the field of interest (the internal magnetic field at Earth's surface) at the epochs of interest. Unfortunately such 'ideal evaluation data' do not exist for the future epochs of 2010 and 2010-2015 and it is even troublesome to obtain high quality independent data for epoch 2005. Attempts to assess the candidate models using either observatory or satellite data are thus hampered by the necessity of propagating the models to suitable comparison epochs as well as the difficulties of separating internal and external field contributions. Other workers have attempted such comparisons (see the evaluations of Thébault, Chulliat and Olsen). However in this document attention will instead focus on comparisons between the models and with mean models based on the computation relevant diagnostic statistics (in both the spectral and physical domains).

DGRF candidate models for main field epoch 2005				
Team	Model	Organization	Data	Comments (parent model etc.)
A	DGRF-2005-A	DTU Space / IPGP/NASA GSFC	Oersted, CHAMP, SAC-C revised obs mon. means	Based on CHAOS-3 in 2005.0 (6th order splines)
B	DGRF-2005-B	NOAA/NGDC & GFZ	CHAMP 2003.5-2006.5	2nd order Taylor series
C	DGRF-2005-C2	BGS	Oersted, CHAMP and obs. hr. means 1999.0-2009.5	revised sub: parent model linear splines (400day knots spacing)
D	DGRF-2005-D	IZMIRAN	CHAMP 2004.0-2006.0	NOC method
E	DGRF-2005-E2	EOST/LPGN/ LATMOS/IPGP	CHAMP & Oersted 2004.5-2005.5	revised sub: based on 12 month model with linear SV
F	DGRF-2005-F	IPGP/EOST/ /LPGN/LATMOS	CHAMP 2004.35-2005.66	2nd order Taylor series
G	DGRF-2005-G	GFZ/IPGP	CHAMP 2001-2009.6 obs. hourly means	Based on GRIMM2x (6th order splines) av. over 1 yr.

Table 1: Summary of DGRF-2005 candidate models submitted for consideration in IGRF 11th generation.

IGRF candidate models for main field epoch 2010				
Team	Model	Organization	Data	Comments (parent model etc.)
A	IGRF-2010-A	DTU Space / IPGP/NASA GSFC	Oersted, CHAMP, SAC-C revised obs monthly means	Based on CHAOS-3 evaluated in 2010.0
B	IGRF-2010-B	NOAA/NGDC & GFZ	CHAMP 2006.5-2009.67	2nd order Taylor series: SV & SA used for 2010.0 estimate
C	IGRF-2010-C2	BGS	Oersted, CHAMP obs. hr. means 1999.0-2009.5	Revised sub: model evaluated 2009.0 MF and SV used to predict 2010 field.
D	IGRF-2010-D	IZMIRAN	CHAMP 2004.0-2009.25	NOC method with extrapol. to 2010 using NOC1,2
E	IGRF-2010-E	EOST/LPGN/ LATMOS/IPGP	CHAMP June/July 2009	Model at 2009.485 extrapol. to 2010 using SV models for 2009, 2010.
F	IGRF-2010-F	IPGP/EOST/ /LPGN/LATMOS	CHAMP 2008.92-2009.61	2nd order Taylor series: SV model used to extrapolate to 2010
G	IGRF-2010-G	GFZ/IPGP	CHAMP 2001-2009.6 obs. hourly means	Based on GRIMM2x MF and SV in 2009 extrapol. to 2010

Table 2: Summary of IGRF-2010 candidate models submitted for consideration in IGRF 11th generation.

The formulae defining quantities of interest are first presented here to avoid ambiguity. The IGRF-11 candidates take the form of Schmidt semi-normalized spherical harmonic models. In what follows g_l^m and h_l^m are used to denote the spherical harmonic (Gauss) coefficients (units nT or nT/yr for SV) associated with the $\cos m\phi$ and $\sin m\phi$ components of candidate models respectively. As is conventional n denotes spherical harmonic degree while m denotes spherical harmonic order. Often we will be concerned with differences between a candidate model i and either some reference mean model or another candidate (labelled j) whose coefficients will be denoted by G_l^m and H_l^m. Much use will be made below of the mean square difference between the vector fields per spherical harmonic degree $P_n^{i,j}$ (see, for example, Lowes (1966) and Lowes (1974))

$$P_n^{i,j} = (n+1)\Big(\frac{a}{r}\Big)^{(2n+4)} \sum_{m=0}^{n} \left[(g_n^m - G_n^m)^2 + (h_n^m - H_n^m)^2\right] \tag{1}$$

where a is Earth's reference spherical radius 6371.2km and r is some radius of interest, for example Earth's surface

SV candidate models for average secular variation during epoch 2010-2015				
Team	Model	Organization	Data	Comments (parent model etc.)
A	SV-2010-2015-A	DTU Space / IPGP/NASA GSFC	Oersted, CHAMP, SAC-C revised obs monthly means	Based on CHAOS-3 SV in 2010.0
B	SV-2010-2105-B	NOAA/NGDC	CHAMP 2006.5-2009.67	2nd order Taylor series: SV in 2009.67 used.
C	SV-2010-2015-C2	BGS	Oersted, CHAMP and obs. hourly means	Revised sub: Av. SV 2005-2009 from parent model used)
D	SV-2010-2105-D	IZMIRAN	CHAMP 2004.0-2009.25	Based on linear NOC extrapolated and compared to obs.
E	SV-2010-2015-E2	EOST/LPGN/ LATMOS/IPGP	Obs. hourly mean used to derive mon. and ann. means	Extrapol. provides 1st diff an. means 1981-2015: SV models by least sq.
F	SV-2010-2015-F	IPGP/EOST/ /LPGN/LATMOS	CHAMP 2008.92-2009.61	2nd order Taylor series: SV from 2008.2 used.
G	SV-2010-2015-G	GFZ/IPGP	CHAMP 2001-2009.6 obs. hourly means	Based on GRIMM2x: linear interpol. SV 2001-2009.5, extrapol to 2012.5.
H	SV-2010-2015-H	GSFC NASA		geodynamo sim: assim. from CALS7K.2, gufm1, CM4, CHAOS-2s

Table 3: Summary of SV-2010-2015 candidate models submitted for consideration in IGRF 11th generation.

or the core-mantle boundary (r=3481km). Summing over degrees n from 1 to the truncation degree N and taking the square root yields the RMS globally averaged difference between the models i and j

$$R_{i,j} = \sqrt{\sum_{n=1}^{N} P_n^{i,j}} \tag{2}$$

It is often very informative to calculate $R_{i,j}$ when j is a model consisting of the arithmetic mean of the candidate models (or some subset of these). Such arithmetic models models will be referred to below as 'mean_m'.

In addition to calculating the $R_{i,j}$ for individual models, it is also possible (see for example the evaluations of C. Beggan) to compute the mean value of the RMS differences $R_{i,j}$ for Q candidate models such that

$$\bar{R} = \frac{1}{(Q-1)} \sum_{candidates\ i \neq j} R_{i,j} \tag{3}$$

Such means of RMS differences will be referred to below in text and figures using the terminology 'mean_dif'.

Analysis of spherical harmonic power spectra is very useful in diagnosing differences in amplitude between models but tells us little about differences in phase. Phase variations between models can be studied as a function of spherical harmonic degree using the quantity known as the degree correlation $\rho_n^{i,j}$ (see, for example p.81 of Langel and Hinze (1998))

$$\rho_n^{i,j} = \frac{\sum_{m=0}^{n} (g_n^m G_n^m + h_n^m H_n^m)}{\sqrt{\left(\sum_{m=0}^{n} [(g_n^m)^2 + (h_n^m)^2]\right)\left(\sum_{m=0}^{n} [(G_n^m)^2 + (H_n^m)^2]\right)}} \tag{4}$$

Rather than averaging over spherical harmonic degree is also possible to study the differences between models coefficient by coefficient. It is then useful to normalize by the power per degree of the reference model and to study the percentage sensitivity $S(n,m)$ (see, for example, Sabaka and Olsen (2006))

$$S(n,m) = 100 \frac{g_n^m - G_n^m}{\sqrt{\frac{1}{2n+1} \sum_{m=0}^{n} [(G_n^m)^2 + (H_n^m)^2]}} \tag{5}$$

where of course the g_n^m and G_n^m in the numerator should be exchanged for h_n^m and H_n^m if the coefficient under study is associated with a $\sin m\phi$ rather than a $\cos m\phi$ term.

Below the DGRF candidate models for epoch 2005, the IGRF candidates for epoch 2010 and the predictive SV candidates for epoch 2010-2015 are each analyzed, with global comparisons presented first, then comparisons in the spectral domain before comparisons in the physical domain. Each group of models will be discussed in turn and recommendations made for relative weightings that might prove useful when deriving the final IGRF models.

2 DGRF-2005

2.1 Root Mean Square (RMS) vector field differences between candidates and mean models

Figure 1 presents the the RMS global vector field differences $R_{i,j}$ between the DGRF candidate models, between an arithmetic mean model 'mean_m' and between an arithmetic mean model excluding candidate D referred to as 'mean_m_no_D'. Blue colours represent close agreement between the models ($R_{i,j} < 6$nT) while red colors represent larger disagreements ($R_{i,j} > 10$nT). Note the symmetry about the diagonal entries (inside the black dividing lines) which is included as a check on the calculation. Figure 1 demonstrates that model D consistently disagrees most with all the other candidates and with the mean model. Removing candidate D from the calculation of the mean also significantly improves the agreement between the mean and the other candidates. An alternative perspective is given by the penultimate row (labelled 'mean_diff') where the arithmetic means $\bar{R}$ of the $R_{i,j}$ (see equation 3) for each candidate model are presented and by final row labelled 'mean_diff_noD' where the same calculation excluding model D is presented. Besides candidate D, candidates C2 and E show the next largest $R_{i,j}$. For completeness the precise numbers used to construct Figure 1 are reproduced below in Table 4.

