

Author response to referee #1

The authors thank Anonymous referee #1 for their detailed comments on the manuscript. We will take the comments into account when revising the manuscript. In this document we provide responses to each of the referee's comments (formatted as italics in indented paragraphs and grouped together when appropriate). At the end of our answers, we indicate in blue where the changes have been made in the revised version.

Line 10: The error comparison for different regions is as such not useful in an abstract. It does not compare to achievable quality in the regular case (all axes working) and it does not relate the error to the measured fields (percentual) – e.g. a deviation of 2 nT in solar wind has a different impact as 12 nT in the high field magnetosphere. Provide more information or less.

We agree and removed that part from the abstract and expanded the discussion section in the revised manuscript accordingly.

In the revised version we changed the abstract in lines 7 to 12 and expanded the discussion in the section Comparison of Results in lines 279 to 303.

Line 25: Please state why this deviation occurs (mechanical alignment/boom etc.)

According to Auster et al. [2008], there is no alignment requirement between the instrument and the spin axis. Fig. 1 shows the evolution of the angles α_{S1} and α_{S2} throughout the mission, demonstrating that they consistently deviate from 90° . We added a short explanation in the revised manuscript.

In the revised version we added a sentence in lines 24 to 25.

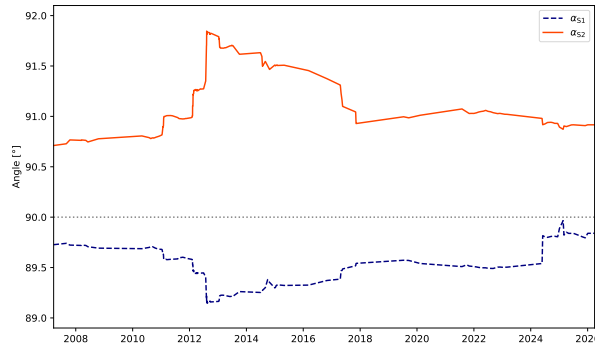


Figure 1: Angles α_{S1} and α_{S2} between the spin plane sensors and the spin axis over the course of the mission.

Line 50:

- *The assumption for accurate spin plane offsets is unclear. The calculation of spin plane offsets requires precise elevation angles, otherwise a constant contribution of the spin axis field (to be expected over many field regions) could be mistaken for spin axis offsets. Precise elevation angles in turn require knowledge of the spin axis field, but this field is in turn recovered by the method below with considerable error.*
- *There is a danger of a circular dependency, so a simple assumption should not be used without analysis (or presented without justification, if the analysis was done, but is not included)*
- *It is clear that there will be an error and that the situation is not perfect. Nevertheless, the assumption without further information is not enough. An IGRF-comparison could provide an estimate (i.e. rather than correlating spin axis variations to spin plane variations, one could correlate IGRF variations to spin plane to find thetas.*
- *It might be that the variations in theta are small enough to be irrelevant, but this should be brought to support the initial claim of spin plane offsets.*

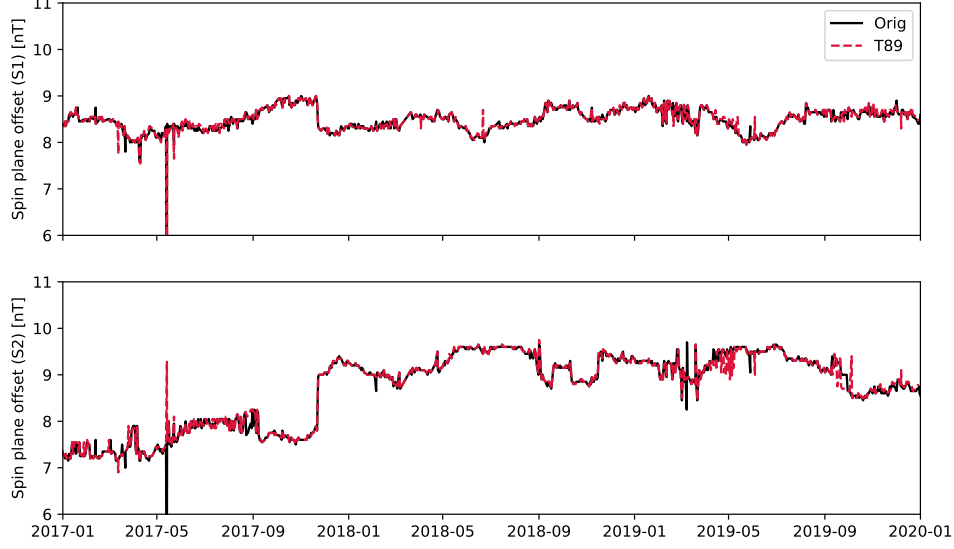


Figure 2: Comparison of spin plane offsets calculated using the standard calibration procedure (black) with those calculated using the T89 model field (red) as a spin axis sensor measurement exemplary from 2017 to 2020.

We agree that there is the danger of circular dependency and that this should not be phrased as it is in the current version. In lines 230 to 243 in the section Discussion of Limitations we address this point, although without analysis. We must emphasize that T89+IGRF (or another model field) is required to obtain an estimate of the spin plane offsets. To demonstrate that these estimates are reliable, we compared the spin plane offsets calculated using the standard calibration procedure with those calculated using the model field as a spin axis sensor measurement. This is shown in Fig. 2 It is clearly visible that the offsets match really well and the median deviations are close to 0 nT for both spin plane components. Therefore we argue that we can in future assume accurate spin plane offsets.

In the revised version we added a discussion of the accuracy to the Discussion of Limitions Section in lines 317 to 321.

Line 51: The angle alpha is not very clear for the outside world. A drawing would help.

Line 132 and 140: The factor alpha has not been explained well, see above.

We agree and have therefore expanded Fig. 1 by adding a third panel illustrating the angles $\alpha_{S1,S2}$ between spin axis and spin plane sensor axes $S1$ and $S2$ (see Fig. 3).

In the revised version we updated Fig. 1. and added a reference to this Figure in lines 50, 95, and 170.

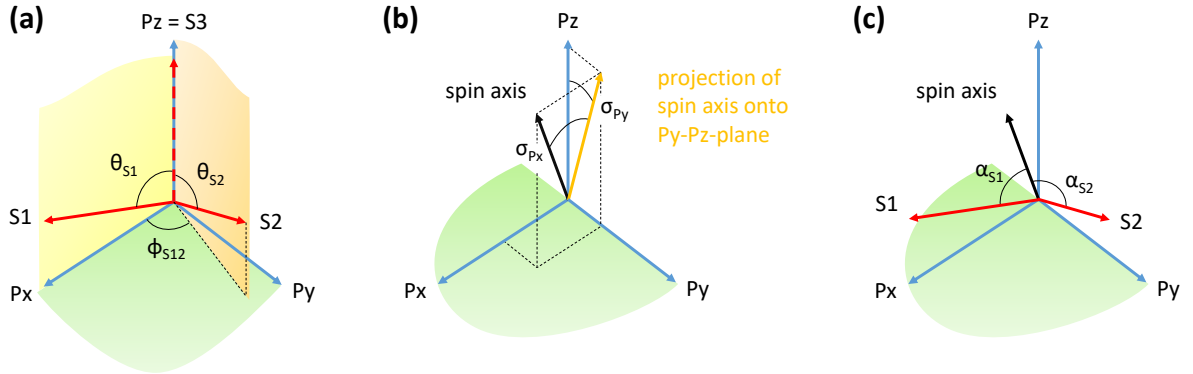


Figure 3: Updated sketch of sensor coordinate system. The panel c) shows the angles α_{S1} and α_{S2} between spin axis and spin plane sensors $S1$ and $S2$, respectively.

