Evaluation of IGRF-11 candidate models

Stefan Maus, 16-Nov-09

Summary of recommendations

DGRF-2005: Despite being based on very different methods and data sources, DGRF candidates A, B and G are in exceptionally good agreement. Since DGRF will be cast in stone, while the IGRF and SV will be revised 5 years from now, I recommend putting aside political considerations for the DGRF and not adding down-weighted poorer candidates to the mix. In order to avoid a large rounding artifact, the DGRF coefficients should be published at the appropriate 0.01 nT precision.

IGRF-2010: Candidates D and E exhibit unrealistic distributions of power, both with degree and azimuth. I recommend weighting these two candidates with 0.5, while giving all other candidates a weight of 1.0.

SV-2010-2015: Candidates E and G overestimate the change of the magnetic field in degrees 6-8. I recommend weighting these two candidates with 0.5. Model H underestimates the change, which is not necessarily detrimental. I therefore recommend including H with the other candidates at full weight.

DGRF standard deviation and comparison with rounding error

Assuming that the candidates are statistically independent, an unbiased estimator of the sample standard deviation *s* is given by $s = \sqrt{\frac{1}{K-1}\sum_k \sum_n (n-1) \sum_m ({}_k g_n^m - \overline{g_n^m})^2}$, from which the error *e* of the mean of K models follows as $e = \frac{s}{\sqrt{K}}$.

Correspondingly, the per-degree sample standard deviation s_n is $s_n = \sqrt{\frac{(n-1)}{K-1}\sum_k \sum_m ({}_k g_n^m - \overline{g_n^m})^2}$ and the per-degree error of the mean e_n follows as $e_n = \frac{s_n}{\sqrt{K}}$.

Ranking the models by their deviation from the mean model, the quantities *s* and *e* are given in Table 1.

Models	Sample standard deviation (nT)	Error of the mean (nT)
B, G	1.56	1.10
B, G, A	1.76	1.02
B, G, A, F	2.46	1.23
B, G, A, F, E2	3.1	1.39

B, G, A F, E2, C2	3.4	1.39
B, G, A F, E2, C2, D	6.3	2.38
Error introduced by rounding the coefficients to 0.1 nT		1.3 nT

Table 1: The error of the mean in the right column provides a realistic estimate of the RMS error of the resulting average model at the Earth surface

As has been pointed out by others, there is a very good agreement between models A, B and G, despite the completely different methodology. It can therefore convincingly be argued that the average of these three models must be very close to the truth. One can see that adding A to the two models B and G reduces the error of the mean despite the larger standard deviation. However, due to the poorer quality of the other models, including them in the average significantly increases the error of the resulting average model.

Figure 1 shows the corresponding per-degree error. While the error curve for models A, B, and G is flat, all other models appear to have problems at higher degrees. As pointed out before, if the DGRF coefficients are published at 0.1 nT precision, the primary error of the 2005 DGRF will be due to the rounding of the coefficients.