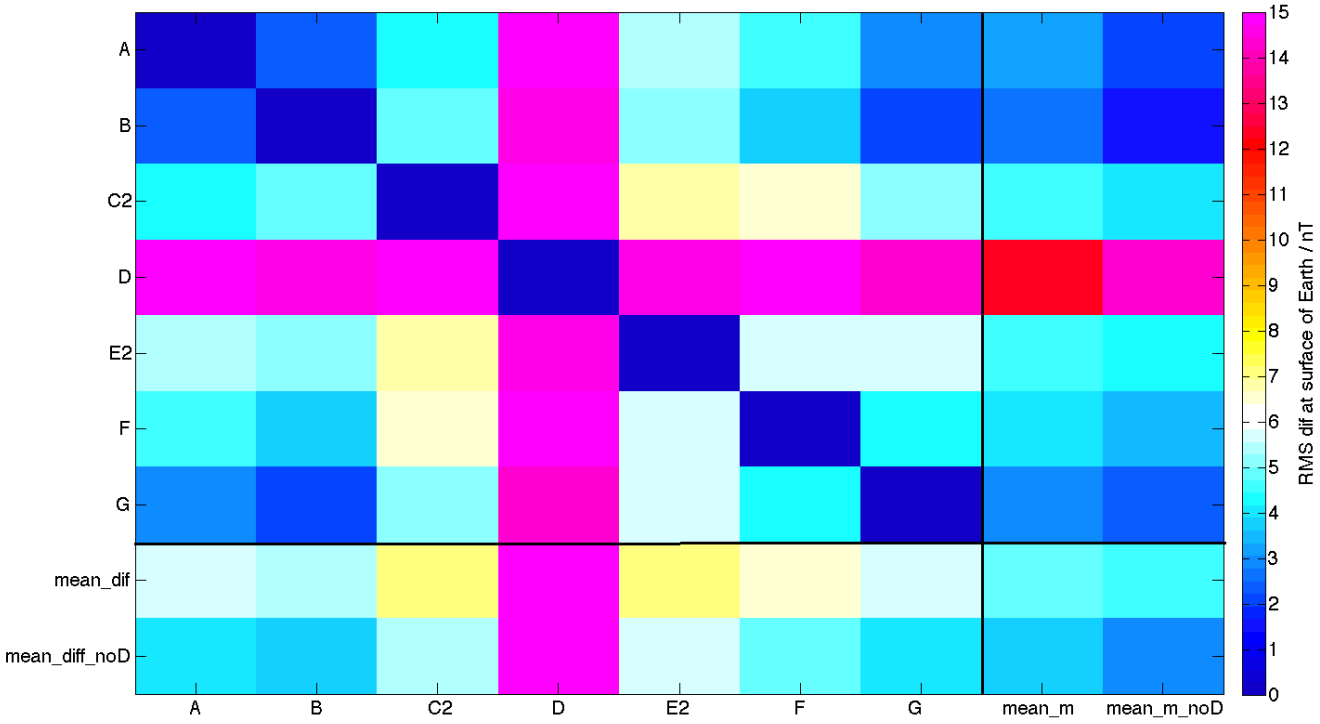

Figure 1: Plot showing RMS vector differences $R_{i,j}$ (see equation 2) in nT between individual DGRF-2005 candidates A, B, C2, D, E2, F, G, the arithmetic mean of all candidates 'mean_m' (in the second column from the right side) and the arithmetic mean model excluding model D 'mean_m_no_D' (rightmost column). The second from bottom row displays the mean of the RMS differences between each model and all other candidate models $\bar{R}$ (see eqn 3) referred to as 'mean_diff' while the bottom row displays the same but excluding model D (referred to as 'mean_diff_noD').

	A	B	C2	D	E2	F	G	mean_m	mean_m_noD
A	0.00	2.29	4.33	14.89	5.44	4.60	2.92	3.13	2.04
B	2.29	0.00	4.82	14.47	5.25	3.77	2.20	2.62	1.67
C2	4.33	4.82	0.00	15.15	6.75	6.50	5.22	4.63	4.05
D	14.89	14.47	15.15	0.00	14.60	14.98	14.36	12.35	14.41
E2	5.44	5.25	6.75	14.60	0.00	5.58	5.65	4.52	4.23
F	4.60	3.77	6.50	14.98	5.58	0.00	4.35	4.06	3.40
G	2.92	2.20	5.22	14.36	5.65	4.35	0.00	2.96	2.35
mean_diff	5.74	5.46	7.13	14.74	7.21	6.62	5.78	4.90	4.59
mean_diff_noD	3.91	3.66	5.52	17.69	5.73	4.96	4.07	3.65	2.96

Table 4: Numerical values of RMS vector field differences $R_{i,j}$ between candidate models and mean models from Figure 1.

2.2 Comparisons in spectral space

Figure 2 presents power spectra of the DGRF candidate models as a function of spherical harmonic degree at Earth's surface and at the core-mantle boundary. The relevant formula is that given in equation 1 but with the reference model G_n^m=H_n^m=0. The spectra of the candidate models are all very similar, almost completely overlapping in these plots. The most noticeable differences occur for candidate D at degree 11 (lower power than the other candidates) and for candidate E2 at degree 13 (higher power than other candidates).

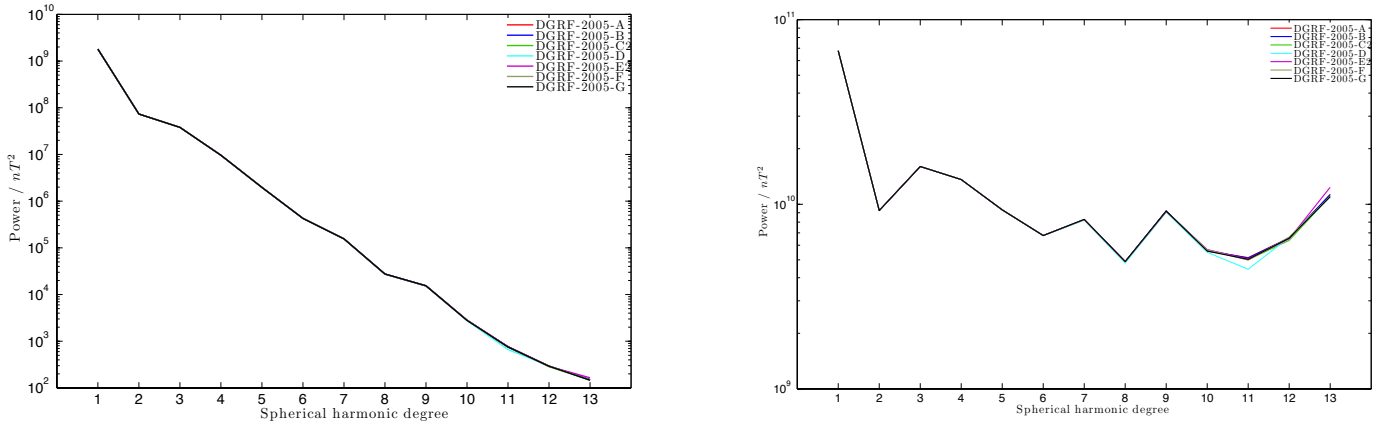

Figure 2: Power spectra of DGRF-2005 candidate models at Earth's surface (left) and at radius 3481km (core-mantle boundary) (right).

Figure 3 presents the degree correlation $\rho_n^{i,j}$ (see equation 4) between the DGRF candidate models and the model 'mean_m' (left plot) and 'mean_m_noD' (right plot). In both plots candidate D is found to be clearly anomalous above degree 9. Candidates C2, E and F depart rather more from both mean models than A, B and G which are very similar and close to the mean models.

Figure 4 presents the power spectra at Earth's surface of the differences (see equation 1) between the candidate models and the arithmetic mean model 'mean_m' (left) and with arithmetic mean model excluding candidate D 'mean_m_noD' (right). Again candidate D lies furthest from both mean models, with models C2 and E next furthest away.

Figure 5 presents the sensitivity matrix S(n,m) (see equation 5) of the candidate models compared to model 'mean_m_noD'. Candidate D possesses major coefficient by coefficient differences to 'mean_m_noD' , particularly above degree 9 and especially in certain coefficients for example, $g_{13}^2, g_{13}^4, g_{13}^{12}, h_{13}^2, h_{12}^4, h_{13}^5$ seem to be particularly different. A, B and G again show only minor sensitivity compared to the 'mean_m_noD' while C2, E2, F generally show slightly larger sensitivities especially at high degrees. The sensitivities are found to be small for all candidates at low degrees.

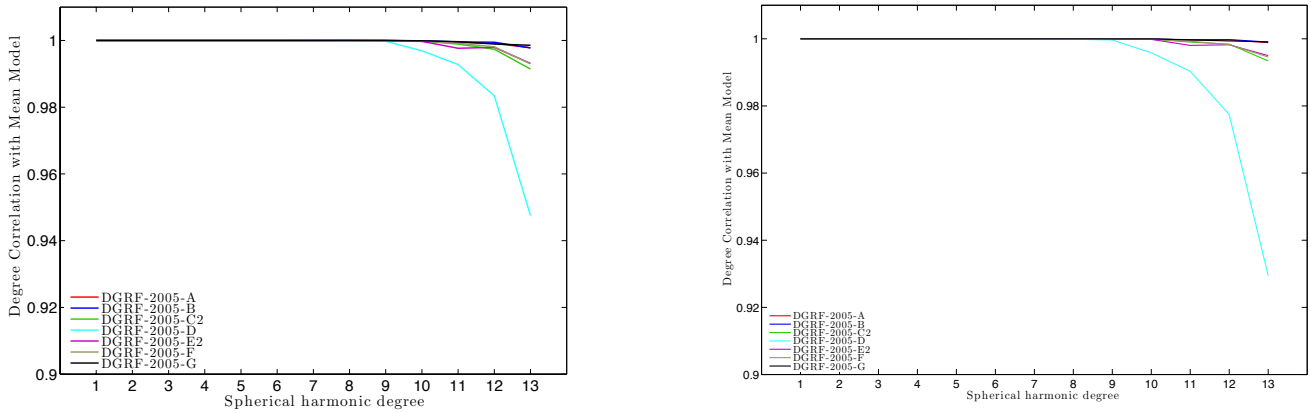

Figure 3: Degree correlation $\rho_n^{i,j}$ (see equation 4) of DGRF-2005 candidate models with the arithmetic mean model 'mean_m' (left) and with arithmetic mean model excluding candidate D 'mean_m_noD' (right).

