# Evaluation of Candidate models IGRF-11

## Summary DGRF Evaluation

- DGRF models are almost statistically equivalent except for D. From the analyses of the power spectra of the difference between the candidates and the mean model we have three groups (A,B,G), (E,F,C), and D. A complementary analysis based on residuals between models and some selected CHAMP data allow us to classify the candidate models in three groups (A,B,G,F,E), C and D. The difference between these two classifications relies on tenuous differences observed when plotting the histograms of residuals.

### *Suggestion*

| |
|---|
| **From the statistical analysis:** |
| **(A,B,G) > (C,E,F) > D** |
| **From comparison between candidates and selected CHAMP data** |
| **(B,F,G,A) > (C,E) > D** |

## Summary IGRF Evaluation

- From power the spectra analyses, all IGRF candidates are statistically equivalent except candidate models D and E.
- The differences between Gauss coefficients and the mean model are systematic for D and E.
- Residual analysis between IGRF 2010 candidates calculated at epoch 2009.0 with the SV model of candidate F shows that candidates A, B, F and G fit the data well, D rather well, and C and E are biased (please interpret this test with caution since models have been reduced with an SV that is itself a candidate for IGRF!).

### *Suggestion*

| |
|---|
| **(A,B,F,G) > (D,C,E)** |

## Summary SV Evaluation

- Power spectra analyses suggest that F and D are OK; A, B, C and E are acceptable while G is different from other models.
- The Gauss coefficients difference to the mean shows that G and A have systematic biases.
- When SV candidates are used with IGRF candidates to predict the field at epoch 2009.0 that is then compared with selected CHAMP data: B,E and F are OK; A and D show a ‘random’ like distribution (slightly skewed) but not centred on 0; they seem acceptable. C and G are skewed, have larger residuals and are not centred on 0.

### *Suggestion*

| |
|---|
| **(E,F) > (B,A,C) > (D,G)** |

# 1- DGRF:

## *1-1 Power spectra*

An average DGRF model is calculated (arithmetic mean of Gauss coefficients). The power spectrum of the between this DGRF and each DGRF candidate model is shown in Figure 1.

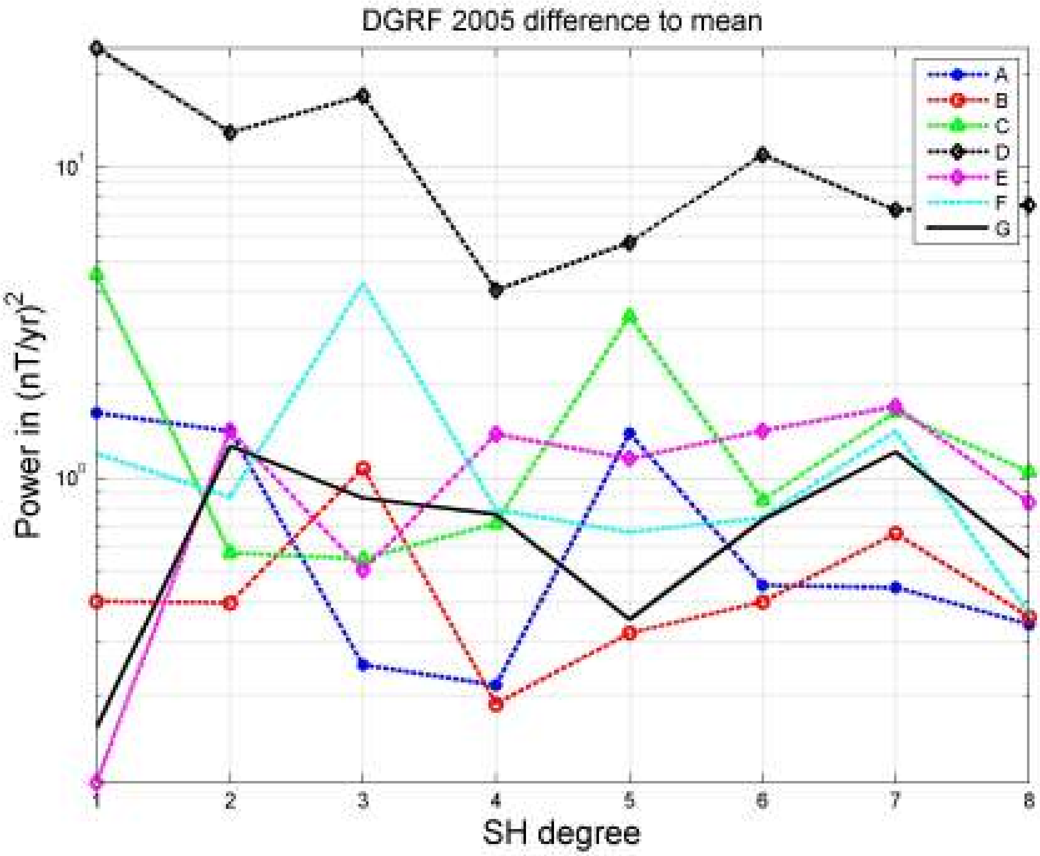



**Figure 1: Power spectra of the difference between the average model and each DGRF candidate.**

With this criterion, the candidate model D may be considered as an outlier. RMS between the average mode and each candidate are:

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| RMS in nT | 3.1 | 2.6 | 4.6 | 12.4 | 4.5 | 4.0 | 2.9 |

## *1-2 Gauss coefficients*

Figure 2 shows the result of a test made following S. Maus evaluation for IGRF-10. It shows the difference to the mean between gauss coefficients. Except for model D, candidates F, C, E may look peaky sometimes but the difference never exceeds 1 nT.

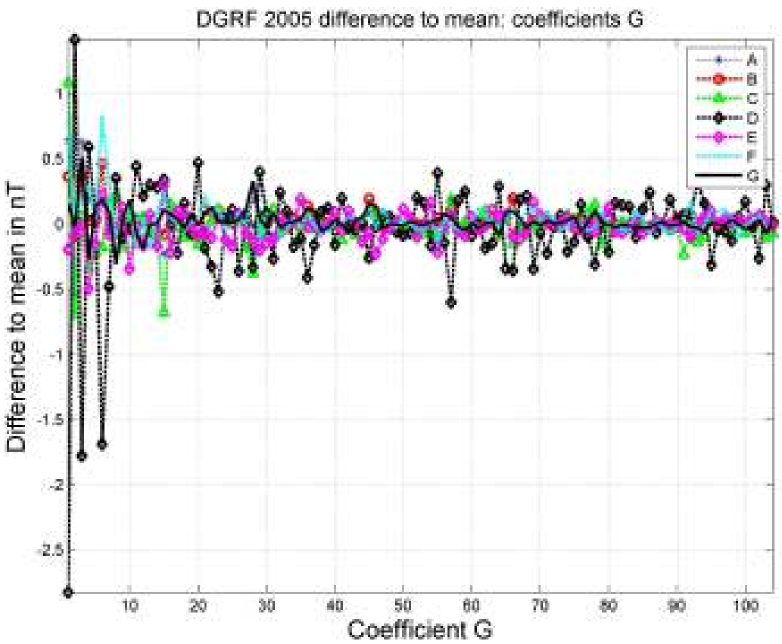



**Figure 2: Difference for each Gauss coefficient in nT. Gnm coefficients.**

The next figure is the same plot but compared to the average model calculated without candidate D. The scatter is lower for all models.