Line 53: The tradeoff is unclear. 2017 is not “close”. This needs more precise justification. No big problem, just needs to be done.

We agree that a justification is necessary. Since 2021, we have occasionally observed disturbances in the magnetic field data that fall outside the scientifically meaningful range but within the calibration range. In the first submission, the year 2017 was chosen somewhat arbitrarily as a representative example to avoid these anomalies. In the revised version, we therefore expanded the validation to a three-year interval from 2017 to 2020. The interval was chosen as a compromise between avoiding these disturbed data, maintaining reasonable temporal proximity to the current mission phase, and capturing variations in the calibration parameters (see Fig. 4). Repeating the analysis over this full interval shows no

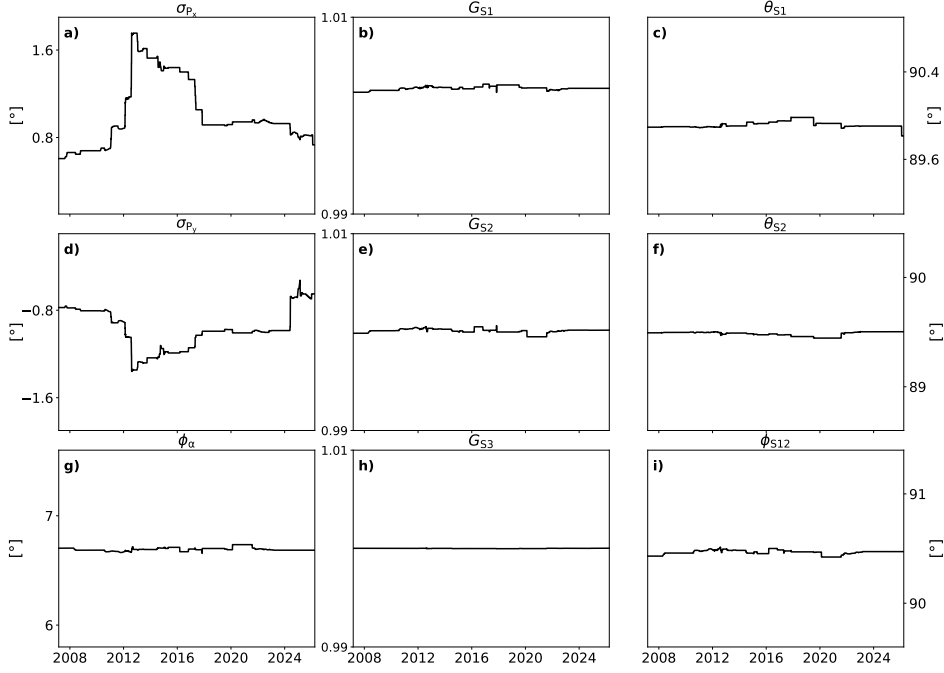


Figure 4: Evolution of calibration parameter: σ_{px} , σ_{py} , and ϕ_α are the rotation angles around the orthogonalized sensor axes (c.f. with blue coordinate system Fig. 1 in the manuscript or Fig. 3 in this reponse). G_{S1} , G_{S2} , and G_{S3} are the gains along the sensor axes. θ_{S1} , θ_{S1} , and ϕ_{S12} are the internal sensor angles as described in Fig. 1 in the manuscript.

significant differences between the years (see Figs. 10 for 2017, 11 for 2018, and 12 for 2019 at the end of the manuscript). We have revised the manuscript accordingly and now justify the chosen time range explicitly.

In the revised version we added a justification in lines 56 to 58 and updated the histograms with data from 2017 to 2020.

Line 85: Provide a citation for constant sensor-internal orthogonality. The statement is true enough, but should be backed up. Past calibration values could help.

We agree and therefore calculated the angles between the sensor axes θ_{S1} , θ_{S1} , and ϕ_{S12} . In Fig. 4 we present the evolution of all calibration parameter over the whole mission time. All parameter, beside σ_{px} and σ_{py} , are very stable over the duration of the mission. σ_{px} , σ_{py} , and ϕ_α are the angles describing the orientation between the FGM and the spin axis direction. We added Fig. 4 to the revised manuscript. In the revised version we added Fig. 4 (Fig. 2 in revised version) and added reference in lines 89, 90, 92 and 101.

Line 95:

- The choice of the wording “spin tone” is very unfortunate. It probably originates from the fact that the contribution of BST would result in a spin tone if despun.

- *What is considered here is the “true” magnetic field in the spin axis (+ errors) that is amplified and used as spin axis field. Even equation 11 is considering a “spinning” part ($A+B$) and a non-spinning part ($C+D$), but the latter is then referred to as spin tone.*
- *In that sense the word “spin tone” is the name of what $A+B$ if they were demodulated, but they are not and used for something else.*
- *(My personal experience was that I first skipped the index ST , just had a look at the equation and things were quite clear. After that I checked what ST means and got confused and started to look for what I have misread, which was nothing)*
- *Please find a new term to make this more clear.*

We agree that the wording is unfortunate. We replaced the wording “spin tone” with “projected spin axis contribution” and renamed the variable from B_{ST} to B_{pSA} . This should make things clearer.

In the revised version we have updated the wording and variable name throughout the manuscript. We refer to the manuscript with track changes to avoid listing all the changes here.

Line 100:

- *The fit is not with respect to B_{raw} – B_{raw} is the desired result. The fit is done with respect to $A/B/C/D$*
- *It is also not clear whether the parameters w_s and ϕ are part of the variable parameters. They probably are, with some limitations on prior knowledge.*

We agree and changed the wording accordingly. We clearly indicate that A , B , C , D , and φ are free fitting parameter, whereas ω_s is provided by the state file of THE.

In the revised version we changed the section describing the fitting process in lines 104 to 123.

Line 104:

- *ϕ is not the phase; it is the angle of the external field that is projected onto the spin plane. This appears as a phase shift in the equation, but does not correspond to the usual term of phase shift, i.e. a fixed time or angle shift. It is a variable, data dependent shift. Please change wording to avoid confusion*
- *Line 104 uses ϕ , while eq. 11 uses only ϕ . Also the use of ω / ω s is inconsistent across equation and text.*

We agree and rewrote the sentence. We also fixed the issue of inconsistency.

In the revised version we changed the description of φ in line 117, renamed all φ_s to φ (115, 117, 118, and 120), and renamed all ω to ω_s (lines 111, 115, 118, and 121).

Line 119: The method for averaging was chosen in time units rather than samples. The impact and justification for this was not given. Samples would generate overcomplete/incomplete spins, while time generates a variable number of samples. The choice was not explained, but will have some (minor) impact.

We agree and compared the different methods. First, we used fixed time windows and disregarded any window containing less than 90% or more than 110% of the expected $2.5 \cdot T_{\text{spin}} \cdot f_{\text{FGL}}$ (≈ 166) samples. Second, we calculated the fits with windows that contain a fixed number of samples and rejected any window where the time difference between the first and last point exceeded $2.55 \cdot T_{\text{spin}}$. The differences of the recovered values of the two methods in Fig. 5a) and Fig. 5b) shows that the outcome of the fitting process is very similar (3 years of data from 2017 to 2020: Median around 0 nT and around 1 nT in width). Fig. 5c) illustrates that the number of samples per window in time units matches the constant sample number quite well, although there are some intervals with a few more points when using a constant time window.

In the revised manuscript we continue using the constant time window approach and discuss the minor discrepancy compared to using a fixed number of samples.

In the revised version we added the discussion/explanation in lines 135 to 141.