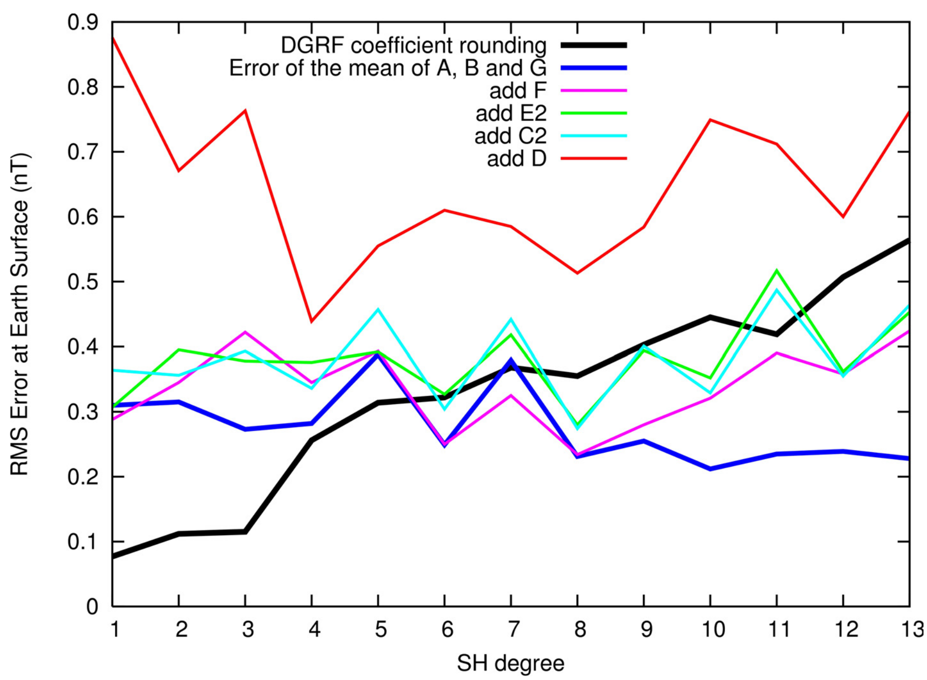

Figure 1: Per-degree error of the mean for various model averages. The blue curve indicates that by far the best DGRF would be obtained by an average of models A, B and G. The black line shows the error that would be introduced by rounding the coefficients to 0.1 nT precision.

IGRF-2010 and SV-2010-2015

While the Gauss coefficients change over time in unpredictable ways, it can be argued that the power distribution per degree and order is not likely to suddenly change very significantly from one epoch to the next. A strong deviation of power distribution from IGRF-2005 and from SV-2005-2010 therefore indicates problems in the modeling approach. Figures 2 and 3 compare the power distribution of the IGRF-2010 candidates with IGRF-2005. Only models D and E stand out as anomalous. Figure 4 shows the per-degree error of the mean, corresponding to Figure 1. Figures 5 and 6 display the distribution of power of the SV-2010-2015 candidates in comparison with SV-2005-2010. In this case, models E and G exhibit high power, meaning that they predict unrealistically strong changes in the field. Model H, on the other hand, underestimates the change of the field. This may however not be bad. A "best estimator" of upcoming change could very well have lower power. Figure 7 shows the per-degree error of the mean for the SV candidate averages.