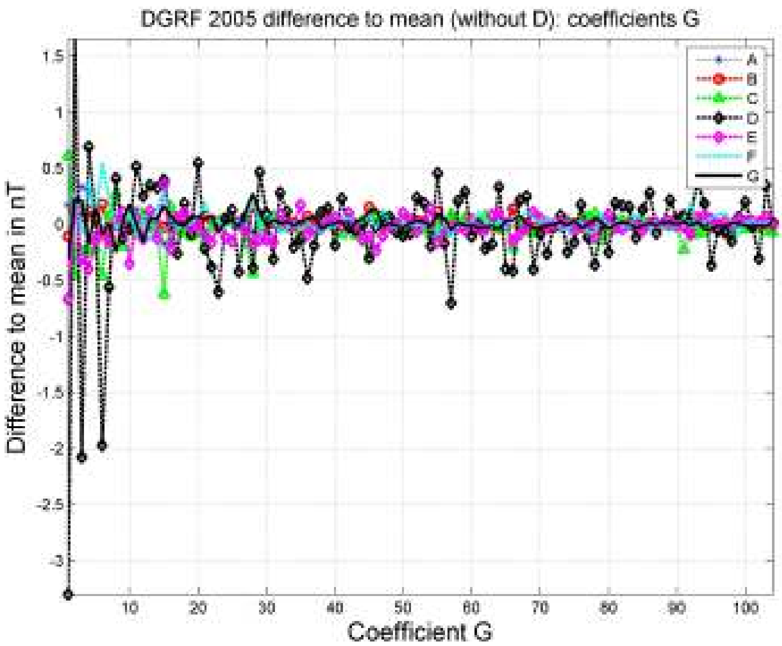



**Figure 3: Same as Figure 2 but without candidate D.**

The difference to the mean is generally lower than 0.5 nT (absolute value). When deriving the IPGP/EOST/LPG/LATMOS candidate model (F) we realized that differences in external field parameterization would easily generate differences of this order of magnitude.

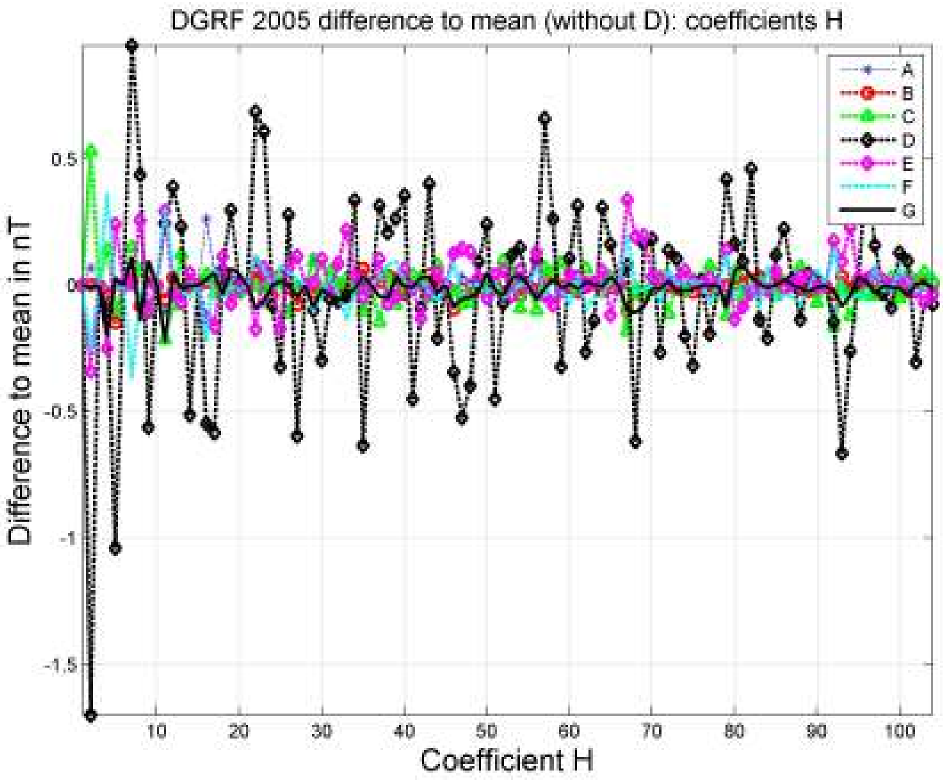



**Figure 4: Same as Figure 3 for coefficients Hnm.**

When model D is discarded in the mean model calculation, RMS are:

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| RMS in nT | 2.0 | 1.7 | 4.1 | 14.4 | 4.2 | 3.4 | 2.4 |

**According to this test, the candidate model D could/should be down-weighted. All other models seem almost statistically equivalent.**

## *1-3 Comparison with an "independent" data set*

We select CHAMP vector and scalar data within a time window covering +/-240 days around 2005.0. The data where corrected following the step described in the notice of model F. The dataset is considered independent from the one used to derive candidates A, C, D, E and G. The SV and external coefficients of model F is considered for correcting for external field and secular variation. It is indeed assumed sufficiently correct for the considered time interval (The F SV candidate model was obtained with the same time interval).

We compare RMS for scalar and vector data independently and for Polar and Mid-latitude regions (Table 1). Model D has strong errors and biases.

**Table 1: RMS between candidate models and selected CHAMP data in 2005.0**

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| δF | -1,9 | -1,8 | -2,3 | -7,6 | -2,7 | -0,8 | -2,2 |
| RMS F [nT] | 6,9 | 6,5 | 8,4 | 13,0 | 7,4 | 6,1 | 6,8 |
| δF \|lat > 50°\| | -3,8 | -3,1 | -5,1 | -12,6 | -4,3 | -1,0 | -3,8 |
| RMS F \|lat > | 8,5 | 7,8 | 10,7 | 16,8 | 9,1 | 7,0 | 8,3 |
| δF \|lat < 50°\| | 0,0 | -0,5 | 0,7 | -2,3 | -0,9 | -0,5 | -0,5 |
| RMS F \|lat < | 4,7 | 4,7 | 4,9 | 6,8 | 5,0 | 4,8 | 4,7 |
| RMS B [nT] | 4,9 | 4,9 | 5,0 | 6,7 | 5,2 | 5,1 | 4,9 |
| δB_r | -0,4 | -0,5 | -0,6 | -0,9 | -0,8 | -0,7 | -0,4 |
| RMS B_r | 5,1 | 5,1 | 5,3 | 8,0 | 5,6 | 5,4 | 5,2 |
| δB_θ | -0,2 | 0,1 | -0,5 | 2,8 | 0,6 | -0,1 | 0,3 |
| RMS B_θ | 5,5 | 5,5 | 5,4 | 7,1 | 5,8 | 5,7 | 5,5 |
| δB_ϕ | 0,2 | 0,2 | 0,2 | 0,2 | 0,3 | 0,2 | 0,3 |