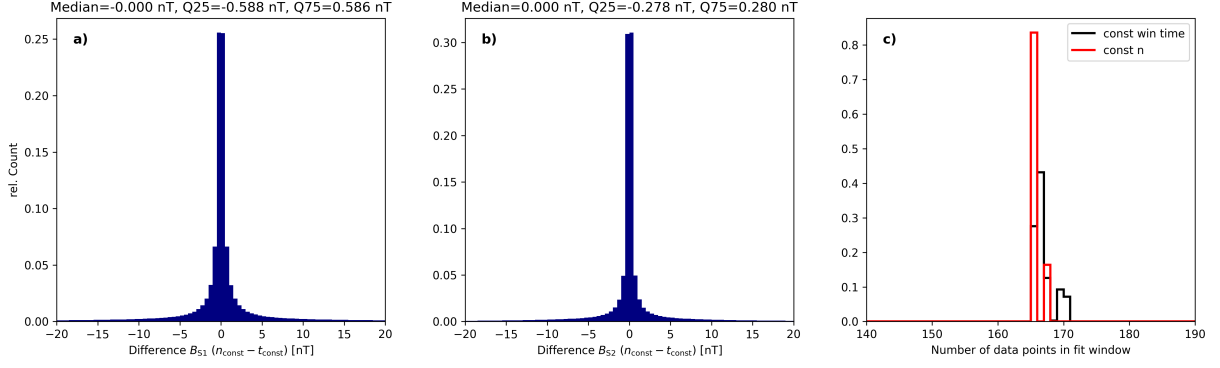


Figure 5: Comparison of recovered spin axis values for both spin plane sensor components B_{S1} (a) and B_{S2} (b) between fits with a fixed number of samples and fits with fixed window length. Panel c) shows the distribution of samples per fit window for fixed numbers in red and for fixed window length in black.

Method in 3.1

The method used in chapter 3.1 is not well explained. The authors are using the following measures to improve the quality of their result:

- Assume a first order least square fit for the magnetic field and considering everything else as error signal.
- Using the fit onto 2.5 spin periods, shifting by 1 spin period (overlap 40%).
- Using a weighted moving average over 2 minutes. Weight increases with constant part of the spin plane component and decreases with the error signal (i.e. true error and higher frequency contributions).

Nevertheless, the formulation of equation Eq. 13 is not clear. BST is defined as a function of t in Eq. 11. Equation 13 also uses $BST(t)$ but it is obvious that this is not the same definition – Here BST is averaged over 2 minutes, while it is only defined for 2.5 spins in Eq. 11.

I assume that the authors average the central values of the respective length-2.5 periods (i.e. $BST(1.25 \cdot t_{spin})$), but this is not stated anywhere. Please clarify.

Assuming equal weights (will be discussed later), the application of a 2-minute weighted average would result in a bandwidth reduction 1/120 Hz, i.e. even though one averaged value per spin is provided, the included bandwidth is much lower. While this corresponds to the abstract statement of providing one estimate per spin, this is still generating the wrong impression of a higher bandwidth. This fact is only stated at the very end (line 264). Nevertheless, it seems that also the authors do not consider this reduction of bandwidth, see comment to Figure 5.

We agree that we have to explain the method in more detail. As correctly pointed out, we used t in Eq. 11 and Eq. 13 although they are not the same. Therefore we replaced t with τ in Eq. 11 confined from 0 to $2.5 \cdot T_{spin}$. In a similar manner we changed Eq. 12. The referee is also right that we are using the central average $B_{pSA} = B_{pSA}(\tau = 1.25T_{spin})$. We expanded the description of the method and fixed the aforementioned inconsistencies. Regarding the bandwidth reduction: We agree that we have to communicate this more clearly. We added the remark in the abstract, in the method section and in the discussion.

We revised Section 3.1 in lines 129 to 166. We added remarks regarding the bandwidth in lines 8, 166, 299, and 350.

The other thing is the topic of weighted averages.

- The authors use a weight that decreases with the error signal. The error signal is the residual between fit and measured signal and contains (a) components with higher frequencies and (b) disturbance/noise and (c) low frequency signals that are not covered by the model of eq. 11 – the latter is the actual fit error the authors are looking for. The presence of higher frequencies is not an error, it is by choice of the bandwidth. The attempt to extract one spin “model” over 2.5 spins means that the actual bandwidth that is relevant

for the fit is clearly less than the raw Nyquist of 2 Hz in FGL. As a result, the resulting fit quality estimate $f_{\min}$ is not only giving the quality of the fit, but also the amount of higher-frequency contribution. The latter is not really an error, it is just an artificial residual that is generated by the choice of bandwidth. It would be more consequent (and it would most probably deliver better results) to reduce bandwidth first and remove the higher-frequency components. This is also justified by the fact that later averaging is massively reducing this bandwidth anyway, so there is no loss as long as the upper bandwidth limit is sufficiently above f_{spin} .

We agree that $f_{\min}$ contains more than the actual error and appreciate the suggestion to use a low-pass filter before fitting. We applied a zero-phase low-pass filter with different cutoff frequencies. In Fig. 6 we present exemplary raw data and low-pass filtered data with different cutoff frequencies.

We selected to continue with a cutoff frequency of $2f_{\text{spin}}$ as this is still resembling the amplitude of the

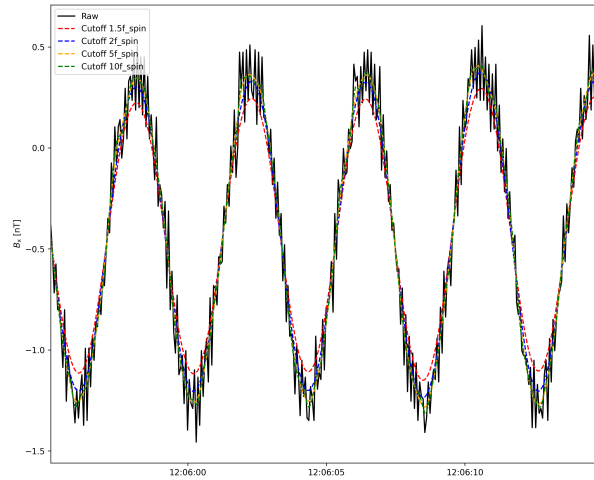


Figure 6: Comparison of raw data (black) and low-pass filtered data with different cutoff frequencies ($1.5f_{\text{spin}}$ in red, $2f_{\text{spin}}$ in blue, $5f_{\text{spin}}$ in yellow and $10f_{\text{spin}}$ in green) at 2017-01-15.

raw signal (contrary to $1.5f_{\text{spin}}$), while ignoring the higher frequent oscillations (contrary to $f_{\text{cutoff}} > 5f_{\text{spin}}$). When we compare the deviation from the original measurements, reducing the bandwidth with the low-pass filter is improving the result a little bit but does not affect the end result a lot. This is visible in Fig. 8 and Fig. 9 when we compare the widths of the distributions (more discussion and explanation in the next answer). In the revised manuscript we added the low-pass filtering as a new step to the implementation and revised the discussion/explanation of $f_{\min}$.

In the revised version the low-pass filtering is added in lines 109 to 111; more details on $f_{\min}$ are added in lines 142 to 147. We also updated the flow chart in Fig. 3 (previously Fig. 2).