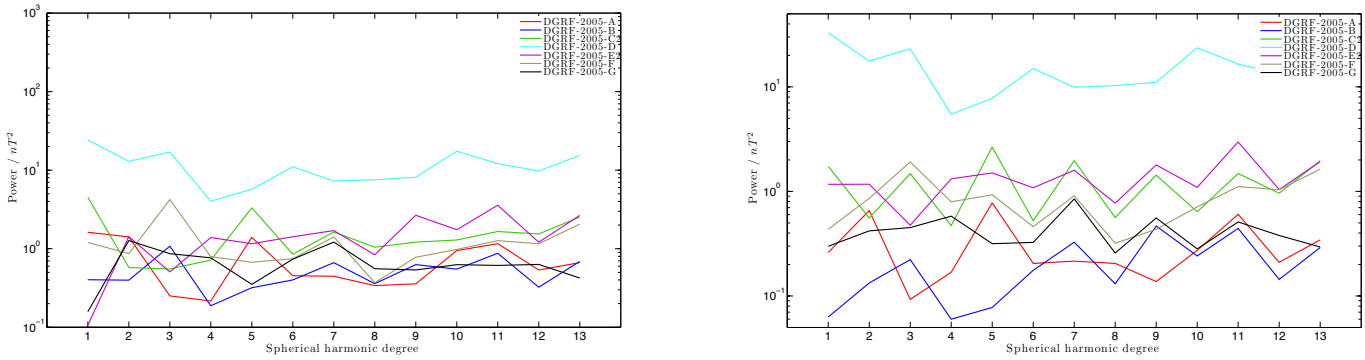

Figure 4: Power spectra of difference between DGRF-2005 candidate models and the arithmetic mean model 'mean_m' (left) and the arithmetic mean model excluding model D 'mean_m_noD' (right).

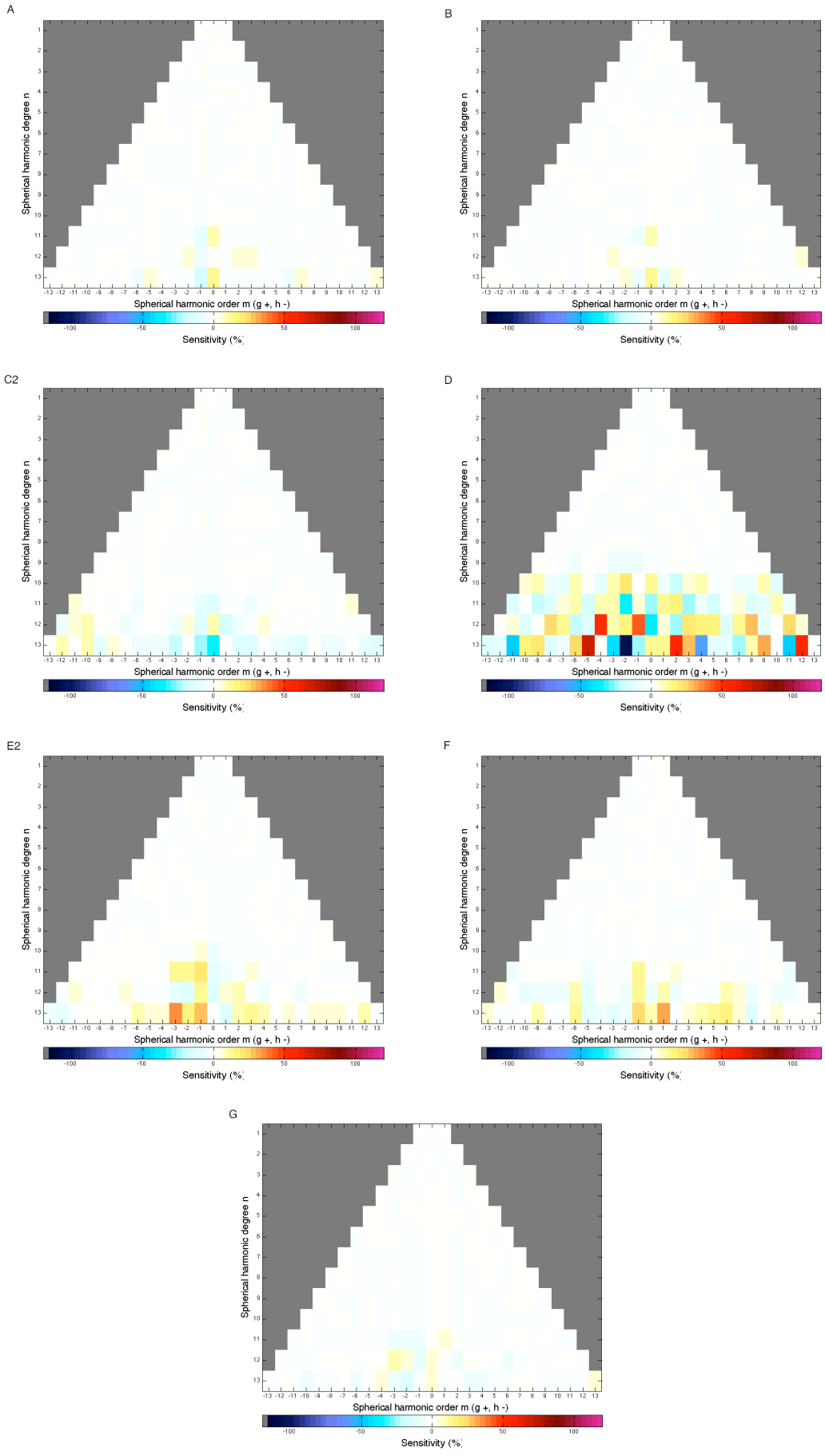

Figure 5: Matrices of sensitivity $S(n, m)$ (see equation 5) for DGRF-2005 models A, B, C2 ,D ,E ,F ,G compared to the arithmetic mean with model D removed ('mean_m_noD').

2.3 Physical space comparisons

Studying the candidate models in physical space yields insight into the geographical locations where differences between the candidate models are localized. In Figure 6 differences in the radial magnetic field at the core-mantle boundary between each candidate model and the model 'mean_m_noD' are presented in Mollweide projection. The comparisons are carried out here at the core-mantle boundary rather than at the Earth's surface because plotting there accentuates the higher degrees which earlier tests have shown is where the differences between candidates are strongest. In addition it is expected that the majority of the surface field modelled in the IGRF candidates will have its physical origin in the core. Visual inspection of Figure 6 reveals once again that candidate D is strikingly different compared to the other DGRF candidates. The differences are seen to be global and are not confined to any particular geographical region. Again models A, B and G are found to have the smallest differences to model 'mean_m_noD'. Model C2 shows differences predominantly in the polar regions and close to a meridian passing through North and South America. Model E2 displays a similar pattern of differences to C2, but with additional large scale differences south of Africa. The differences between candidate F and 'mean_m_noD' are less prominent and confined to the polar regions.

2.4 Discussion

Based on all the tests presented above, candidate D appears consistently different in both the spectral domain (with certain spherical harmonic coefficients apparently anomalous- see Figure 5) as well as in physical space where global problems are observed. The task force should therefore consider excluding candidate D from the calculation of the final DGRF model. Candidate models C2, E2 and F are also noticeably different from 'mean_m_noD' especially at high degree, while candidates A, B and G are rather similar according to all the tests performed here.

Recommendation: D x 0; C2,E,F x 1/2; A,B,G x 1.

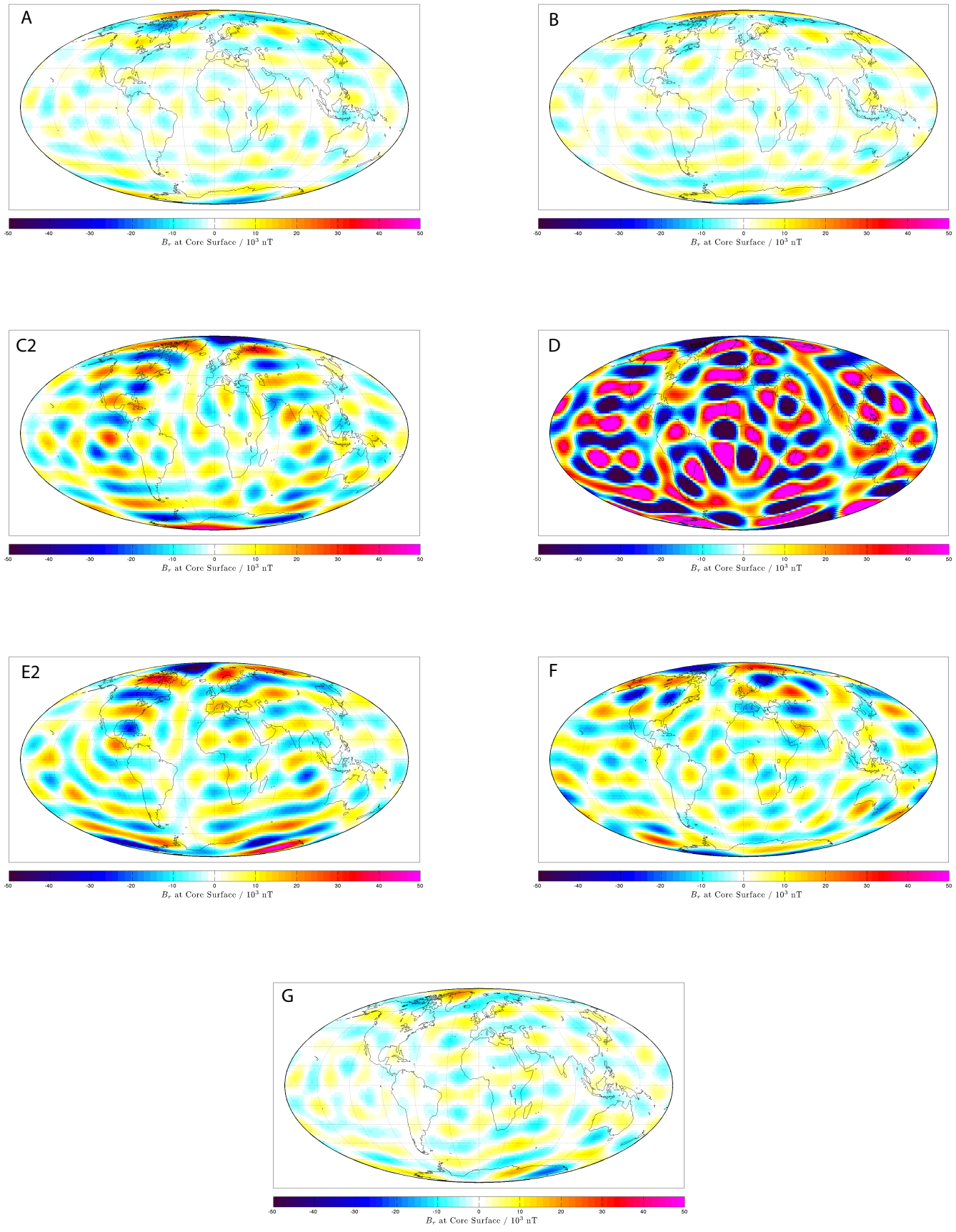

Figure 6: Radial component of magnetic field of the difference between each DGRF-2005 candidate model and the arithmetic mean model excluding candidate D (‘mean_m_noD’) plotted at radius 3481km (the core-mantle boundary) in Mollweide projection.