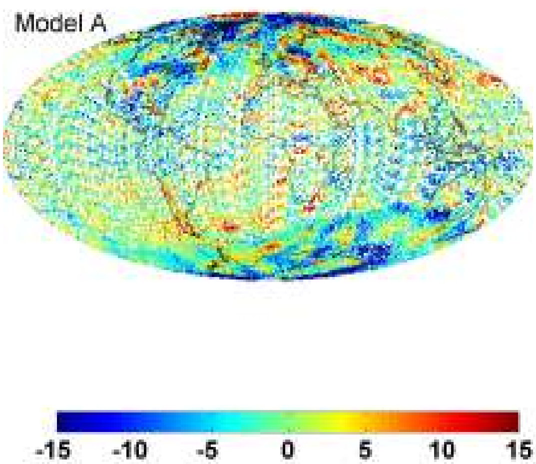



**Figure 5: Residuals between a selection of CHAMP data and model A for F in nT.**

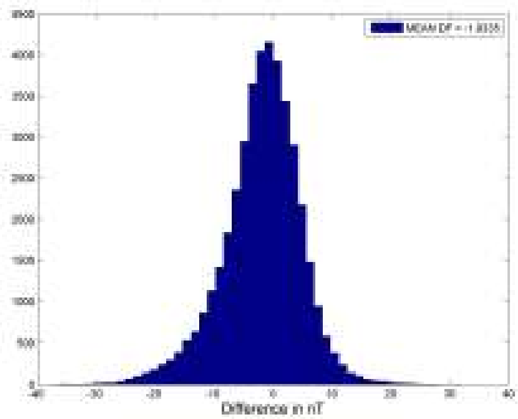



**Figure 6: Histrogram of residuals between model A and selected CHAMP DATA.**

The histogram seems OK (also for vector components) even though is **very slightly skewed** towards negative values. Some high values are visible near the frontier +/-55° in the residual map. This is the overlapping area between mid and Polar Regions that bore different data selection and correction. This feature will be often observed.

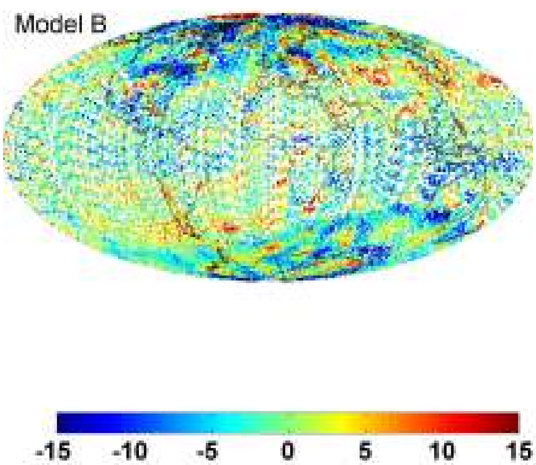



**Figure 7: Same as figure 5 for model B.**

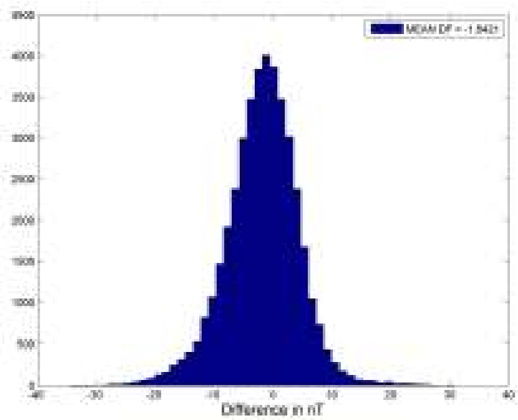



**Figure 8: Same as Figure 6 for model B.**

**Remark:** The histogram seems OK (also for vector components) even though **one must keep in mind that the selected CHAMP data were partly used for deriving model B**.

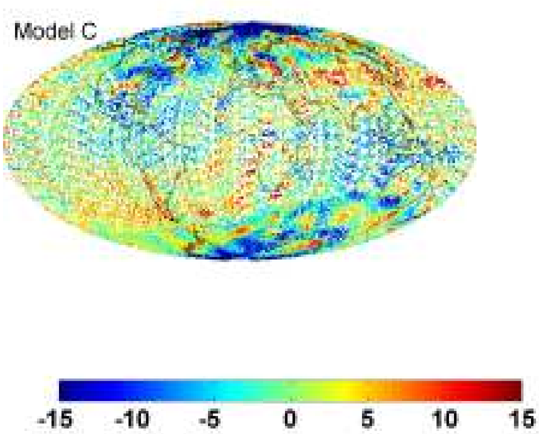



**Figure 9: Same as figure 5 for model C.**

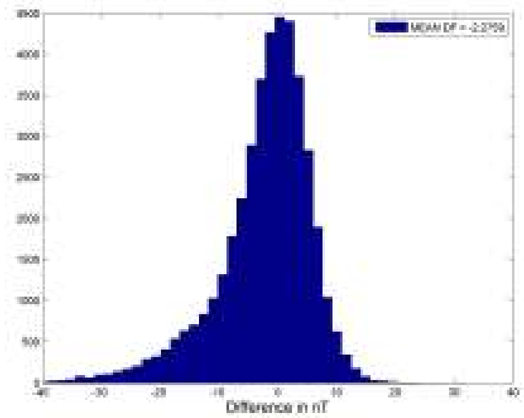



**Figure 10: Same as Figure 6 for model C.**

**Remark:** The histogram seems OK (also for vector components) even though it is **skewed** towards negative values.

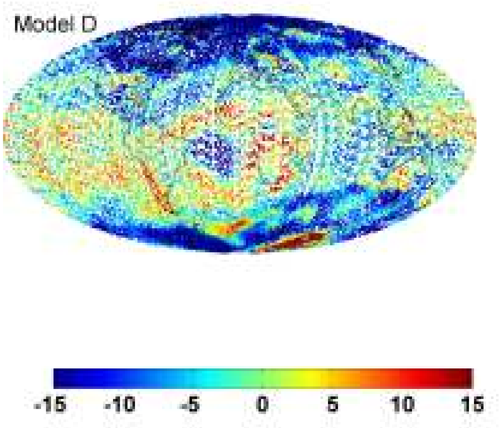



**Figure 11: Same as figure 5 for model D.**

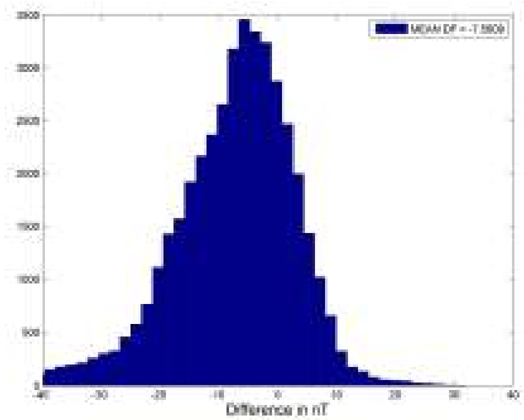



**Figure 12: Same as Figure 6 for model D.**

**Remark:** The histogram is problematic. It is **skewed towards negative values and has a wrong mean**. There is a systematic bias as can be seen on Figure 11.