- The other topic is the weighting with the value of A , i.e. the constant spin plane field. While of course such a field could increase some errors, it might also be correlated with the spin axis field – i.e. high spin plane and axis fields come at the same time. While the first degrades results, the second is actually improving results. IMO the weights should be revisited. In addition there might also be a correlation between field magnitude and field variation, i.e. $f_{\min}$. This would require analysis before specifying an error criterion. One could also discuss the justification of this scaling. The quality measure would only increase the weight of high field regions in the case there are also low field regions within the same 1-minute interval. In all other cases, the field is (typically) consistently either high or low over 1 minutes, so the scaling will not have any effect.
- Generally, the application of such an error criterion over such a long time (1 minute or 15-ish spins) is not advisable. Generally, as also done in most filtering applications, the weight of a sub-estimate should be reduced with the time distance to the currently averaged point. While its quality of a point with high time distance may be higher, its validity for the current time is typically lower. The used approach here could (theoretically) mean that

a single spin with high field could dominate a complete 1-minute interval. There are 2 options out of this time problem:

- Scale with time distance and/or
- Decrease output cadence to 1 minute instead of T_{spin} (even then additional distance scaling might be appropriate)

We have revised the weighting and agree that the factor A^2 has no significant influence, as it does not vary strongly within one-minute intervals. We therefore compared three different weighting schemes: A^2/f_{min} , $1/f_{\text{min}}$, and $1/f_{\text{min}}$ with additional time-distance scaling. The results show that using A^2/f_{min} or $1/f_{\text{min}}$ leads to very similar recovered spin axis values (see the blue histogram in Fig. 7). Adding time-distance scaling changes the results more, as illustrated by the orange and black histograms in Fig. 7.

When comparing the recovered data with the averaged original FGS data (see response to the comment

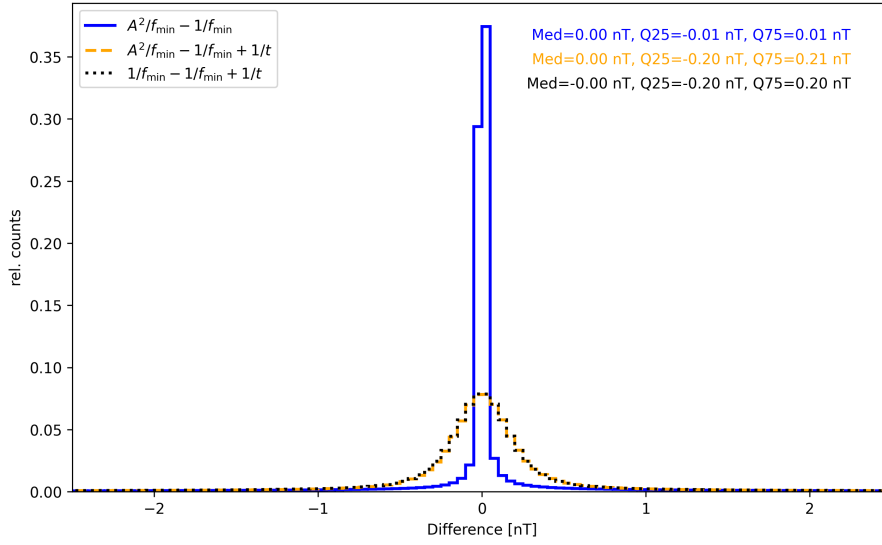


Figure 7: Comparison of recovered spin-axis values obtained with different weighting schemes for the years 2017 to 2020. The blue line compares A^2/f_{min} with $1/f_{\text{min}}$, the dashed orange line A^2/f_{min} with $1/f_{\text{min}}$ including additional time-distance scaling, and the black dotted line compares $1/f_{\text{min}}$ with and without additional time-distance scaling.

on bandwidth reduction of the original measurements) for 2017 to 2020, the resulting histograms differ only slightly between the weighting schemes A^2/f_{min} and $1/f_{\text{min}}$ with time-distance scaling (see Figs. 8 and 9). The version including time-distance scaling performs slightly better, as indicated by its smaller interquartile ranges. We therefore chose to use $1/f_{\text{min}}$ together with time-distance scaling as the revised weighting scheme.

In the revised version we changed the paragraphs describing the weighting in lines 156 to 166 and changed/added Eqs. 13, 14, 15, and 16.

Line 150: The impact of α_{S1} and $S2$ is relevant and the method is correct, but there is no error estimate and no method is presented.

We agree and expanded the paragraph. We estimate the uncertainty $\Delta\alpha$ by comparing our calculated α_{fit} values with the model field with the values directly from the calibration file α_{cal} from previous years when we could still use the routine calibration methods. We are using the median difference between these two from 2017 to 2020:

$$\Delta\alpha \approx \text{median}(|\alpha_{\text{cal}} - \alpha_{\text{fit}}|). \quad (1)$$

This yields uncertainties of $\Delta\alpha_{S1} = 0.045^\circ$ and $\Delta\alpha_{S2} = 0.058^\circ$. We also expanded the explanation of how the updated values of σ_{P_x} and σ_{P_y} are calculated. In addition, we now quantify in the discussion section how the uncertainty in α translates into the uncertainty of the recovered field.

In the revised version we added the comparison and more details in lines 185 to 195, and the influence

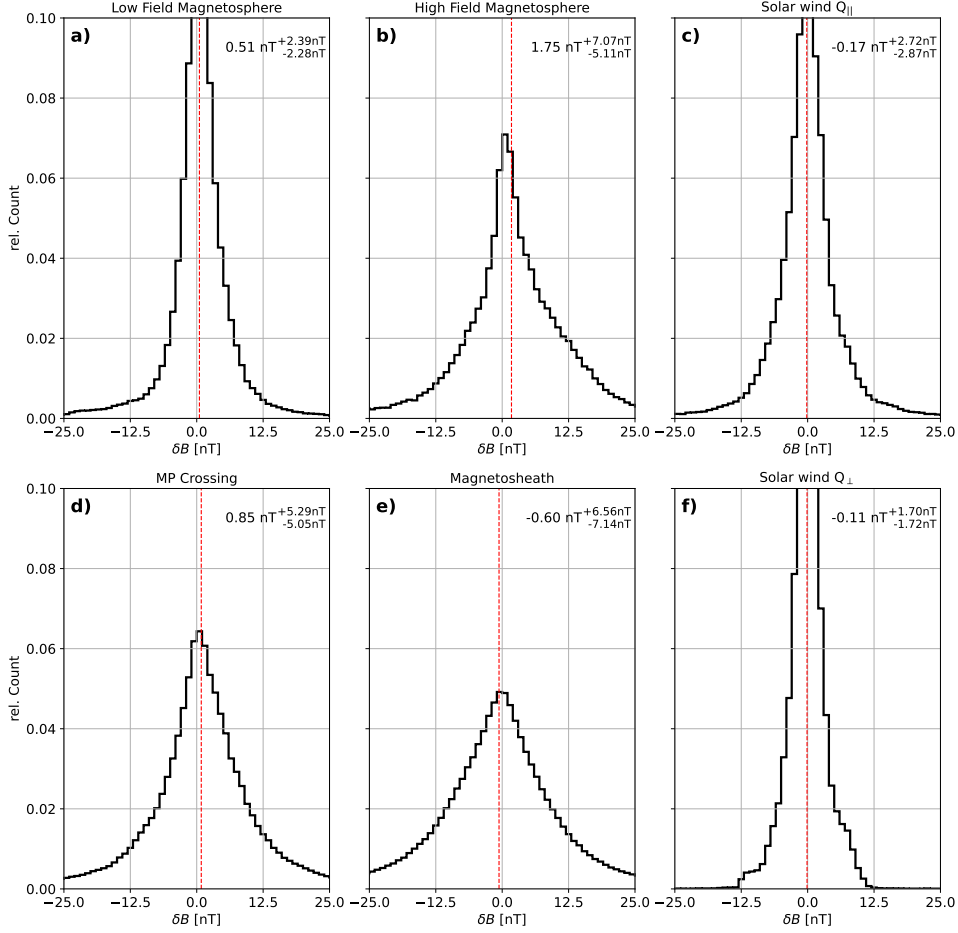


Figure 8: Histograms of the deviation δB between recovered (B_r) and original (B_{FGS}) spin-axis measurements in six different regions. Here we used $A^2/f_{\min}$ as weight and applied a simple moving average to both recovered and original data.

on the recovery in lines 329 to 336.