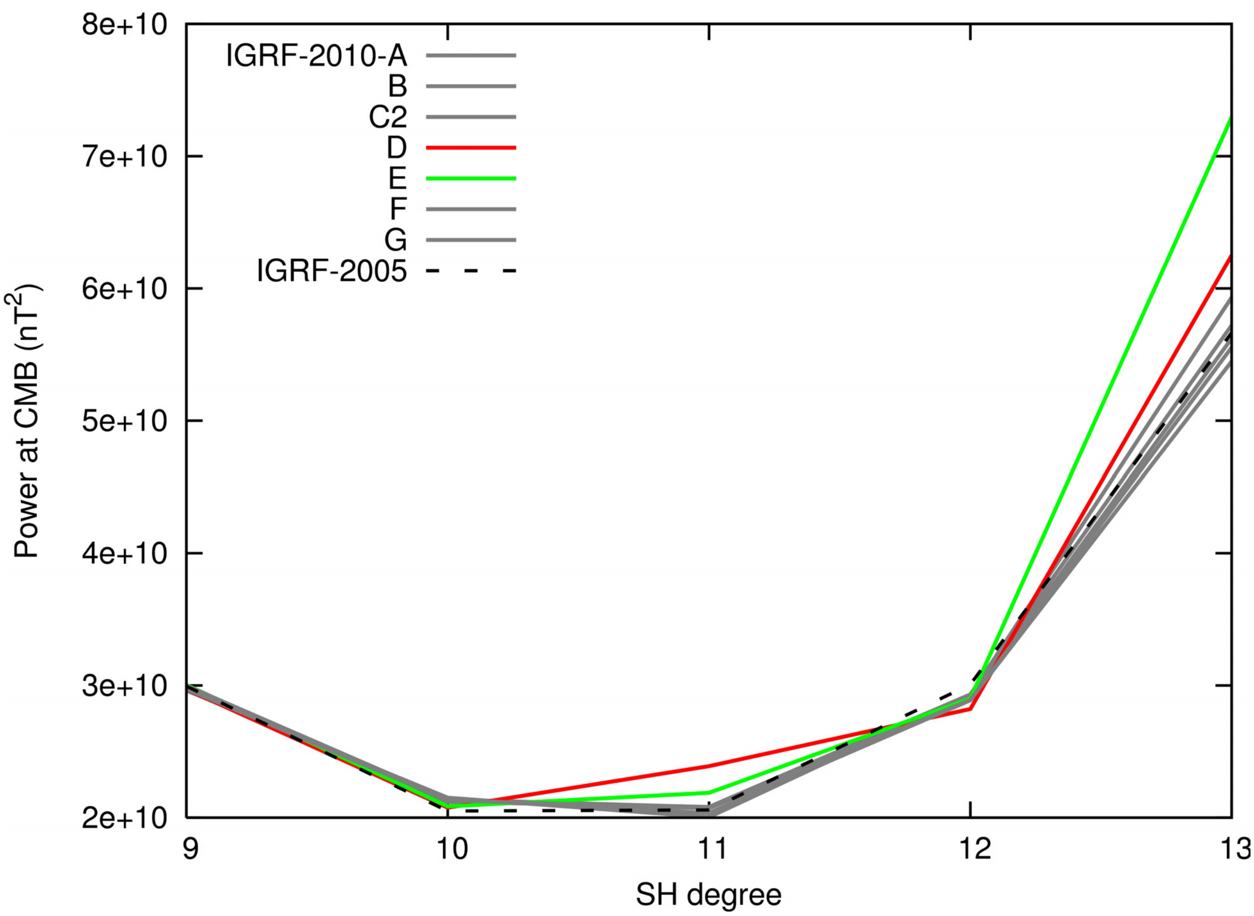

Figure 2: Lowes/Mauersberger power in degrees 9-13 at the C/M boundary. The corresponding power of the IGRF-2005 is given as a dashed line. Models D and E deviated significantly from the expected power.