3 IGRF-2010

3.1 Root Mean Square (RMS) vector field differences between candidates and mean models

Figure 7 displays the root mean square vector field differences $R_{i,j}$ between the IGRF candidates for epoch 2010 and between the candidates and the arithmetic mean model 'mean_m' which is shown in the rightmost column. The bottom row shows $\bar{R}$ the mean differences $R_{i,j}$ (excluding the zero value for the difference between candidates and themselves - see equation 3). The associated numerical values are also reproduced below in Table 5. As expected the differences between the IGRF-2010 candidates is larger than between the DGRF-2005 candidates, with the mean of the $R_{i,j}$ between candidates and the model 'mean_m' being approximately 8.5nT compared to 4.9nT for the DGRF candidates. From Figure 7 it can be seen visually that models D and E display the largest differences to the other candidates and to the model 'mean_m'. Regarding the other candidates, G and B seem to be most similar to each other and to the mean while A, C2 and F show slightly larger differences. It nonetheless seems difficult on this evidence alone to definitively separate A,C2, F and B,G into two distinct categories.

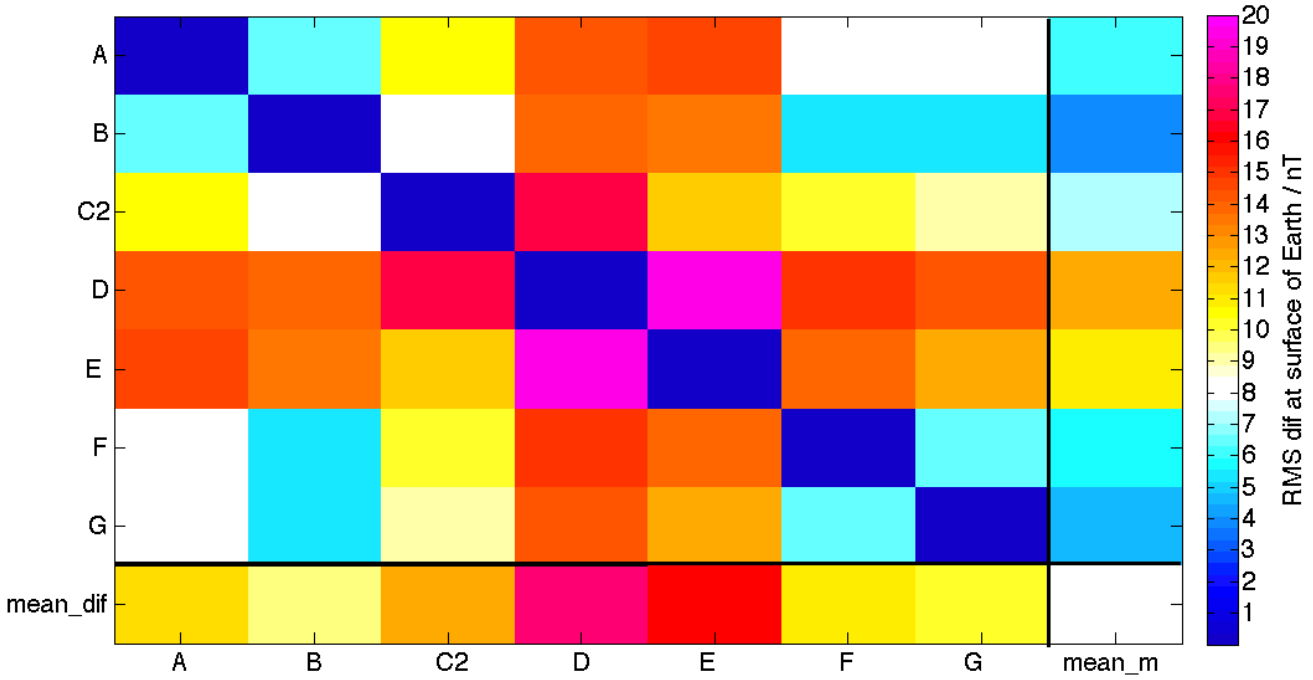

Figure 7: Plot showing RMS vector differences $R_{i,j}$ (see equation 2) in nT between individual IGRF-2010 candidates A, B, C2, D, E, F, G and the arithmetic mean of all candidates (in the rightmost column) referred to as 'mean_m' . The bottom row displays the mean of the RMS differences between each candidate model and all other candidate models $\bar{R}$ (see equation 3) referred to as 'mean_diff'.

3.2 Spectral comparisons

Figure 8 presents the power spectra of the IGRF epoch 2010 candidates as a function of spherical harmonic degree. The power spectra of the candidate models at Earth's surface are again very similar. Candidates E and D are found to have slightly different power from other candidates at degree 13 at Earth's surface; differences are apparent above degree 10 particularly for candidates E and D at the core-mantle boundary.

Figure 9 shows the degree correlation $\rho_n^{i,j}$ between the candidates and the arithmetic mean model 'mean_m'. Candidate D shows the largest differences above degree 10, candidate E also shows considerable differences while candidates C2, F and G show smaller deviations from 'mean_m' with candidates A and B being closest to mean in this test.

Figure 10 presents the spectra (as a function of spherical harmonic degree) of the difference between the candidates and 'mean_m'. Candidates D and E and C2 show the largest differences, particularly at high degrees. Candidates G

	A	B	C2	D	E	F	G	mean_m
A	0.00	6.30	10.59	14.22	14.80	8.19	8.19	6.30
B	6.30	0.00	8.11	13.86	13.35	5.21	5.42	3.83
C2	10.59	8.11	0.00	16.93	11.82	10.04	8.93	7.12
D	14.22	13.86	16.93	0.00	19.44	15.04	14.25	12.29
E	14.80	13.35	11.82	19.44	0.00	14.04	12.44	10.93
F	8.19	5.21	10.04	15.04	14.04	0.00	6.59	5.75
G	8.19	5.42	8.93	14.25	12.44	6.59	0.00	4.59
mean_diff	11.43	9.35	12.26	17.67	16.14	10.81	10.07	8.47

Table 5: Numerical values of RMS vector field differences $R_{i,j}$ in units nT between candidate models and mean models from Figure 7

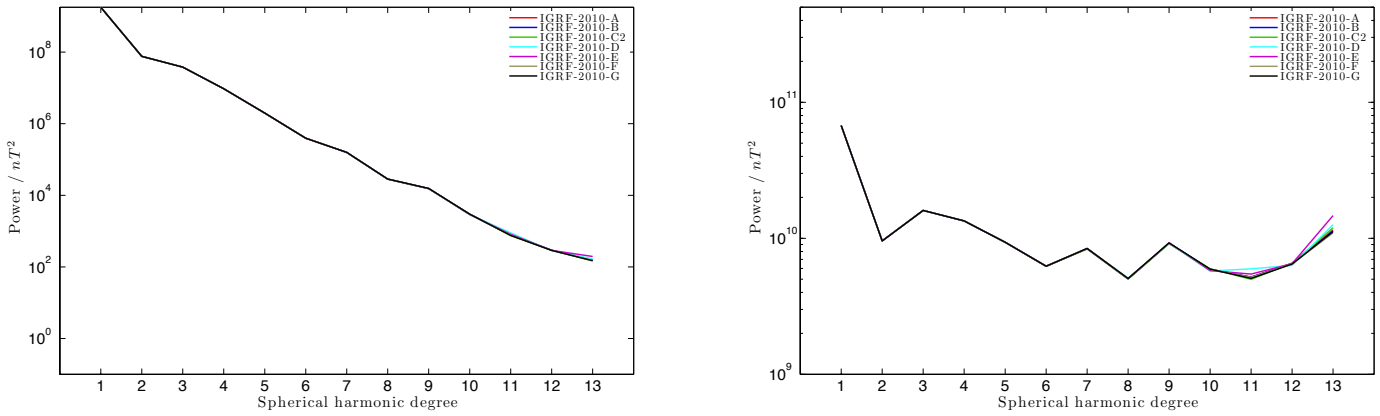

Figure 8: Power spectra of IGRF-2010 candidate models at Earth's surface (left) and at radius 3481km (core-mantle boundary) (right).