Line 159:

- *There is no real error/sensitivity analysis. The impact of having wrong calibration parameters is not discussed. The impact of amplifying the tiny spin axis part within the spin plane with a value from several tens to 100 (e.g. for elevation=1°) is not discussed – this would e.g. also scale sensor noise and have a relevant impact on the expectable quality of the result. It cannot become any better than that.*
- *As a result, the authors only present an empirical error of the given data set, which is an upper bound for a given data set. Nevertheless, it is not known if this error is different for other data sets and what the sensitivity of these bounds relative to other calibration parameters is. I am aware that not all of this should be written in the paper, but there is some room to improve.*

We agree that the discussion of uncertainties should be expanded. In the revised manuscript, we therefore added an estimate of the lower bound of the recovery uncertainty due to the sensor noise. As the reviewer correctly noted, the noise of the spin plane sensors is amplified by the same geometric factor as the recovered signal, i.e. by $1/|\cos \alpha|$, when projected onto the spin axis. Using the sensor amplitude spectral density near the spin frequency, $S_B^{1/2}(f_s) \approx 10\text{-}15 \text{ pT}/\sqrt{\text{Hz}}$, and an effective bandwidth of $\Delta f = 1/60 \text{ Hz}$ for the final one-minute product, we estimate a lower-bound noise level of approximately 80-120 pT for the recovered spin-axis field (details in the revised manuscript). Since the sensor noise increases toward lower frequencies, this estimate is optimistic. Nevertheless, it remains well below the

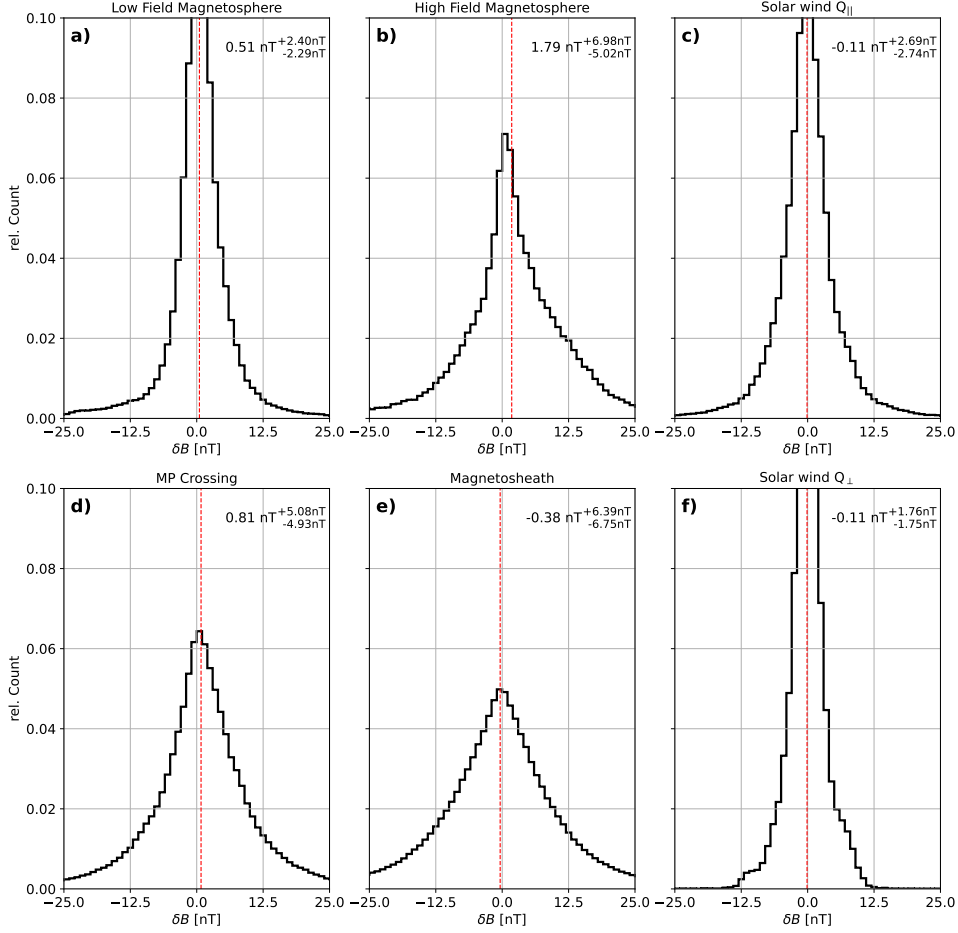


Figure 9: Histograms of the deviation δB between recovered (B_r) and original (B_{FGS}) spin-axis measurements in six different regions. Here we used $1/f_{\min}$ as weight and applied a moving average scaled with time distance to both recovered and original data.

interquartile ranges of the residual distributions derived from comparison with calibrated data, indicating that sensor noise is not the dominant limiting factor of the recovery.

[In the revised version we added the noise analysis in lines 286 to 303.](#)

Figure 3:

- The chosen time + amplitude scales of the plot are not suitable to actually demonstrate the quality of the recovery. An overview plot is great, but detail plots would be needed.
- In addition, the choice of unfiltered FGS vs. the recovered signal suggests differences that are in fact caused by the higher bandwidth of FGS – it would be necessary to reduce the bandwidth of FGS to the one of B_r before doing this comparison. This would most probably show better results.
- The same applies to figure 4

Figure 4:

- It might be worth to reduce bandwidth and to use time lags for comparing the 2 spacecraft

We agree that comparing the recovered data with FGS data of different bandwidth is not suitable. We therefore reduced the bandwidth of FGS data for both THE and THA by averaging in the same manner as described above for the recovered data. In addition, we added more panels with details to Fig. 3 and Fig. 4 in the revised manuscript. We chose not to apply a time lag on THA because they observe the same features nearly simultaneously, and no systematic time offset is evident in the updated Fig. 4.

In the revised version we changed Figures 5 and 6 (their number changed because we added plots).

Figure 5:

- Also this figure has the problem of different bandwidth between Br and FGS. The shown errors are a mixture of actual error and high frequency content. With the general inverse power laws, the existence of higher frequency signals with lower amplitude is more probable and so the error distribution is potentially highly related to the high frequency components rather than the actual error.
- This plot has identical scaling on the x axis. This choice is valid, but it is not really giving a quality estimate for the residual – e.g. in solar wind the error will be small per se, as there is not a lot of signal anyway. Even an opposite signal might result in an error of $4 nT$.

In addition, the SNR of the tiny bit of spin axis field in the plane axes is probably quite bad, it is probably not that much above the noise. It might be worth to add some “SNR limit line” that helps to differentiate between errors in recovery and the limitations of noise that is just unavoidable.

We agree that we also need to reduce the bandwidth for this comparison. We also discuss not only the absolute, but also relative errors in the revised manuscript. In addition, we like the idea of adding a SNR limit line. However, the estimation of a lower limit due to sensor noise results in values too close to zero to be meaningful in plot and we chose to not include them (see comment above).

In the revised version we expanded the discussion with relative errors in lines 279 to 285 and added a discussion about the lower bound due to sensor noise in lines 286 to 303.

Line 283:

- This statement should not be given without proper analysis (e.g. impact on currently used calibration method, errors and parameter accuracy). One example is that every axis offset would require a mixture of Hedgecock and spin calibration output – but those require different external magnetic field conditions and are therefore executed on different data sets and times.
- Suggestion for a change: “The impact of such an intentional tilt would require detailed analysis though, as it might impact the sensitivity and errors of calibration during normal operation scenarios.”

We agree and appreciate the suggestion. We revised the statement.

In the revised version we changed the paragraph in lines 373 to 377.