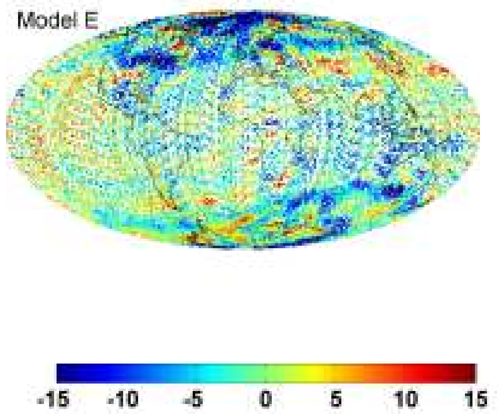



**Figure 13: Same as figure 5 for model E.**

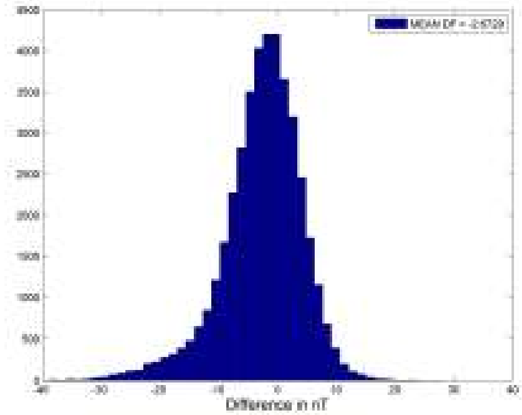



**Figure 14: Same as Figure 6 for model E.**

The histogram seems OK even though is **slightly skewed** towards negative values.

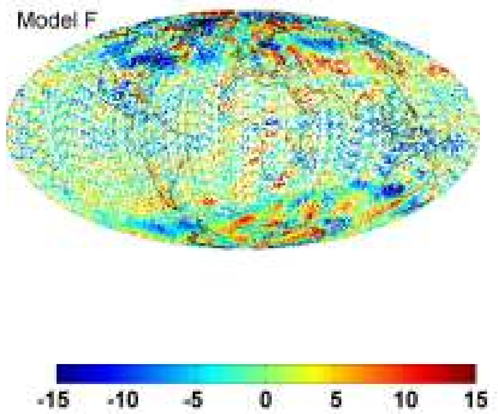



**Figure 15: Same as figure 5 for model F.**

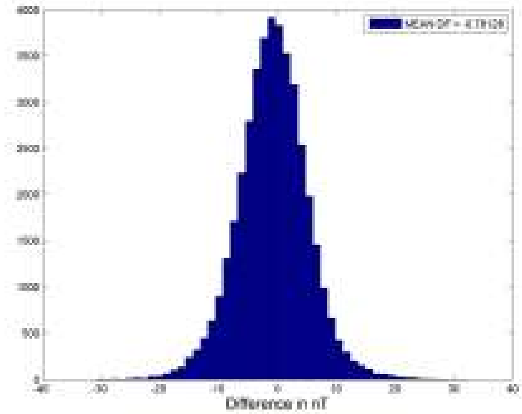



**Figure 16: Same as Figure 6 for model F.**

The histogram is OK but **the model is not tested against completely independent data**.

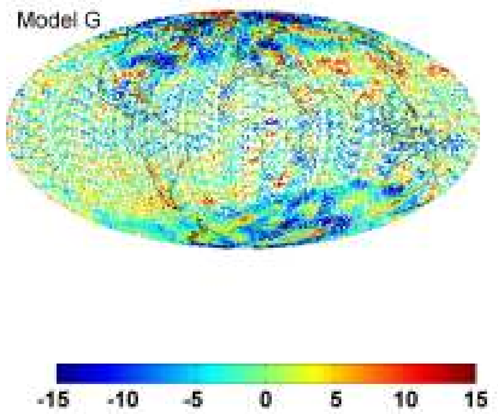



**Figure 17: Same as figure 5 for model F.**

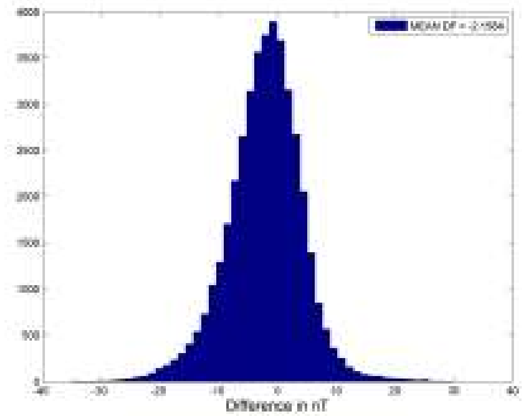



**Figure 18: Same as Figure 6 for model G.**

The histogram seems OK (also for vector components).

## *1-4 Concluding remarks of section 1*

- The statistic comparison between candidates and the average DGRF model showed that model **D has a problem**.
- The statistical comparisons between Gauss coefficients of candidates and the average DGRF model showed that A, B and G have better RMS than C, E, and F (D being discarded).
- Comparisons between candidate and independent data (for A, C, D, E and G) show that all models but D have the correct level of RMS and mean. They may be considered as statistically equivalent.
- However, histograms of residuals showed some differences. In particular, B, F and G are correct, A is very slightly skewed, C and E are skewed. Note again, that F was derived with the same dataset and B probably with a subset of it.
- The skewness may be due to small external field contamination in DGRF candidates C and E, providing, of course that the subset of considered data is itself well corrected for external fields.

**From the statistical analysis:**
**(A,B,G,) > (C,E,F) > D**
**From the comparison with selected CHAMP data**
**(B,F,G,A) > (C,E) > D**

**However, selected CHAMP data and candidate F are not independent so it expected that B and F fits the data well. We may thus prefer the first solution: (A,B,G,) > (C,E,F) > D or simply (B,F,G,A,C,E) > D**

# 2- IGRF candidate models

## *2-1 Power spectra*

The same tests are carried out. The power-spectrum of the difference between the mean IGRF and each IGRF candidate model is plotted in Figure 19. The RMS to the mean are:

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| RMS in nT | 6.3 | 3.8 | 7.1 | 12.3 | 10.9 | 5.8 | 4.6 |