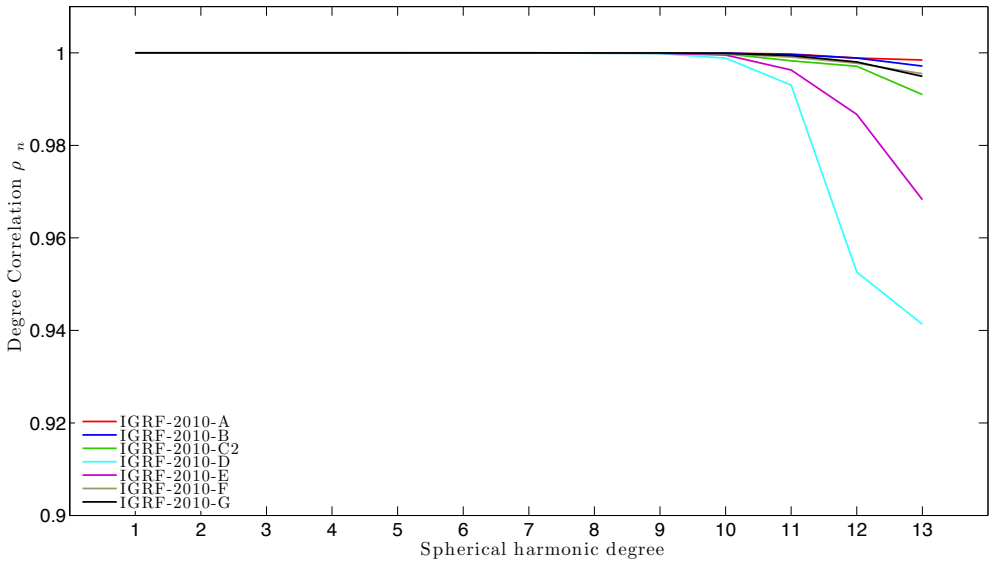

Figure 9: Degree correlation of IGRF-2010 candidate models with the arithmetic mean model 'mean_m' .

and F show similar levels of difference to 'mean_m' at high degrees, though candidate F also shows slightly larger differences at low degrees. Candidate B is consistently close to 'mean_m' while model A shows a noticeable large difference in degree 1.

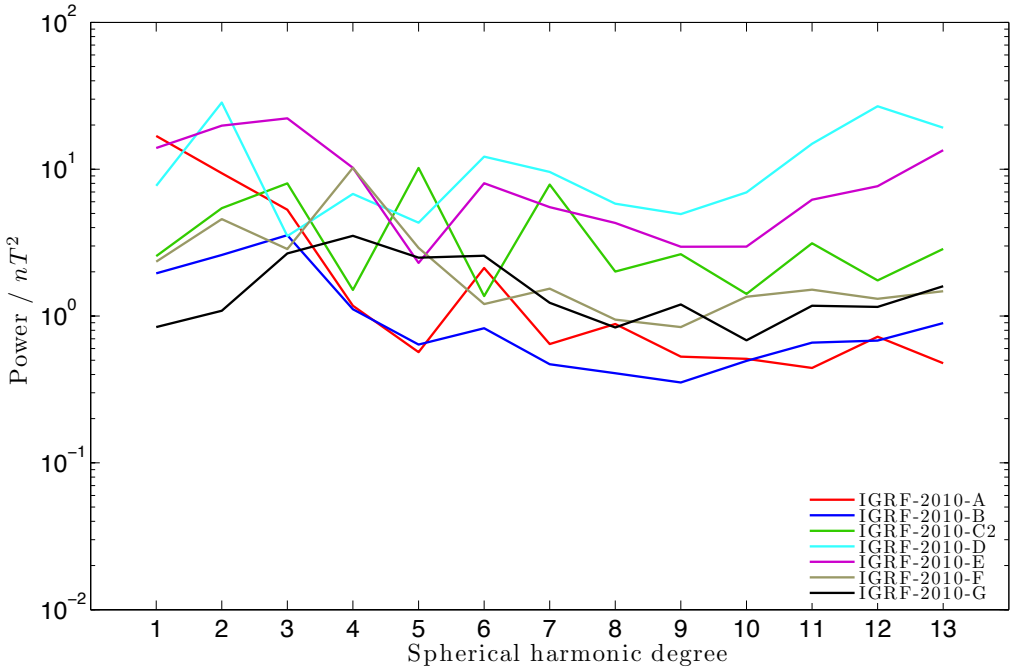

Figure 10: Power spectra of difference between IGRF-2010 candidate models and the arithmetic mean model 'mean_m' .

Considering the plots of sensitivity $S(n, m)$ in Figure 11 between the IGRF-2010 candidate models and 'mean_m', it is noticeable that candidates A, B and G show only very small differences from the mean model. Candidate C2 is found to possess systematic differences in the zonal (axisymmetric m=0) terms for degrees 11, 12 and 13. Candidate E shows major ($S > 50$) differences in the coefficients $g^{12}_{12}, g^{1}_{13}, g^{2}_{13}, h^{1}_{11}, h^{1}_{12}, h^{13}_{13}$ and especially h^{1}_{13} where $S \sim 150$. Candidate F shows differences compared to the mean of the opposite sign to Candidate E in coefficients $h^{1}_{11}, h^{1}_{12}, h^{1}_{13}$ though of smaller amplitude. Candidate D shows the largest differences over many coefficients, particularly in degrees 11, 12 and 13 with the largest discrepancies seen in orders m=0 and 1.

3.3 Physical space comparisons

Differences between IGRF epoch 2010 candidate models and model 'mean_m' in the radial magnetic field at the core-mantle boundary are presented in Figure 12. The largest differences are observed for candidate D, particularly in the northern hemisphere. The differences are rather global, with the exception of a region under Southern America. Candidate E also displays differences from 'mean_m' locally in access of 10nT, particularly at equatorial latitudes and in the polar regions. Candidate C2 displays strong differences from the mean model in the polar regions. Candidates A, B, F, and G show more minor differences, in all cases the differences are largest in the polar regions.

3.4 Discussion

The comparisons above suggest that candidates D and E are most different from the other candidates and from 'mean_m'. These should therefore be down-weighted compared to the other models in the calculation of the final IGRF for epoch 2010. Candidates C2, F and A (the later because of discrepancies in degree 1) could also perhaps be slightly down-weighted with respect to candidates B and G which seem on the basis of the comparisons above to be the most trustworthy candidates.

Recommendation: D, E x 1/4; C2,F,A x 1/2; B,G x 1.

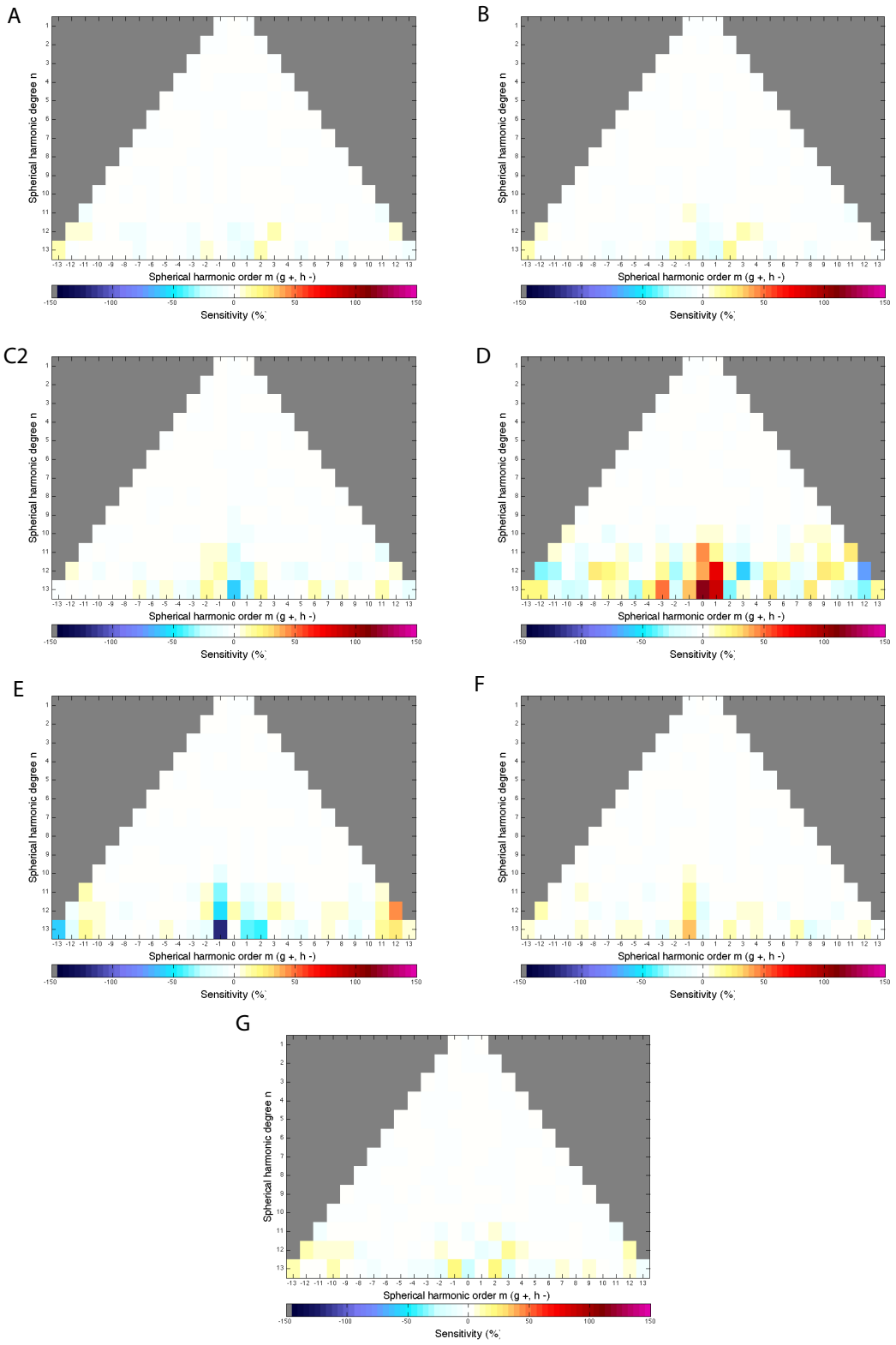

Figure 11: Matrices of sensitivity $S(n, m)$ (see equation 5) for IGRF-2010 candidate models A, B, C2 ,D ,E ,F ,G compared to the arithmetic mean model ('mean_m').