1 Comparison of Recovery for Different Years

References

Auster, H., Glassmeier, K., Magnes, W., and W. Baumjohann, O. A., Constantinescu, D., Fischer, D., Fornacon, K., Georgescu, E., Harvey, P., Hillenmaier, O., Kroth, R., Ludlam, M., Narita, Y., Nakamura, R., Okrafka, K., Plaschke, F., Richter, I., Schwarzl, H., Stoll, B., Valavanoglou, A., and Wiedemann, M. (2008). The themis fluxgate magnetometer. *Space Sci. Rev.*, 141:235–264.

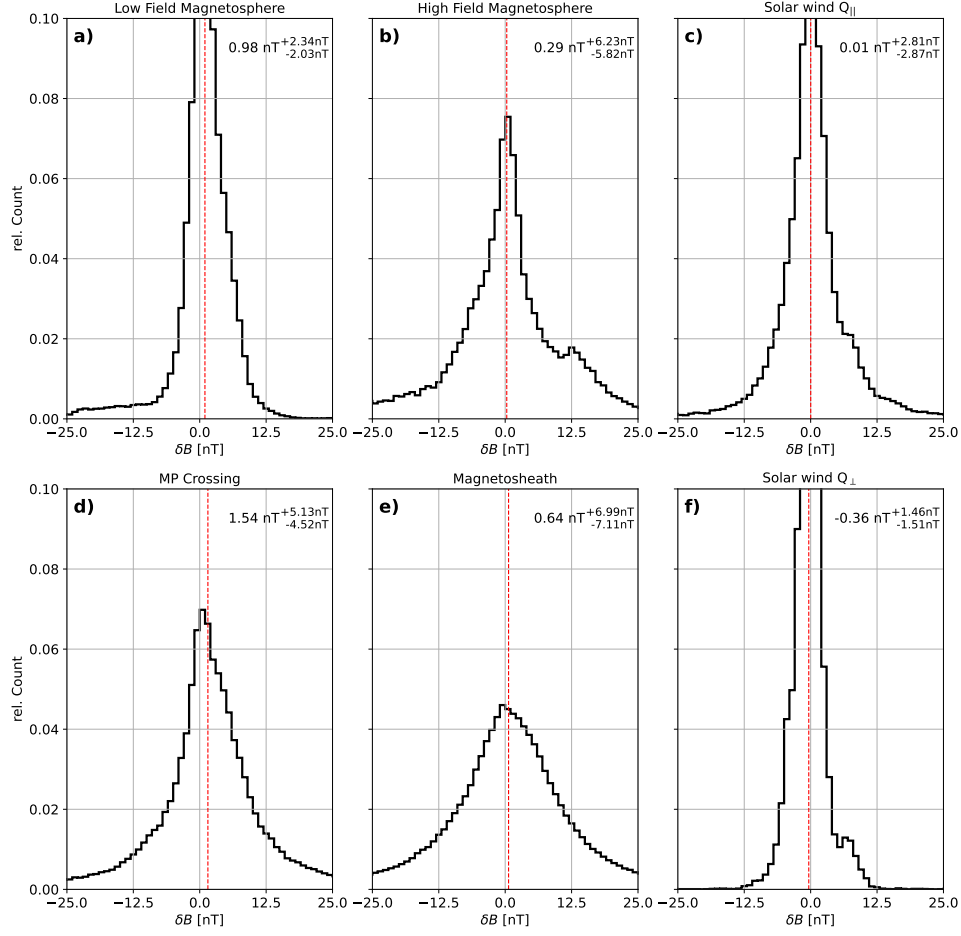


Figure 10: Histograms of the deviation δB between recovered (B_r) and original (B_{FGS}) spin-axis measurements in six different regions for the year 2017. Here we use A^2/f_{min} as weight and applied a simple moving average to both recovered and original data (a more detailed discussion of various weighting schemes will follow later).

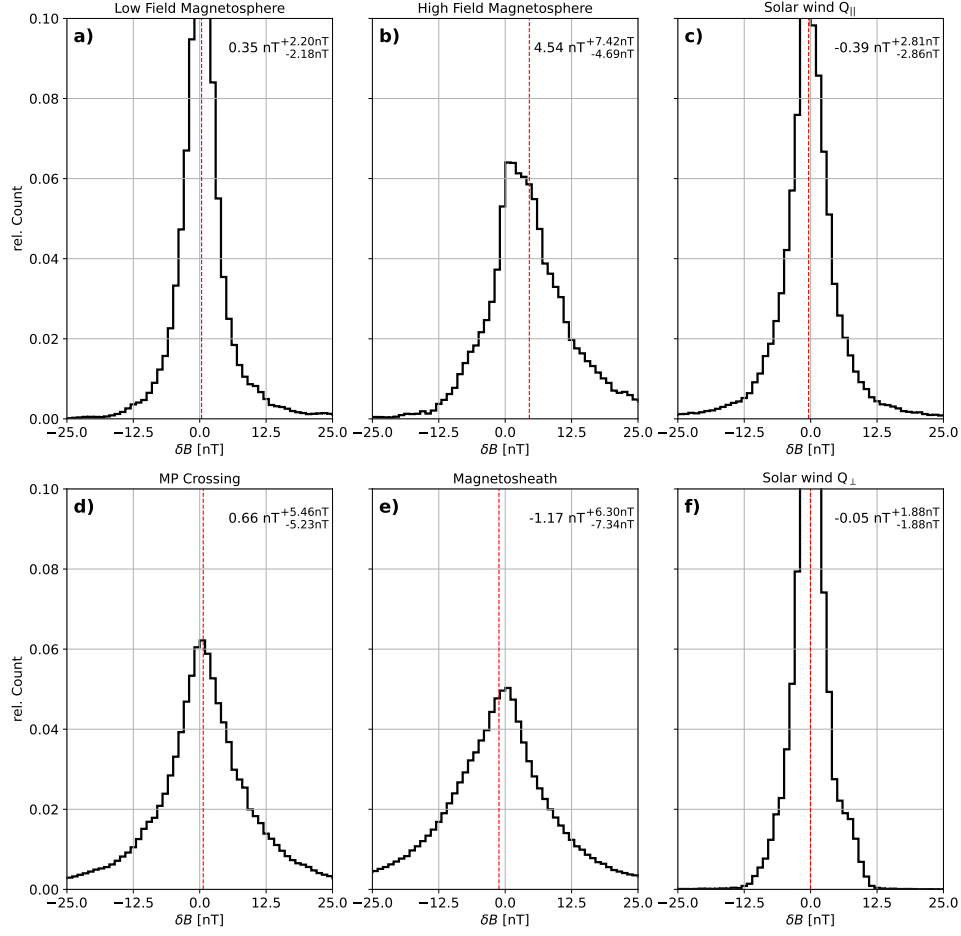


Figure 11: Histograms of the deviation δB between recovered (B_r) and original (B_{FGS}) spin-axis measurements in six different regions for the year 2018. Here we use $A^2/f_{\min}$ as weight and applied a simple moving average to both recovered and original data (a more detailed discussion of various weighting schemes will follow later).