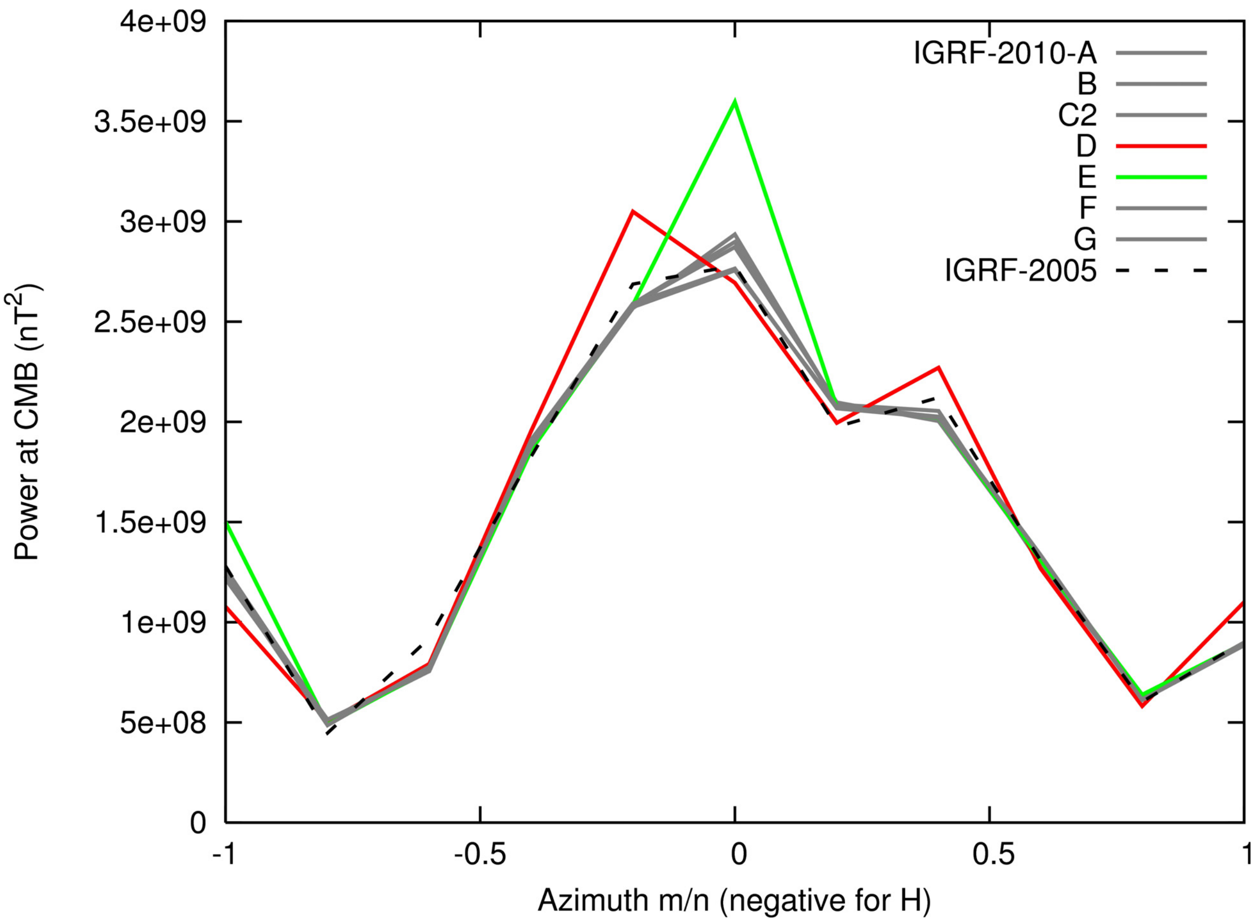

Figure 3: Azimuthal distribution of power at the C/M boundary. Models D and E again exhibit unrealistic deviations from the expected power distribution.

Models	Sample standard deviation (nT)	Error of the mean (nT)
B, G	3.8	2.7
B, G, F	4.1	2.4
B, G, F, A	4.8	2.4
B, G, F, A, C2	5.6	2.5
B, G, F, A, C2, E	7.2	2.9
B, G, F, A, C2, E, D	8.5	3.2
Error introduced by rounding the coefficients to 0.1 nT		1.3 nT

Table 2: The errors of the mean confirm that only the addition of the last two models E and D has a detrimental effect on the accuracy of the average. Models E and D should therefore be down-weighted.