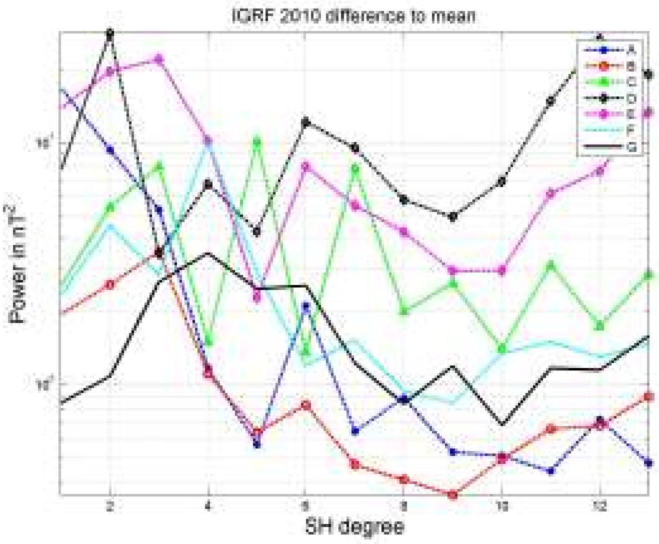



**Figure 19: Same as Figure 1 but considering IGRF Gauss coefficient.**

## *2-2 Gauss coefficients*

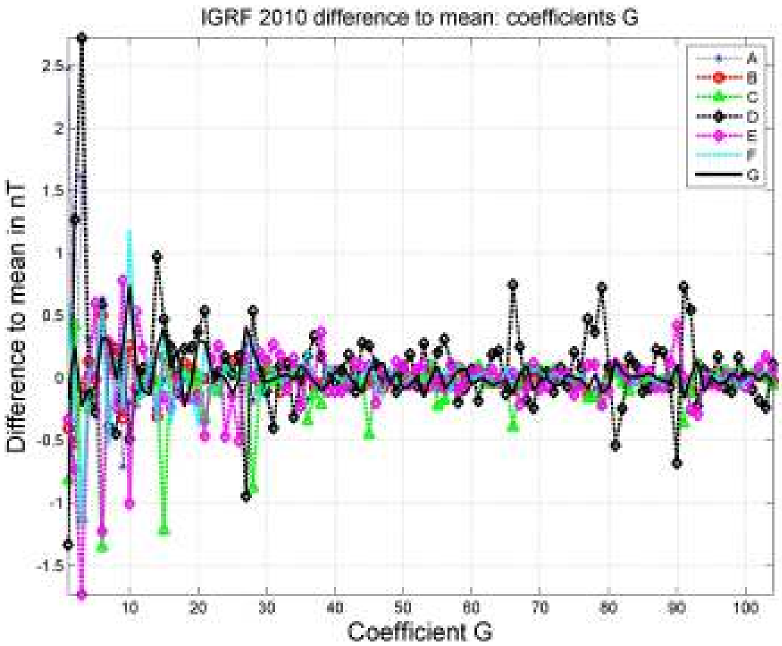



**Figure 20: Same as Figure 2 with all candidates included.**

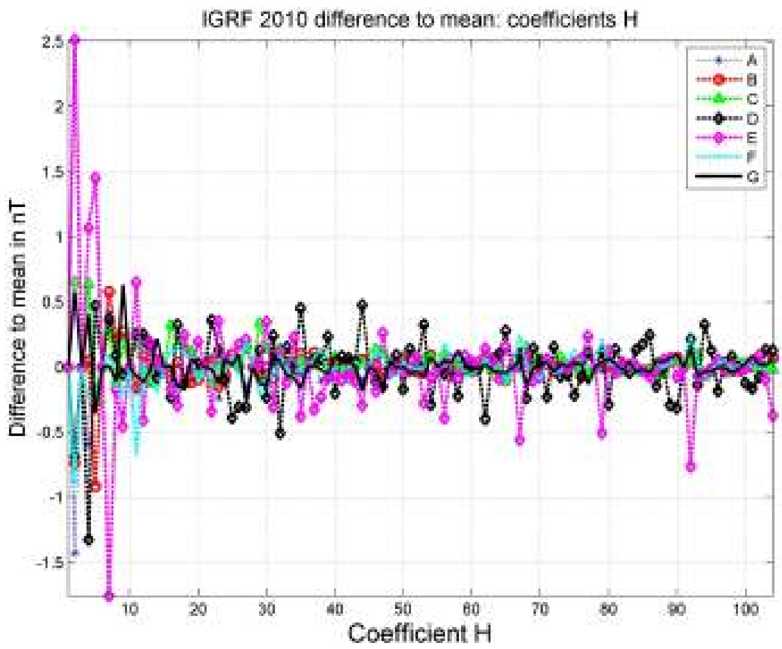



**Figure 21: Same as Figure 3 with all candidates included.**

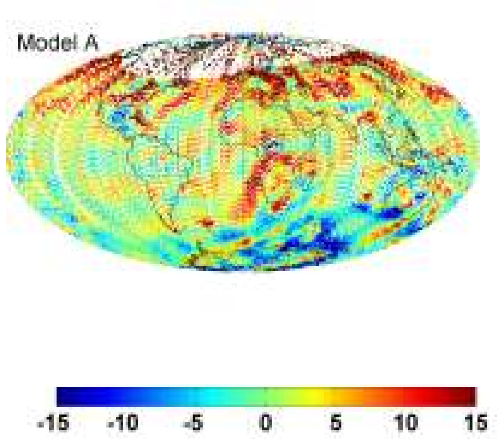
Gauss coefficient differences for models D and E have a large scatter and seem to have a systematic difference: the sign is not zero. Model D, for instance, has more positive differences (for coefficients G). For E, differences are mainly positive for G and negative for H at low SH degree.

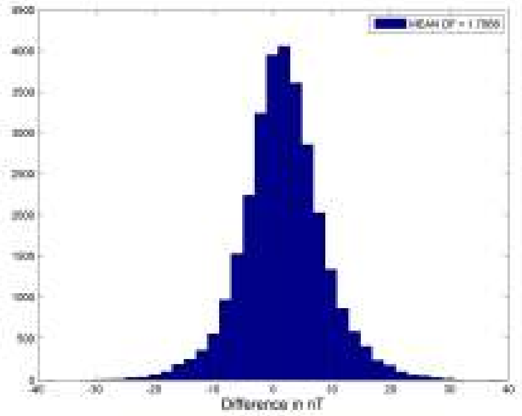
## *2-3 Comparison with independent data*

Again, not all models are completely independent from the considered data **(the test may be biased towards models B and F**). The selected data are centred at 2009.0 +/-240 days. The external Gauss coefficients are those obtained for candidate model F. SV model of candidate F is used because according to the various accompanying explanation SV candidates are not necessarily derived from one parent model and may be averaged over a longer time interval. **This test is thus subjective but model F is average over the same time interval, is valid for epoch 2009.0, which is not the case for some other candidate models.**

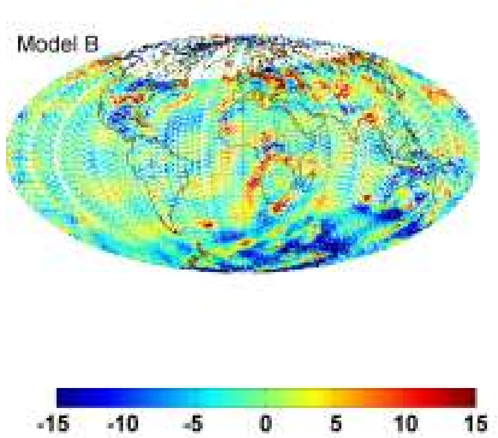
**Table 2: Statistics between candidates and selected CHAMP data in 2009.0**

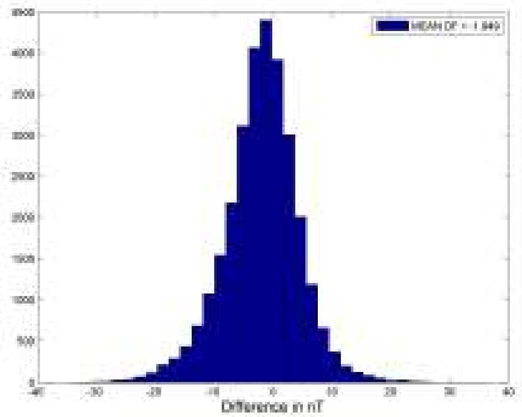
| | A | B | C2 | D | E | F | G |
|---|---|---|---|---|---|---|---|
| δF | 1,7 | -1,9 | -4,4 | -2,6 | -2,7 | -0,7 | -1,8 |
| RMS F [nT] | 7,3 | 6,8 | 11,5 | 9,0 | 9,8 | 6,6 | 6,8 |
| δF \|lat > 50°\| | 0,7 | -3,6 | -9,4 | -3,6 | -4,7 | -1,7 | -3,1 |
| RMS F \|lat > | 8,6 | 8,4 | 16,2 | 11,1 | 13,2 | 7,8 | 8,4 |
| δF \|lat < 50°\| | 2,6 | -0,6 | -0,2 | -1,7 | -1,0 | 0,1 | -0,7 |
| RMS F \|lat < | 6,0 | 5,1 | 5,0 | 6,8 | 5,7 | 5,3 | 5,1 |
| RMS B [nT] | 5,7 | 5,1 | 4,8 | 7,0 | 6,2 | 5,1 | 5,1 |
| δB_r | 0,8 | 0,5 | -0,2 | 0,7 | -0,8 | 0,7 | 0,4 |
| RMS B_r | 6,3 | 5,6 | 5,3 | 8,2 | 7,3 | 5,7 | 5,7 |
| δB_θ | -1,9 | 0,3 | 0,6 | 1,1 | 0,5 | -0,3 | 0,3 |
| RMS B_θ | 6,4 | 5,6 | 5,2 | 7,3 | 6,1 | 5,5 | 5,4 |
| δB_ϕ | 0,0 | 0,0 | 0,1 | 0,2 | 0,0 | 0,0 | 0,1 |