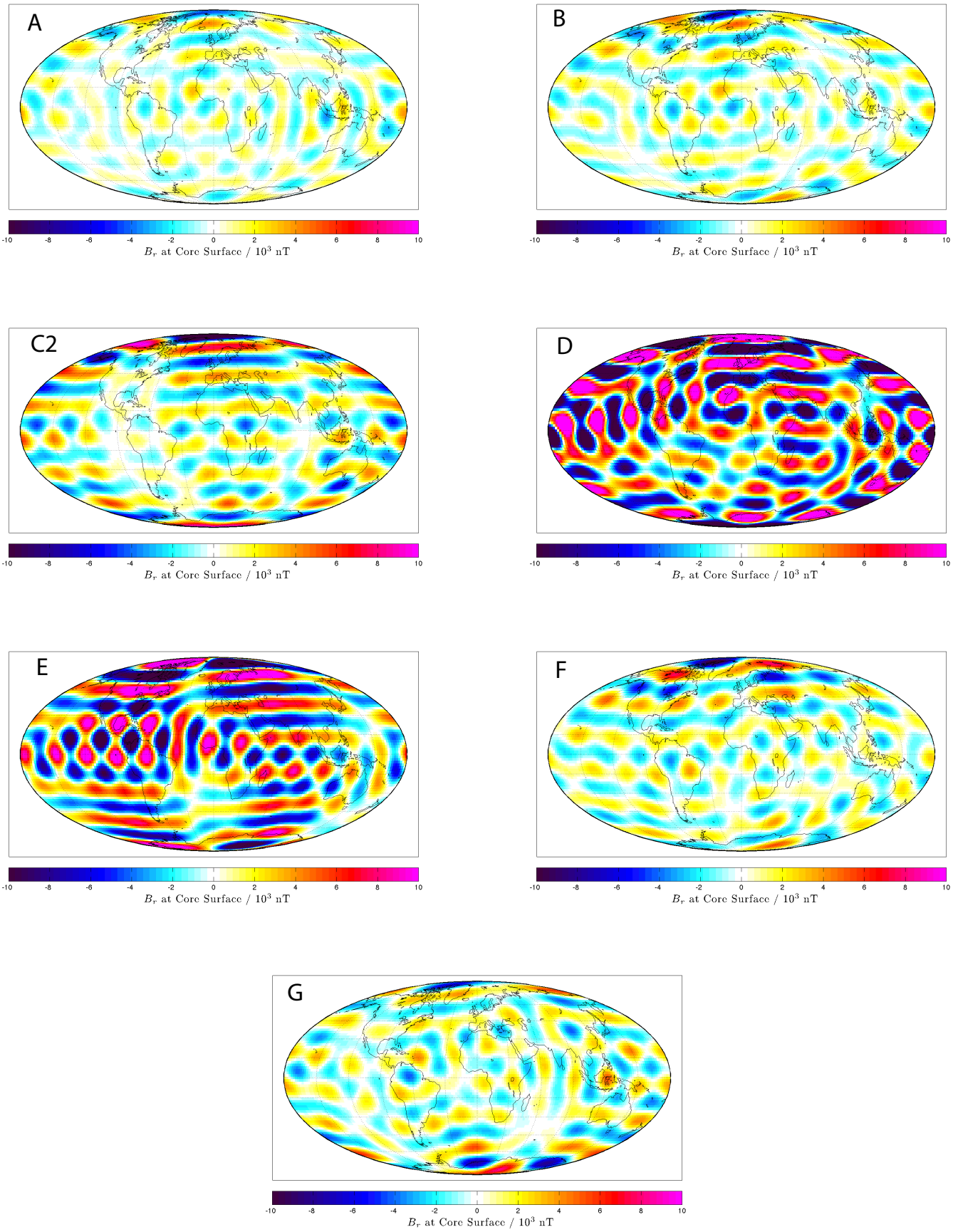

Figure 12: Radial component of magnetic field of the difference between each IGRF-2010 candidate model and the arithmetic mean model 'mean_m' plotted at radius 3481km (the core-mantle boundary) in Mollweide projection.

4 SV-2010-2015

4.1 Root Mean Square (RMS) vector field differences between candidates

Once again, the analysis of the candidates (this time the average predictive SV to degree 8 for the epoch 2010-2015) will begins with consideration of a compilation of calculations of $R_{i,j}$ for the global RMS vector field differences between the candidates and an arithmetic mean model 'mean_m'. The results of this analyses are shown graphically in Figure 13 with colors again encoding the magnitude of $R_{i,j}$; the numerical values are additionally given in Table 6. In this case there is much more diversity and spread in the predictions of candidate models compared to the earlier analysis of the main field candidates. Candidates D and F are closest to the mean model, followed by candidates B and H, then come candidates A and E with candidate G most different from the mean model.

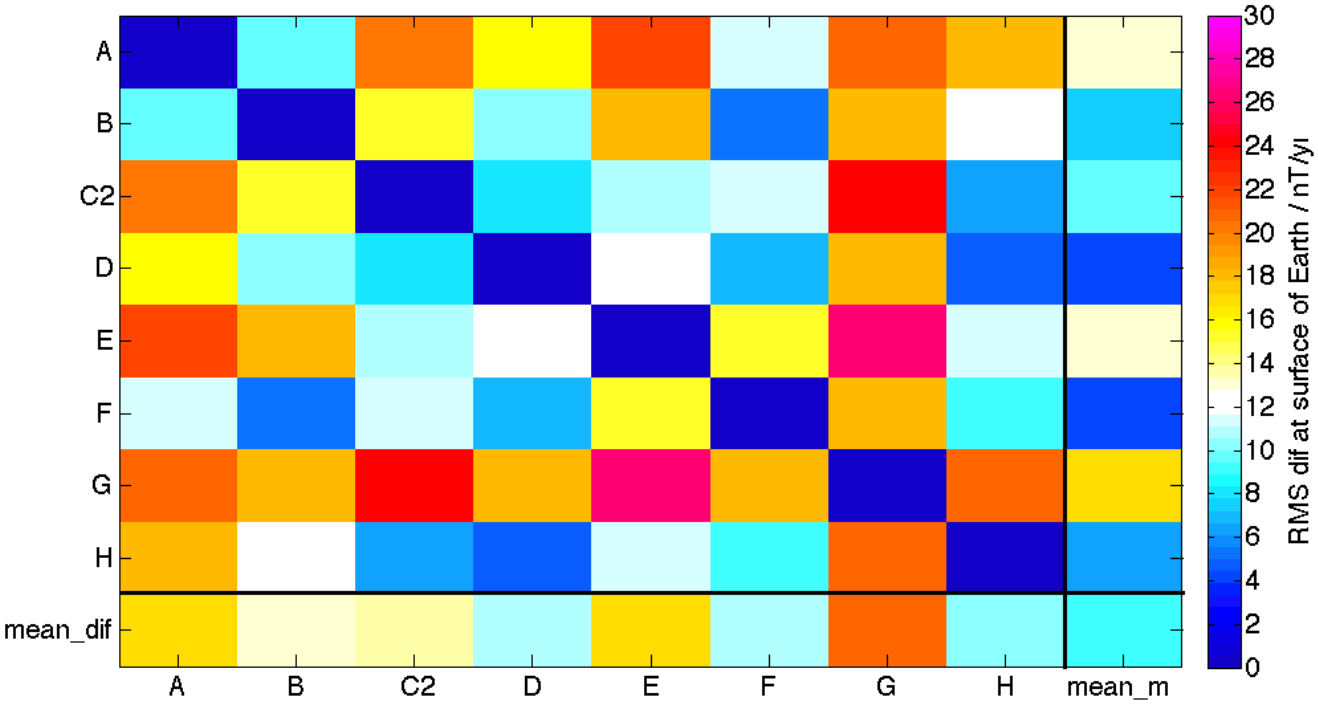

Figure 13: Plot showing RMS vector differences $R_{i,j}$ (see equation 2) in nT/yr between individual SV-2010-2015 candidates A, B, C2, D, E, F, G,H and the arithmetic mean of all candidates (in the rightmost column) referred to as 'mean_m' . The bottom row displays the mean of the RMS differences between each candidate model and all other candidate models $\bar{R}$ (see equation 3) referred to as 'mean_diff'.

	A	B	C2	D	E	F	G	H	mean_m
A	0.00	9.98	20.23	15.93	22.16	11.45	20.99	18.07	12.79
B	9.98	0.00	15.36	10.17	18.31	5.11	18.12	12.48	7.43
C2	20.23	15.36	0.00	7.96	11.05	11.44	24.22	6.47	9.67
D	15.93	10.17	7.96	0.00	12.73	6.71	18.17	4.68	4.14
E	22.16	18.31	11.05	12.73	0.00	15.33	26.29	11.62	12.89
F	11.45	5.11	11.44	6.71	15.33	0.00	18.14	8.98	4.11
G	20.99	18.12	24.22	18.17	26.29	18.14	0.00	20.72	16.93
H	18.07	12.48	6.47	4.68	11.62	8.98	20.72	0.00	6.60
mean_diff	16.97	12.79	13.82	10.91	16.78	11.02	20.95	10.38	9.32

Table 6: Numerical values of RMS vector field differences $R_{i,j}$ in units nT/yr between candidate models and mean models from Figure 13

4.2 Spectral comparisons

Considering the power spectra of the predictive SV candidate models both at Earth's surface and the core-mantle boundary in Figure 14 there appears to be little to choose between the models which are now widely spread rather than tightly bunched. Nonetheless, candidate E noticeably possess a very different spectral slope for degree 5 to 8 (degree 7 appears particularly anomalous) while candidate G contains noticeably high power above degree 5.