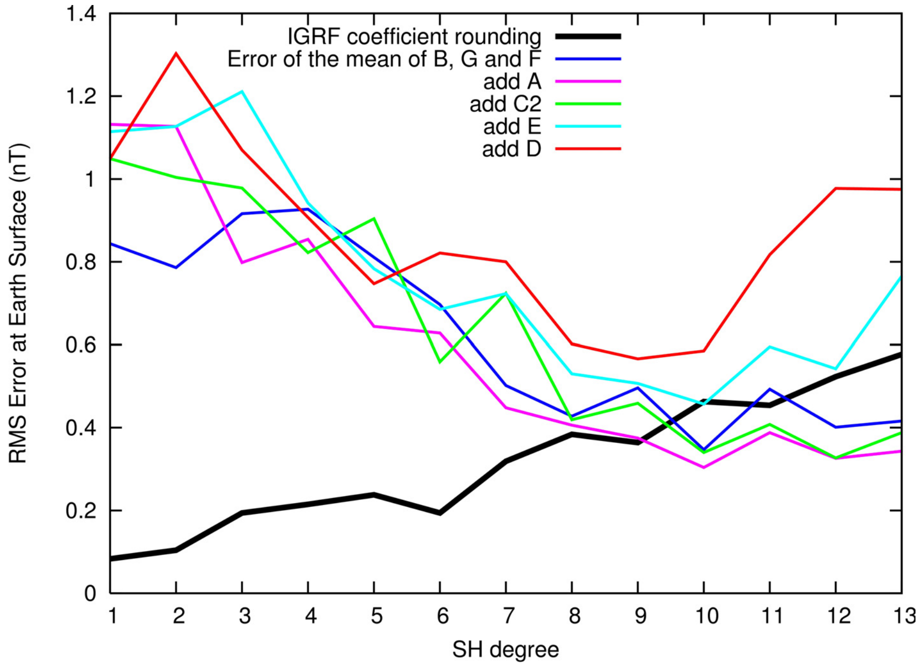

Figure 4: Per-degree error of the mean for IGRF candidate averages. Models E and D have an adverse effect at high degrees. Rounding the coefficients to 0.1 nT precision will have a significant adverse effect at degrees 8-13 not only for DGRF but also for IGRF-2010.

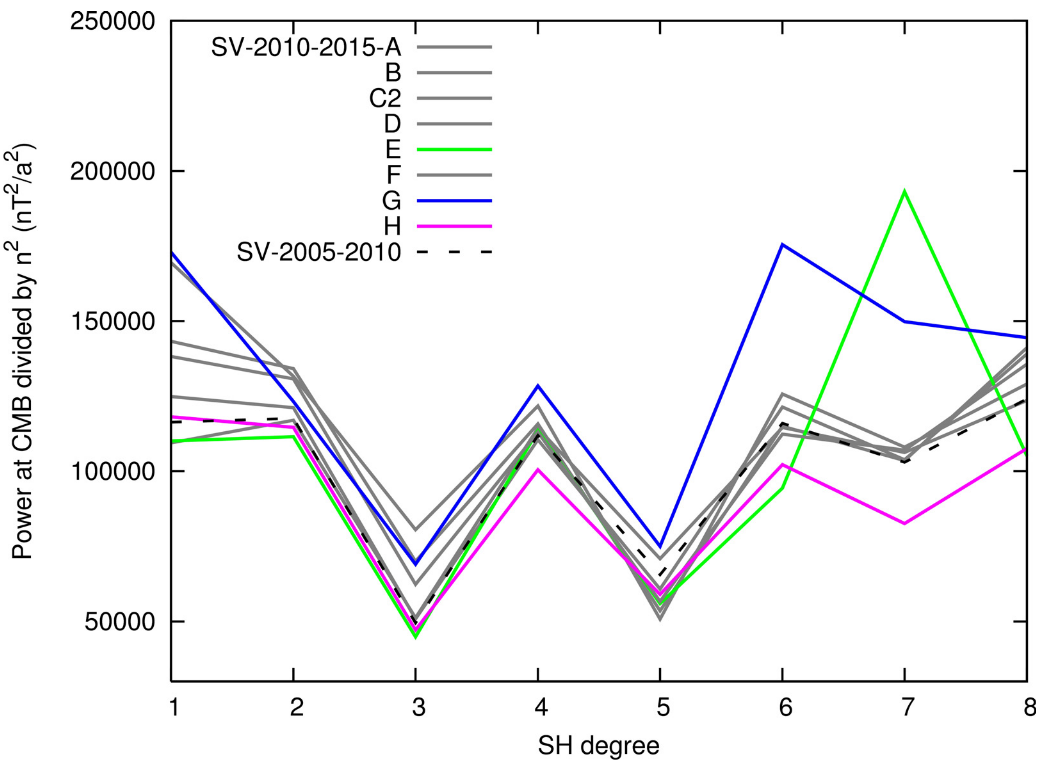

Figure 5: Lowes/Mauersberger power of the SV at the C/M boundary. The corresponding power of the SV-2005-2010 is given as a dashed line. The power in models E and G is unrealistically high at degrees 6-8, while the power in model H is unrealistically low.