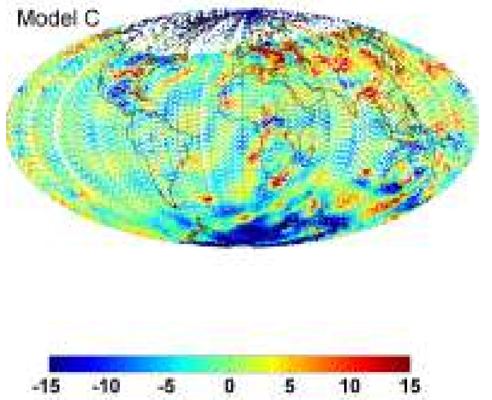



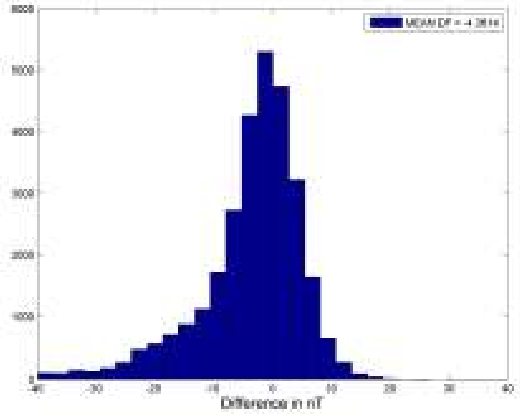

Model A is OK. The histogram looks good but maximum towards positive values (~1 nT). This may be an effect of the considered SV not fully applicable to IGRF candidate A.



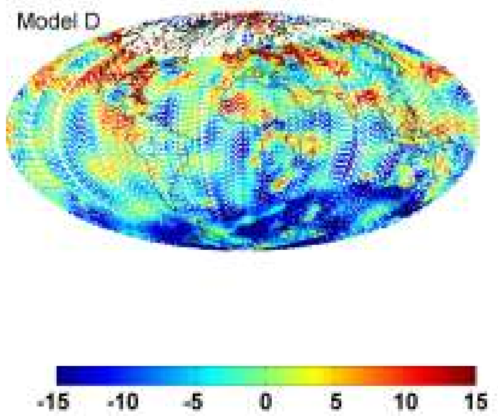



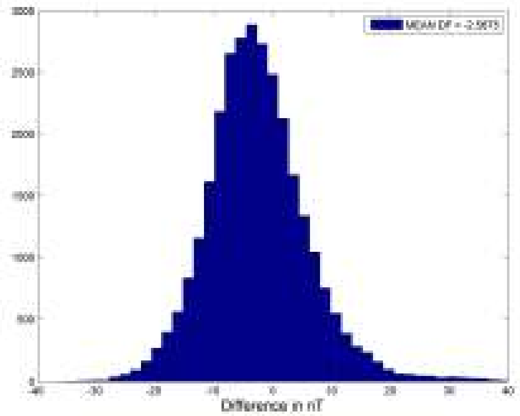

Model B is OK. The histogram looks good and centred near 0.



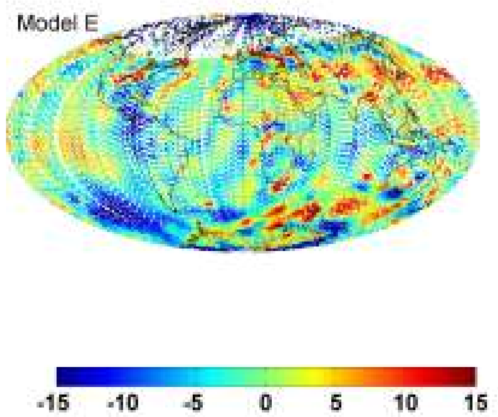



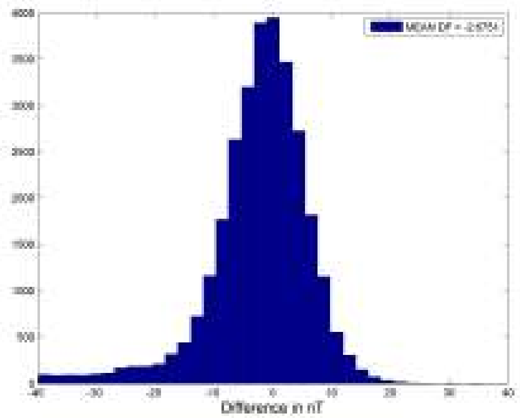

RMS are high for F. The histogram of F residuals is skewed and its mean not 0.



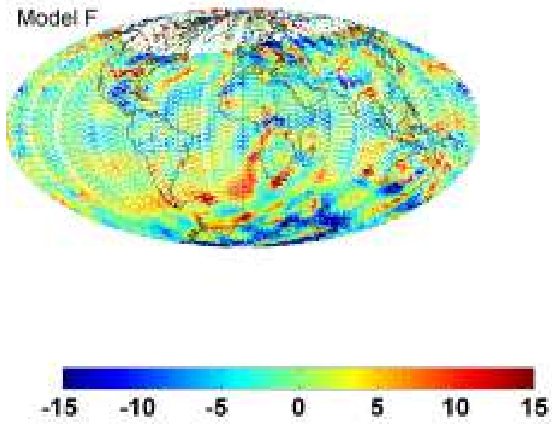



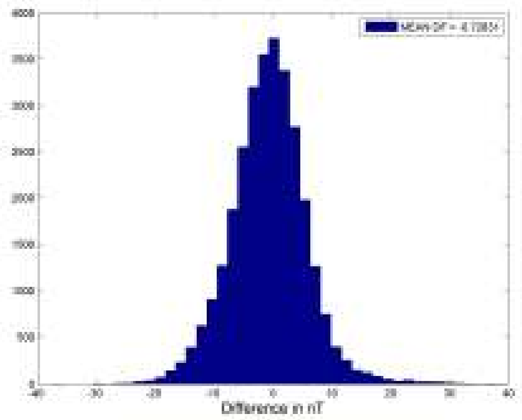

The histogram is acceptable but slightly skewed toward positive values. The mean is shifted towards negative values. Residuals are large in the Southern hemisphere.