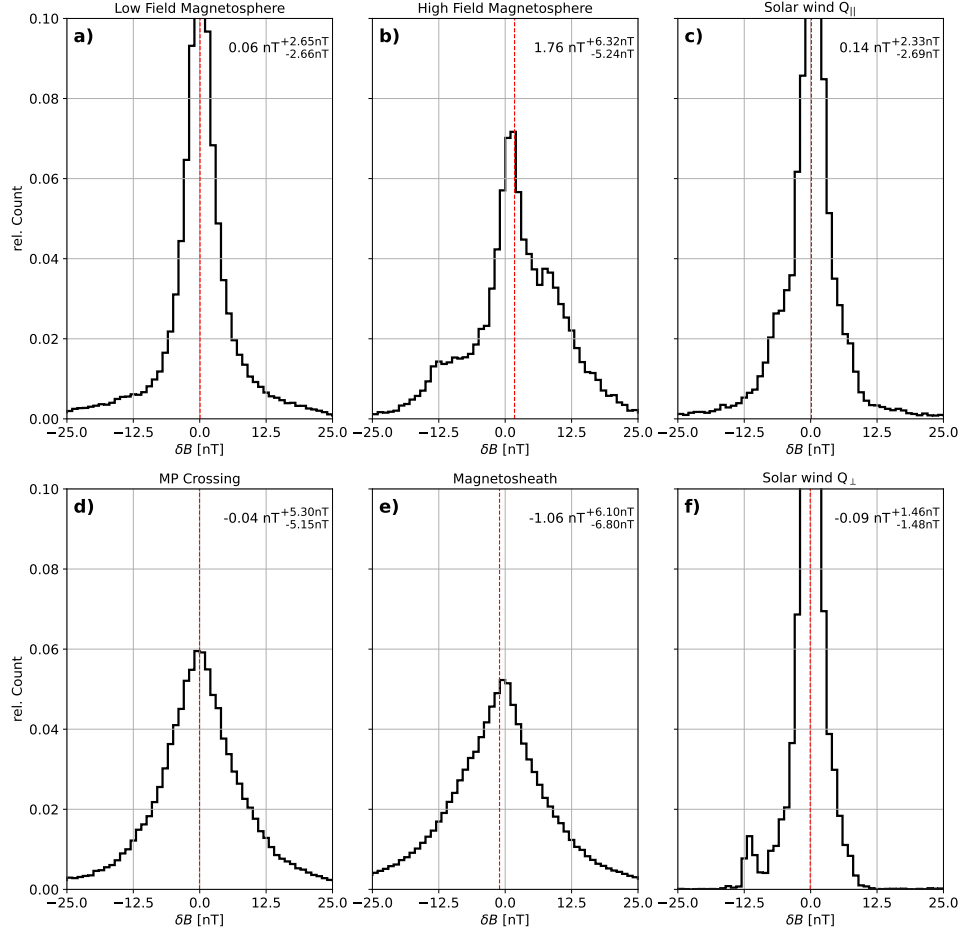


Figure 12: Histograms of the deviation δB between recovered (B_r) and original (B_{FGS}) spin-axis measurements in six different regions for the year 2019. Here we use $A^2/f_{\min}$ as weight and applied a simple moving average to both recovered and original data (a more detailed discussion of various weighting schemes will follow later).

Author response to referee #2

The authors thank Anonymous referee #2 for their detailed comments on the manuscript. We will take the comments into account when revising the manuscript. In this document we provide responses to each of the referee's comments (formatted as italics in indented paragraphs and grouped together when appropriate). At the end of our answers, we indicate in blue where the changes have been made in the revised version.

Please do not use the term "offset" in connection with the measured magnetic field value, p. 2, line 46. It is very confusing considering that this term is reserved for quantifying the systematic error of the sensor. Replace it with, for example, the term "contribution" or similar. Likewise, the term "despun coordinate" is inappropriate and very confusing.

If I understand it correctly, the projections of the BSA field component measured along the spin axis onto the axes of the other sensors represent that permanent, non-removable contribution. However, this does not exhibit sinusoidal behavior with the ω_{spin} frequency, as stated on p.2, line 47. This is also evidenced by relation (2), where the only time dependence of this contribution is related to the time variability $B_{\text{SA}} = B_{\text{SA}}(t)$. Please clarify this part of the text.

We agree that the wording of this paragraph is unfortunate and does not explain the situation well. We rephrased this paragraph, replacing the wording "spin tone" with "projected spin axis contribution". We also renamed the variable from B_{ST} to B_{pSA} throughout the manuscript.

In the revised version we have reformulated the paragraph in lines 42 to 48. In addition, we updated the wording and variable name throughout the manuscript. We refer to the manuscript with track changes to avoid listing all the changes here.

For the sensor configuration as described in Abstract and Introduction, the second step in the coordinate transformation is redundant because the probe spin axis coincides with the Pz-axis, as it is depicted in Figure 1(a). The presentation of Figure 1(b) and the matrix Σ (which actually should be an identity) is very confusing for the reader. Please remove this contribution to the explanation, or present it as a general case of rotating the sensor design with appropriate commentary for the current case.

We agree that the description of the sensor system needs improvement. Accordingly, we added a third panel, Fig. 1(c), to illustrate the angle α between spin plane sensor axes S1 and S2 and the spin axis (see Fig. 1). However, we kindly disagree that the Pz-axis and the spin axis coincide which should be visible in Fig. 1(b) with the spin axis in black and the Pz-axis in blue.

In the revised manuscript we explain the sensor system and transformation between the different

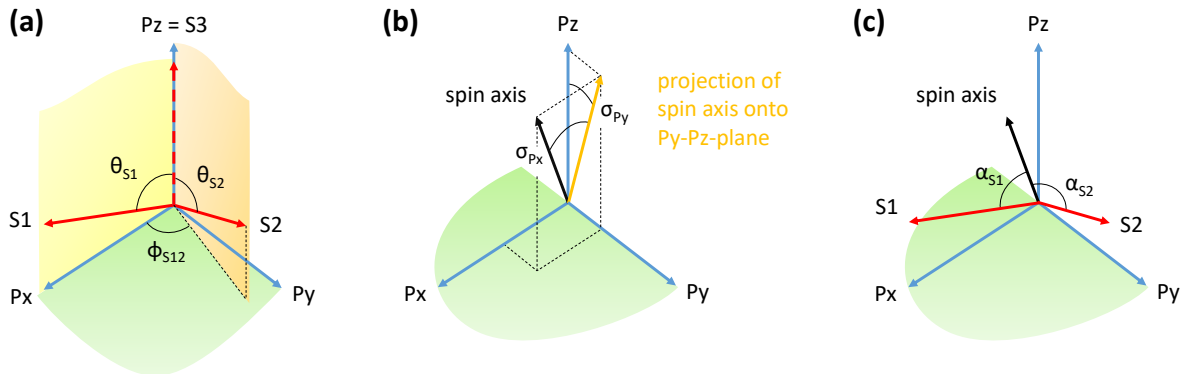


Figure 1: Updated sketch of sensor coordinate system. The panel c) shows the angles α_{S1} and α_{S2} between spin axis and spin plane sensors.

coordinate systems in more detail.

We updated Fig. 1 in the revised manuscript and rewrote the paragraph explaining the coordinate transform in lines 72 to 93.

Please justify the use of the transformation Φ , apparently $\phi_a = \phi_a(t)$.

There were no requirements for the orientation of the FGM instrument onboard the THEMIS spacecraft [see Auster et al., 2008]. Therefore the rotation around the spin axis just comes from the position of the FGM onboard the spacecraft, and ϕ_a is needed to get from the sensor system to a physically meaningful spacecraft-fixed coordinate system, from which further transformations into other coordinate systems, such as geocentric systems, can be performed. In Fig. 2 we present the evolution of all 9 calibration parameter for THE over the mission duration. As can be seen in Fig. 2(g) ϕ_a stays quiet constant the whole time.

We added Fig. 2 to the revised manuscript and justified Φ in lines 86 to 88.