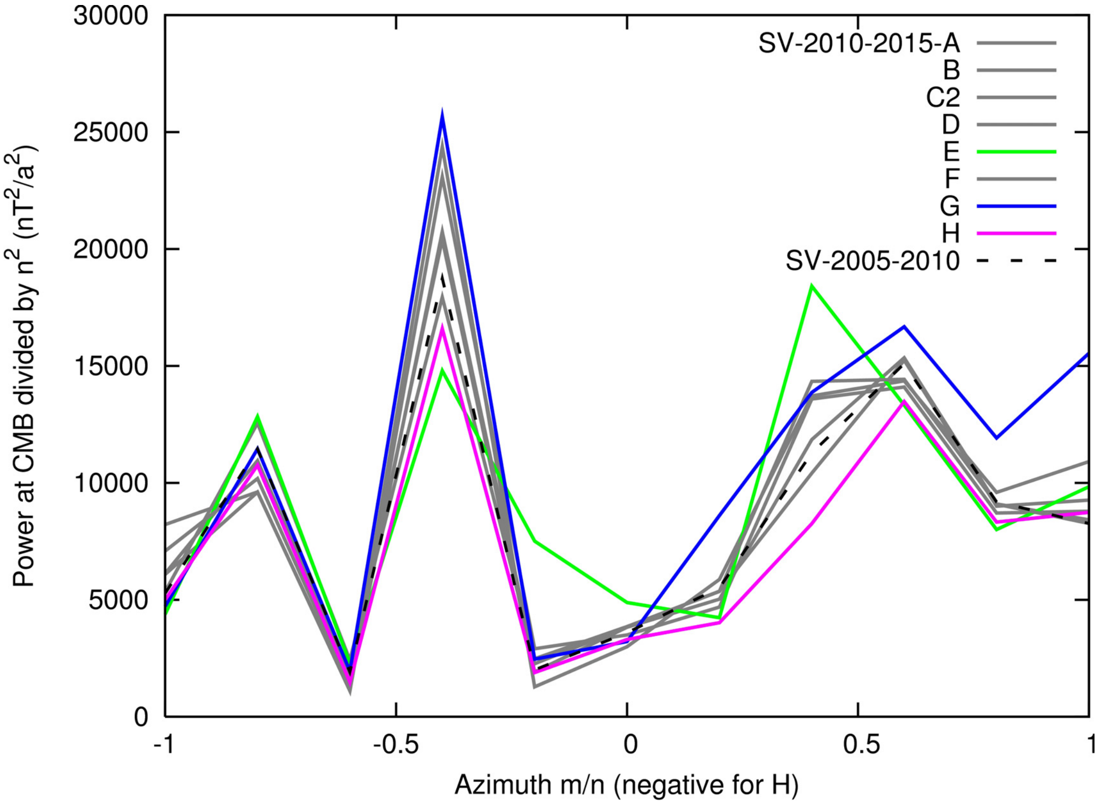

Figure 6: Azimuthal distribution of power at the C/M boundary.

Models	Sample standard deviation (nT/a)	Error of the mean (nT/a)
D, F	4.7	3.4
D, F, H	5.0	2.9
D, F, H, B	6.0	3.0
D, F, H, B, C2	6.7	3.0
D, F, H, B, C2, A	8.4	3.4
D, F, H, B, C2, A, E	9.3	3.5
D, F, H, B, C2, A, E, G	11.0	3.9
Error introduced by rounding the coefficients to 0.1 nT		0.7 nT/a

Table 3: Sample standard deviation and errors of the mean for SV candidate averages.