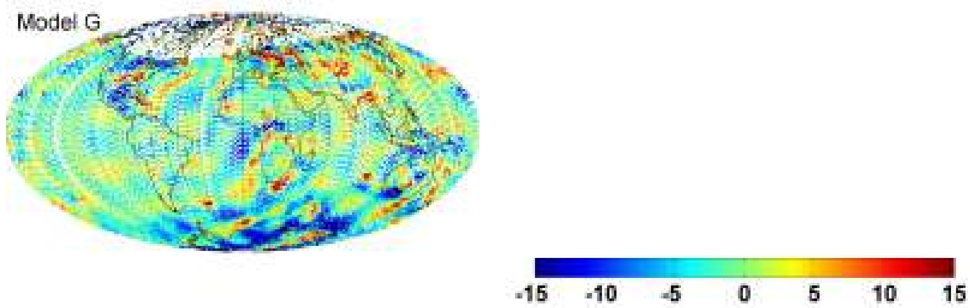



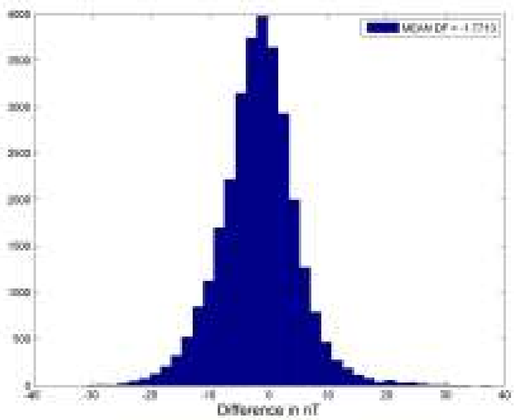

The histogram is narrower than for model D but slightly skewed toward negative values. The mean is shifted towards negative values. A large negative patch is observed in the Pacific.




OK but large positive stripe crossing the southern atlantic ocean.




OK also when considering all criteria (mean and shape of histrograms on F, Br, Btheta and Bphi).

## *2-4 Concluding remarks of section 2*

- From power spectra analyses, all IGRF candidates are statistically equivalent but candidate D and E.

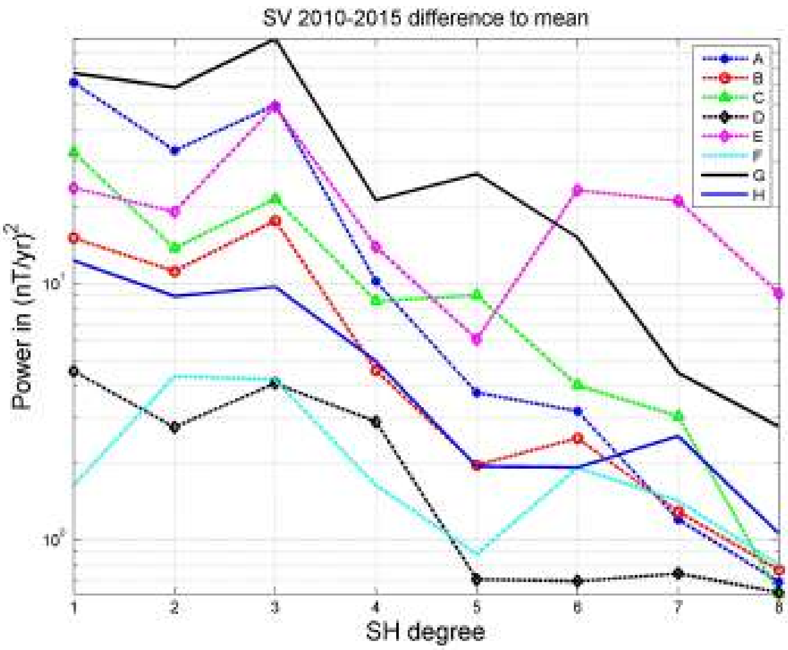
- The differences between Gauss coefficients and the mean IGRF model suggest that D and E have indeed systematic differences.
- The residual analysis between IGRF 2010 candidates calculated at epoch 2009.0 with the SV model of candidate F show that: Candidates A, B, F and G fit the data well, D rather well, and C and E are biased (**please interpret this test with caution**).

According to these **non-exhaustive** criteria:

| **(A,B,F,G) > (D,C,E)** |
|---|

# 3- Secular Variation models 2010-2015

## *3-1 Power spectra*




The RMS vector difference (nT/yr) between models and the arithmetic mean is:

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| RMS in nT/yr | 12.8 | 7.4 | 9.7 | 4.1 | 12.9 | 4.1 | 16.9 | 6.6 |

Removing candidate G provides:

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| RMS in nT/yr | 13.0 | 7.9 | 8.4 | 4.4 | 11.7 | 4.4 | 19.3 | 5.9 |

Following this criterion, candidates A, G and E are not in very good agreement with the arithmetic mean.

## *3-2 Gauss coefficients*

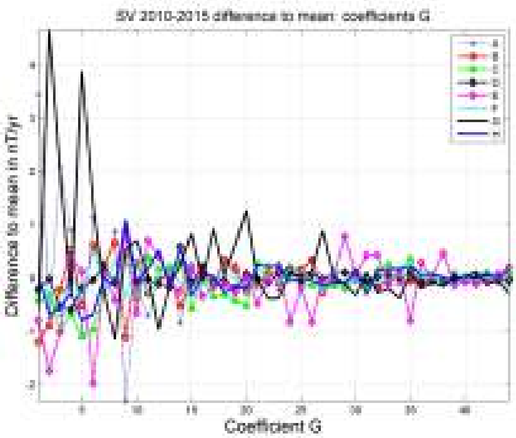



**Figure 22: Same as Figure 20 for SV candidates.**

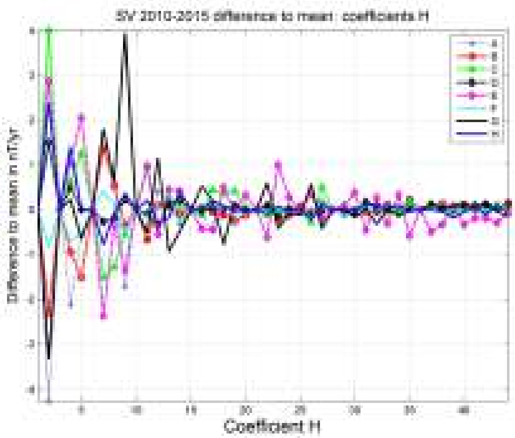



**Figure 23: Same as Figure 21 for SV candidates.**

Model A and G have a systematic error for coefficients G and H at low degree (mostly positive for G, negative then positive for H).

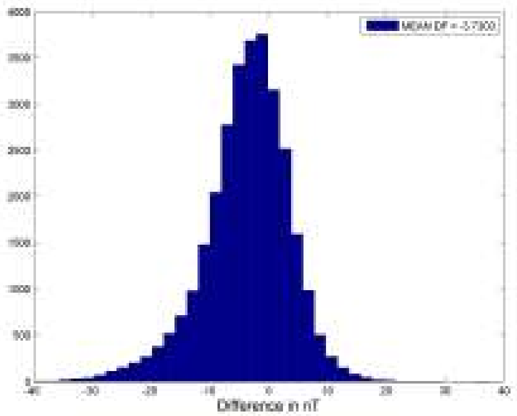
## *3-3 Extrapolation to epoch 2009.0*

We now use the same data set but IGRF and SV candidates are now considered. We assume that SV candidates allow some hind cast, at least up to 2009.0. The difference between the selected CHAMP data and the reduced candidate models are given in Table 3.