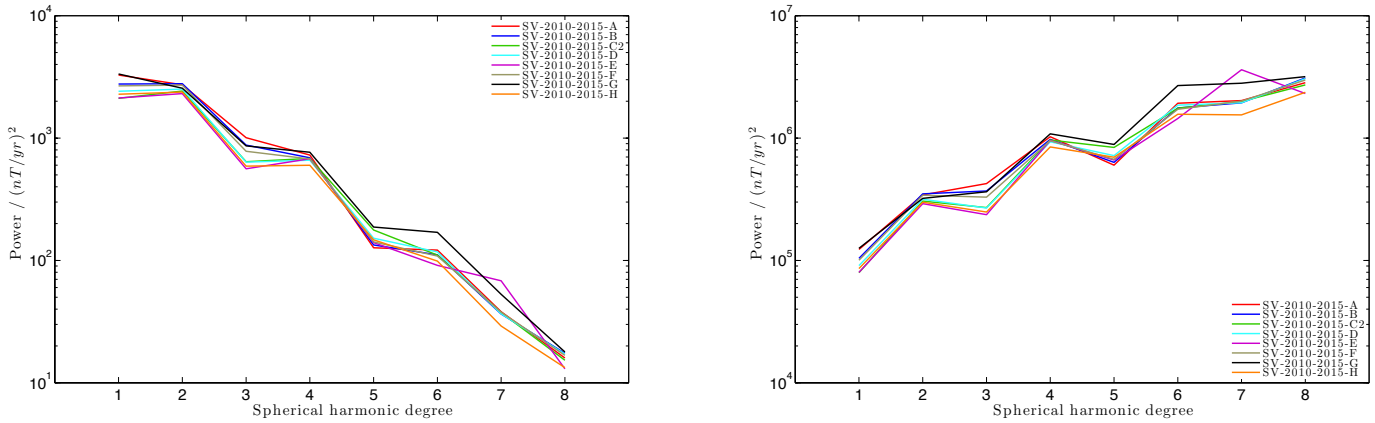

Figure 14: Power spectra of SV-2010-2015 candidate models at Earth's surface (left) and at radius 3481km (core-mantle boundary) (right).

The degree correlation $\rho_n^{i,j}$ between the candidate models and model 'mean_m' shown in Figure 15 shows the phase of candidate E is very different from the other models above degree 5 while candidate G is also noticeably different in degrees 3, 5 and 8. Candidate C2 also has systematically lower correlations to the mean model that the remaining models, but it is less clearly distinct than candidates E and G.

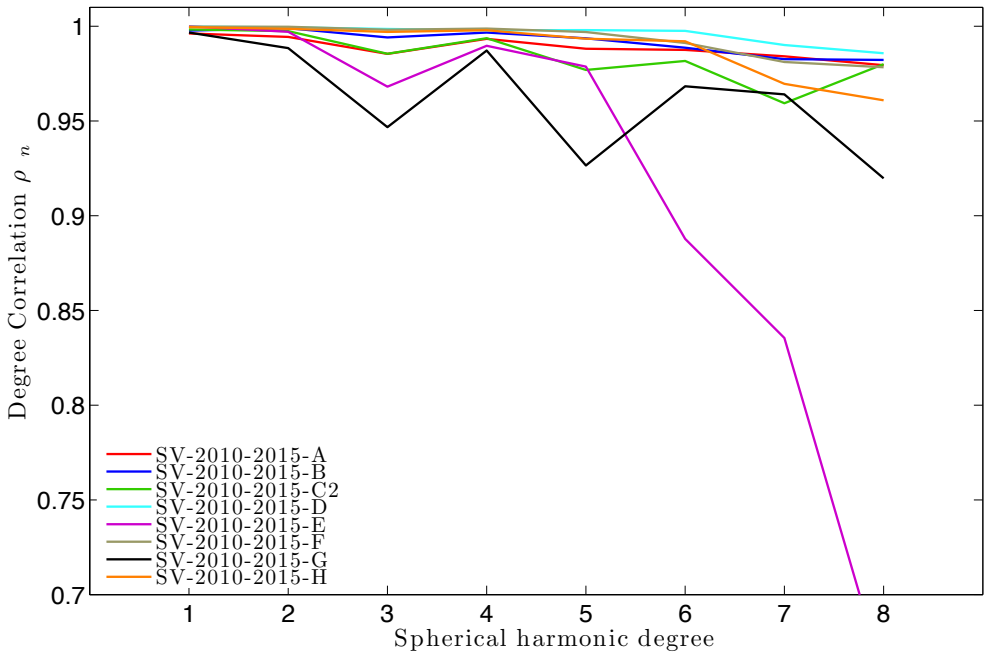

Figure 15: Degree correlation $\rho_n^{i,j}$ of SV-2010-2015 candidate models with the arithmetic mean model 'mean_m' .

The power spectra of the differences between the candidates and 'mean_m' is presented in Figure 16. It displays similar trends to those found in the earlier tests with candidate G possessing consistently higher power, candidate E with anomalous behaviour above degree 5 and candidates F and D closest to 'mean_m' having smallest differences. It is also noticeable that candidate A possesses a steeper spectrum than the other candidates with relatively high power

in degree 1 and 2 and comparatively low power in degree 7 and 8.

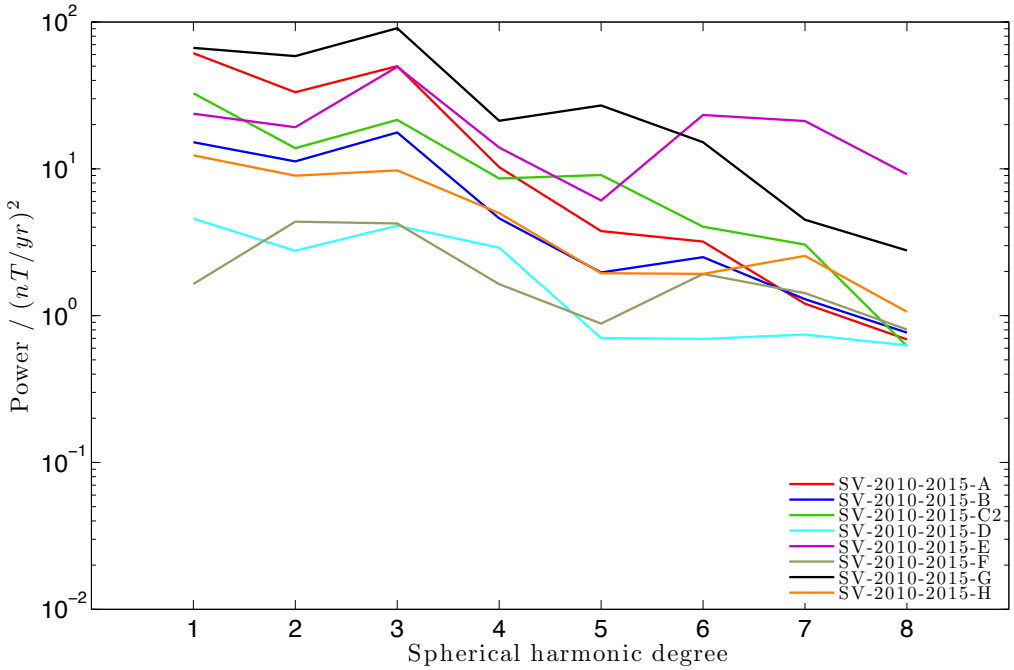

Figure 16: Power spectra of difference between SV-2010-2015 candidate models and the arithmetic mean model 'mean_m' .

Figure 17 presents sensitivity plots between the SV candidates and model 'mean_m'. These plots show much greater differences in the coefficients of all degrees compared to similar plots for main field candidates where differences were largely confined to degrees greater than 10. Candidate E is found to possess strongly anomalous g_7^7, h_7^7, h_8^1 and g_8^0 coefficients. Candidate G also possesses numerous anomalous coefficients, for example h_3^3, g_5^5, g_6^6, g_8^8 and h_8^6. Many of the largest differences for candidate G seem to occur in its sectoral terms. Candidate C2 possesses an anomalous g_7^7 term but seems to be otherwise be typical while candidate F is found to possess differences in the g_7^1 and g_8^0 coefficients compared to the mean model. Candidate A has remarkable large differences from the mean model in g_0^1 and h_1^1 as well as in g_3^3 and g_8^8. Candidate B is found to have h_8^8 and h_8^2 coefficients with large sensitivity. Candidates D and H are generally close to the mean model, though D possess a slightly anomalous h_8^1 and H possess a h_8^8 rather different from the mean model.

4.3 Physical space comparisons

Looking at the differences in the predicted change in radial magnetic field at the core-mantle boundary between the candidates and model 'mean_m' (see Figure 18) reveals that candidate E is globally anomalous with the local differences to 'mean_m' exceeding 25nT/yr in many locations. The differences between candidate G and the mean model are most visible at low latitudes under the Atlantic hemisphere, where important field change is known to have taken place for several centuries. The differences between candidates H, C2, B and 'mean_m' are also most striking at low latitudes. Accurate determination of the evolution of flux features in this region of the core surface is crucial in obtaining accurate secular variation predictions - it will be of great interest in the upcoming five years to see whether any of the candidates (including H which based on an approximation of core MHD) performs better than the 'mean_m' - it is unfortunately not currently possible to make a prior judgment on this matter. Candidates A and D display rather weaker differences from the 'mean_m', at least in terms of localized secular variation foci at the core-mantle boundary.