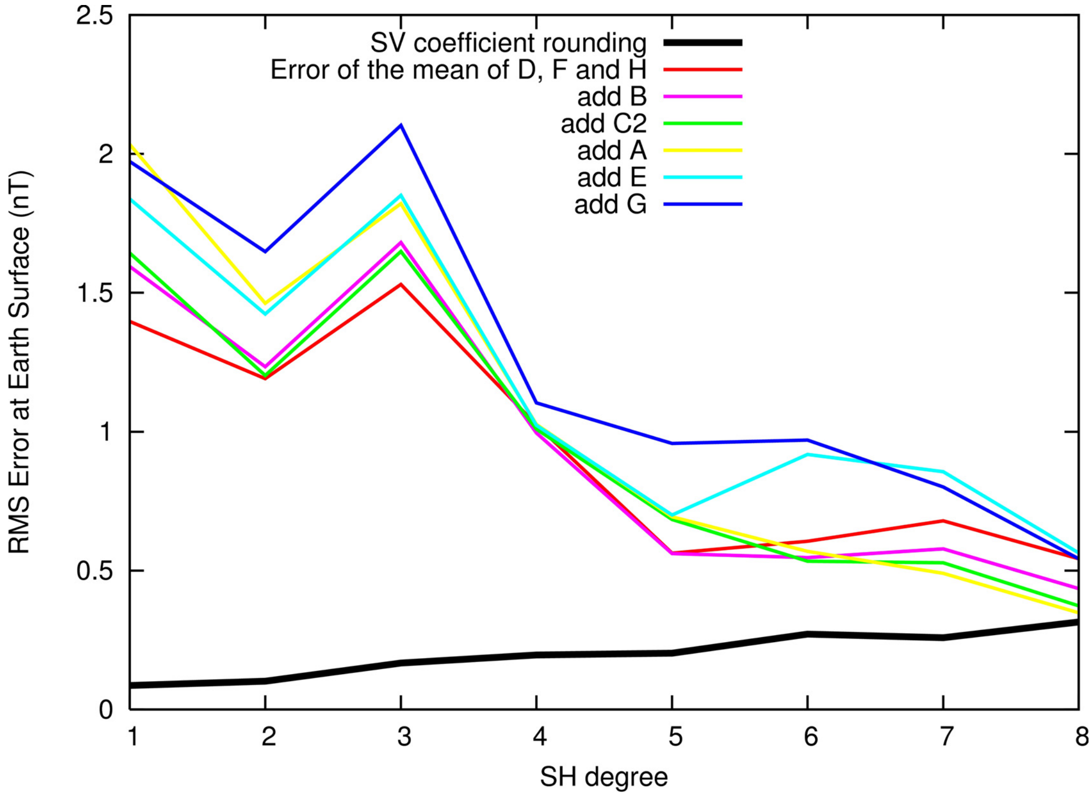

Figure 7: Per-degree error of the mean for SV candidate averages. Only models E and G appear to have an adverse affect. Rounding the coefficients to 0.1 nT precision would only have a significant adverse effect if the SV coefficients were provided to degree 13.

Conclusions and recommendation

The excellent satellite coverage of the past decade could allow us to generate a DGRF-2005 with an unprecedented RMS accuracy of 1.0 nT. However, this accuracy can only be achieved if the three models A, B and G are selected and all others discarded.

For IGRF-2000 and SV-2010-2015 the ranking of the models is less obvious and uncertainties in the field evolution probably outweigh much of the discrepancies between models. The distribution of power with degree and order ("azimuth") shows, however, that IGRF-2010 candidates D and E have artifacts and should be down-weighted. Correspondingly, SV candidates E and G overestimate the change of the magnetic field and should therefore be down-weighted. Model H underestimates the power of the SV, which is not necessarily detrimental and should therefore not be a reason for down-weighting.

What is clear is that the improved accuracy of our observing systems demands that we improve the precision of the coefficients from 0.1 nT to 0.01 nT. This should already be done for DGRF-2005, while the IGRF-2010 and SV-2010-2015 can be published at the old resolution without creating too much harm. However, IGRF-12 should certainly be published at 0.01 nT precision, in particular since the SV will likely be extended to degree 13.