**Table 3: RMS between IGRF calculated at 2009.0 and selected CHAMP data. SV candidates are considered for reducing IGRF candidates to 2009.0**

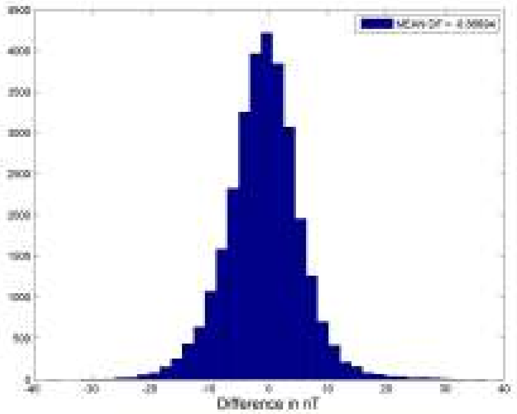
| | A | B | C | C2 | D | E | F | G |
|---|---|---|---|---|---|---|---|---|
| δF | -3,7 | -0,9 | 1,2 | -2.7 | -2,2 | -0,6 | -0,7 | -2,3 |
| **RMS F [nT]** | **8,2** | **6,5** | 61,0 | **9.9** | **9,8** | **8,4** | **6,6** | **10,0** |
| δF \|lat > 50°\| | -5,6 | -1,8 | 5,7 | -6.3 | -3,2 | -0,8 | -1,7 | -5,1 |
| RMS F \|lat > 50°\| | 10,3 | 7,8 | 80,4 | 13.0 | 12,3 | 10,5 | 7,8 | 10,8 |
| δF \|lat < 50°\| | -2,2 | -0,1 | -2,4 | 0.3 | -1,5 | -0,4 | 0,1 | 0,0 |
| RMS F \|lat < 50°\| | 6,1 | 5,0 | 38,2 | 6.2 | 7,1 | 6,2 | 5,3 | 9,4 |
| **RMS B [nT]** | **5,7** | **4,9** | 35,9 | **7.0** | **7,6** | **6,2** | **5,1** | **9,5** |
| δB_r | 0,0 | 0,2 | -0,8 | 0.5 | 1,0 | 0,2 | 0,7 | 0,1 |
| RMS B_r | 6,4 | 5,4 | 45,3 | 8.5 | 9,1 | 7,2 | 5,7 | 12,2 |
| δB_θ | 1,0 | -0,2 | 1,5 | 0.0 | 0,9 | -0,6 | -0,3 | 0,4 |
| RMS B_θ | 6,0 | 5,3 | 29,2 | 6.9 | 7,7 | 6,6 | 5,5 | 7,3 |
| δB_ϕ | 0,1 | 0,1 | 0,2 | 0.1 | 0,0 | 0,2 | 0,0 | -0,1 |

## Model A



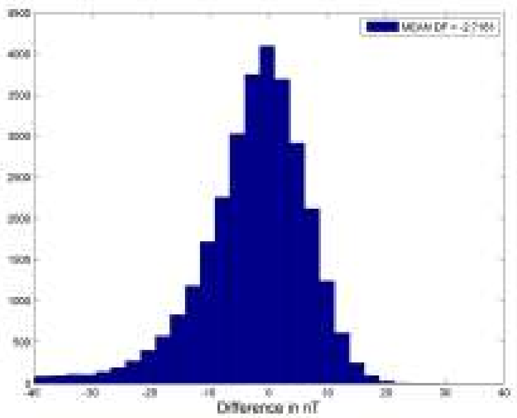

Histogram is skewed with a larger amount of negative residuals. It is not centred around 0.

## Model B



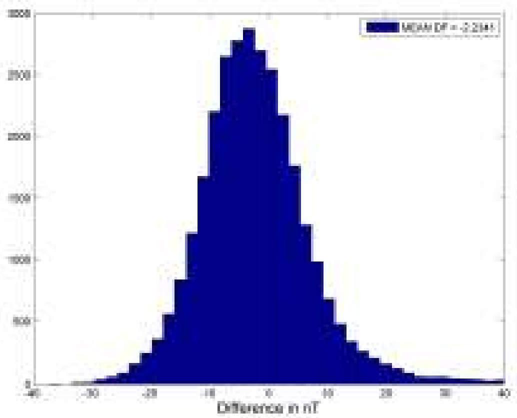

Good histogram with reasonable mean and standard deviation.

## Model C



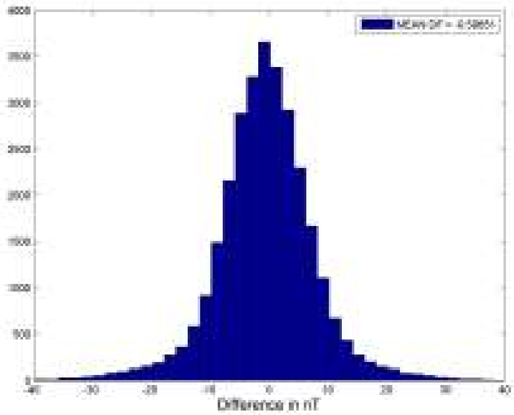

The histrogram is skewed. A large amount o negative residuals are seen over Antarctica.

## Model D



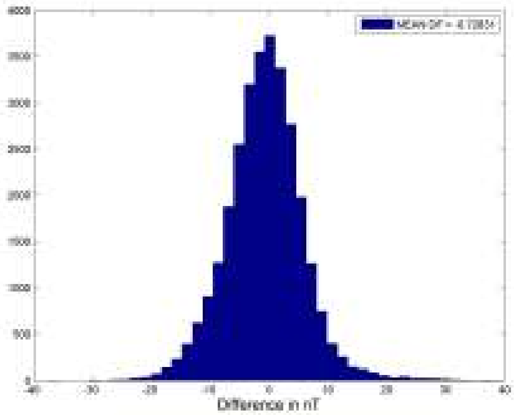

Same characteristics as for C but residuals are located over both polar caps.

## Model E



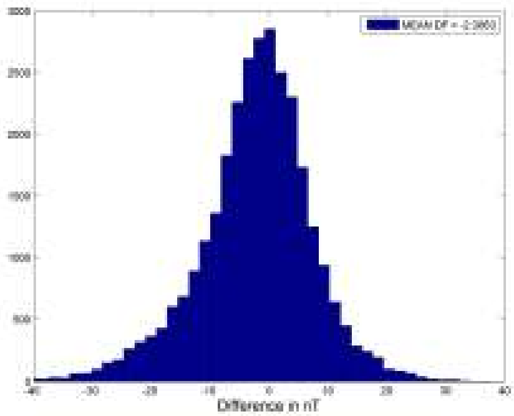

Model E seems OK.

## Model F




Model F seems OK.

## Model G




The histogram is skewed towards negative values. The mean is significantly different from 0.

**Model H:** unavailable without IGRF candidate – See comparisons of A. Chulliat.

## 3-4 Concluding remarks of section 3

- SV Power spectra analyses suggest that F and D are OK; A, B, C and E are acceptable while G is a bit different from other models.
- Gauss coefficient difference to the mean show the G and A have systematic biases compared to the mean SV model.
- When SV candidates are used to predict IGRF candidate to epoch 2009.0, B,E and F are OK. A, C, D and G have different characteristics: A and D shows a 'random' like distribution (slightly skewed) but not centred on zero; they seem acceptable. C and G are skewed, have larger residuals and are not centred on zero.

According to these **non-exhaustive** criteria:

**(E,F) > (B,A,C) > (D,G)**

**E. Thébault**

**End of document**