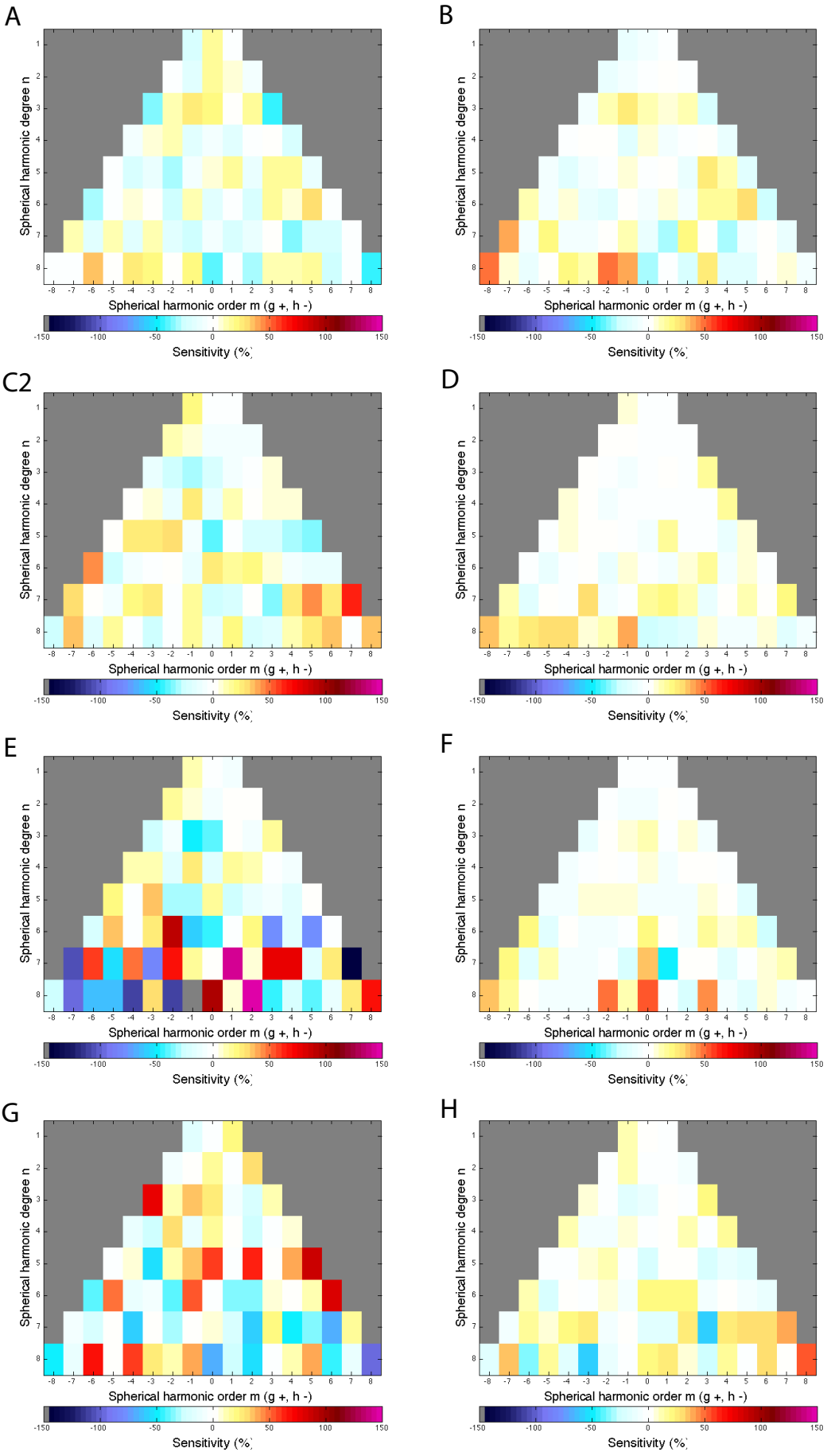

Figure 17: Matrices of sensitivity $S(n, m)$ (see equation 5) for SV-2010-2015 candidate models A, B, C2 ,D ,E ,F ,G compared to the arithmetic mean model ('mean_m').

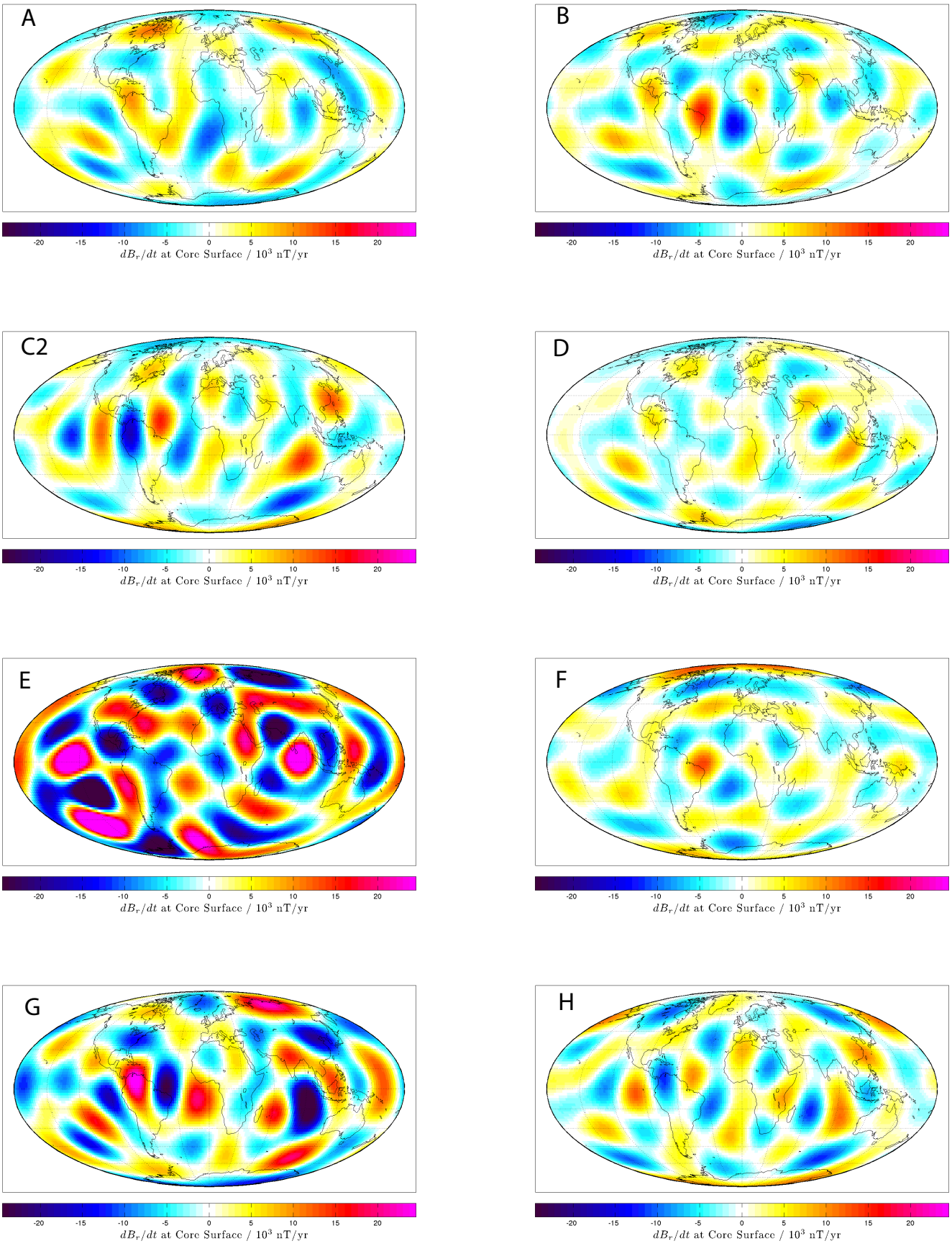

Figure 18: Radial component of magnetic field of the difference between each SV-2010-2015 candidate model and the arithmetic mean model (‘mean_m’) plotted at radius 3481km (the core-mantle boundary) in Mollweide projection.

4.4 Discussion

The decisions regarding the weighting of the predictive SV candidates is more challenging than for MF models because the candidates are much more spread. One must therefore be very careful in making assessments using an arithmetic mean model that may be of dubious relevance. Nonetheless, in the tests performed here it seems that candidate E is consistently different from the other candidates, particularly at spherical harmonic degree greater than 5 and also in physical space (see Figures 15, 16, 17, 18). The task force should therefore consider down-weighting candidate E. The reason for its anomalous behaviour above degree 5 is unclear, but may related to reliance on observatory data with incomplete global coverage. Candidate G is also consistently different with stronger power particularly in the some of the sectoral coefficients; unless convincing arguments are provided as to why it should be as reliable as the other candidates it should perhaps also be down-weighted. It is difficult to make very strong arguments for relative weighting of the other candidates; D, F and H are closer to the mean model but it is difficult to argue here that the mean model obviously the best model. The differences between SV candidates are largely the result of differences in the extrapolation procedures and certainly not random samples so purely statistical reasoning may be flawed.

Recommendation: E,G x 1/2; A, B, C2, D, F, H x 1

5 Concluding remarks

This study has focused on statistical comparisons of candidate models, following the example of the most recent previous IGRF evaluations by Maus et al. (2005). Comparisons with independent data should also be taken seriously, and if defensible factored into the final deliberations. Explanations for deviations of certain candidates from most other candidates should be sought in the data selection and modelling strategies. If there are defensible reasons why certain candidates should be distinct in particular ways, then down-weighting schemes could be modified to take this into account so that such candidates are not unfairly penalized. Discussions along these lines will no doubt take place in amongst the task force in the days ahead.

It is remarkable that the submitted candidate models are generally in good agreement. Model differences for retrospective epochs such as for DGRF 2005 are small and apparently now largely a question of the criteria used to select satellite data, the pre-processing (filtering, corrections etc.) applied and of the parameterization of the external field. For the predictive models such as IGRF for epoch 2010, the method of extrapolation into the near future still appears to play an significant role (for example whether one uses spline extrapolation from the end of a time-dependent model, uses quadratic terms in a Taylor series expansions, or fits a trend to data from a longer interval). Accurate determination of predictive SV remains the major challenge in the IGRF process; a large scatter in the submitted candidate models was again present in the IGRF-11 SV candidates. It will be of interest over the next 5 years to see whether data assimilation methods based on approximations of core physics (for example SV candidate H) are yet at the stage where they can yield more useful forecasts than more traditional statistical extrapolation strategies.

References

Langel, R. A., Hinze, W. J., 1998. The Magnetic Field of the Earth's Lithosphere: The Satellite Perspective. Cambridge University Press.

Lowes, F., 1966. Mean-square values on the sphere of spherical harmonic vector fields. J. Geophys. Res. 71, 2179.

Lowes, F., 1974. Spatial power spectrum of the main geomagnetic field. Geophys. J. R. Astron. Soc. 36, 717–730.

Maus, S., Macmillan, S., Lowes, F. J., Bondar, T., 2005. Evaluation of candidate geomagnetic field models for the 10th generation of igrf. Earth. Planet. Space 57, 1173–1181.

Sabaka, T., Olsen, N., 2006. Enhancing comprehensive inversions using the swarm constellation. Earth. Planet. Space 58, 371–395.