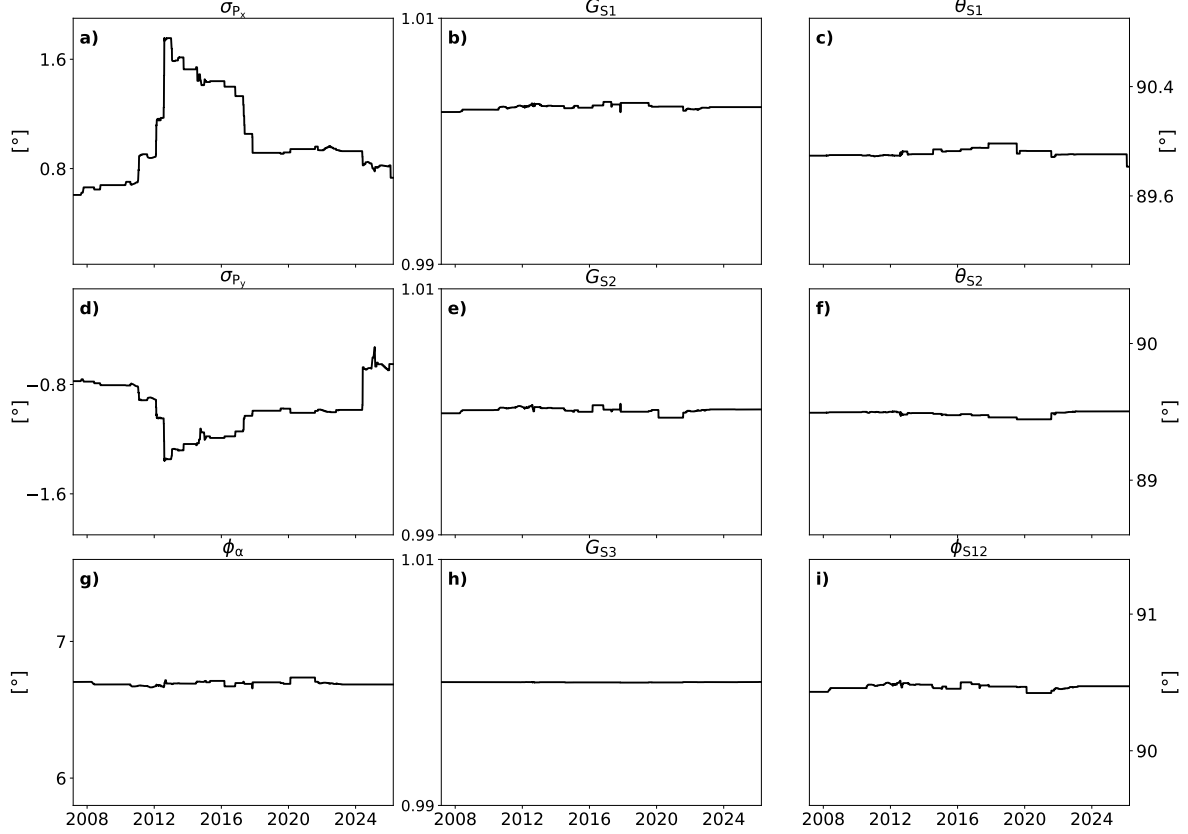


Figure 2: Evolution of calibration parameter: σ_{px} , σ_{py} , and ϕ_α are the rotation angles around the orthogonalized sensor axes (c.f. with blue coordinate system Fig. 1 in the manuscript or Fig. 1 in this reponse). G_{S1} , G_{S2} , and G_{S3} are the gains along the sensor axes. θ_{S1} , θ_{S2} , and ϕ_{S12} are the internal sensor angles as described in Fig. 1 in the manuscript.

It is appropriate to present the conversion by the $\mathbf{G}$ matrix first, since this is applied to the data that has not yet been transformed.

We agree and reordered the presentation of the matrices.

The gain matrix is now described in lines 72 to 76.

The origin of relation (8) is unclear, the angles α_{Si} apparently $\alpha_{S1} = 90^\circ - \theta_{S1}$, $\alpha_{S2} = 90^\circ - \theta_{S2}$ (what is α_{S3} ?). Please provide the derivation and explain also with regard to the reservation in point 1.

We agree that the angles α_{S1} , α_{S2} and α_{S3} have not been introduced well. They are the angles between the sensor axes S1, S2 and S3 and the spin axis, respectively. As mentioned above, we will add a third panel to Fig. 1 and will expand the explanation of the sensor system. Eq. 8 is simply the projection of the spin axis onto the sensor axes, using the standard relation for the angle between two vectors,

$\cos \theta = \vec{x} \cdot \vec{y} / (|\vec{x}| \cdot |\vec{y}|)$. Using the transformation matrices, the calculation of the angles can be expressed as given in Eq. 8.

We update Fig. 1 in the revised manuscript and rephrased the paragraph before Eq. 8 in lines 95 to 100.

Please specify that formula (11) is applied to (9) and (10) individually. From the fitting procedure, I assume, the unknown values $A_1, B_1(t), C_1, D_1(t), \varphi_1$ and $A_2, B_2(t), C_2, D_2(t), \varphi_2$ arise and the value $B_{\text{ST}}(t)$ is common to both expressions. Please demonstrate in detail how this is achieved. In relation (11) also correct for the linear increase in amplitude with time.

We agree that we have to be more precise. We point out that A, B, C, D and φ are free fitting parameters and might differ between the two components, whereas ω_s is provided by the state file of THE.

Applying the fit to both spin plane sensor axes data results in two independent estimates for B_{pSA} (formerly B_{ST} , see comment above). Referee #1 pointed out, that we used t in Eq. 11 and Eq. 13 although they are not the same. Therefore we replaced t with τ in Eq. 11 with τ ranging from 0 to $2.5T_{\text{spin}}$. To obtain one B_{pSA} value for one fit window, we are then using the central average $B_{\text{pSA}} = B_{\text{pSA}}(\tau = 1.25T_{\text{spin}})$. We will expand the description of the method and will fix the aforementioned inconsistencies.

In the revised manuscript we rewrote this paragraph and added more details and explanations.

We revised subsection 2.3 in lines 104 to 123.

p. 5, line 111: You now have recalculated the components B_{raw1} and B_{raw2} (and not $(B_{\text{S1}}$ and $B_{\text{S2}})$), according to relations (9) and (10). Please consider the designation for raw data and treated data and be consistent in their use in the following text.

We agree that we need to be more precise and careful when naming variables. In this sentence, we intended to inform the reader that we calculate B_{pSA} independently for both spin plane sensor axes, as mentioned in the comment above, with details on the implementation to follow later in the manuscript. We revised this paragraph to avoid confusion and resolved inconsistencies in the naming of our variables. We revised the paragraph in lines 125 and 126 and placed the part with independent calculation in lines 153 to 155.

p.5, line 115: An average reader is not familiar with the devices you mention. This requires a more thorough description of the functionality of the device whose data you use, and possibly also a picture with a description.

We can add a more detailed description, but we would prefer not to include another image, since it is already included in the cited work and this manuscript will contain many illustrations nonetheless.

We revised the paragraph in lines 129 to 134.

p.5, line 120: Please provide an example of this procedure using relations (12) and (13) with specific values and a specific result and with an appropriate comment.

We included a figure and an explanation to illustrate the fitting and shifting process in detail. This is exemplary shown in Fig. 3.

We added Fig. 3 to the manuscript (in the revised manuscript it is Fig. 4 as we added other figures as well). We revised the paragraphs explaining the implementation in lines 135 to 152.

What is the origin of the amplitude $A(t)$ in the weight function $w(t)$, p. 6, line 126? What are the criteria for its magnitude? Please, provide details.

The parameter $A(t)$ stems from the idea to not down weight intervals with large magnetic fields. But since $A(t)$ is not changing meaningful over the duration we average (1 minute), we changed the weighting to $1/f_{\text{min}}$ and include a scaling with time distance.

We revised the paragraph for the weighting in lines 156 to 165.

The angle α ($\alpha_{\text{S1}}, \alpha_{\text{S1}}$) does not come from measurements or from the derivation of the matrix C - on the contrary, it must enter the calculations as a data known from the technical settings of the device. Please clarify this fact on p. 6, 7. If this is not the case, I consider it im-

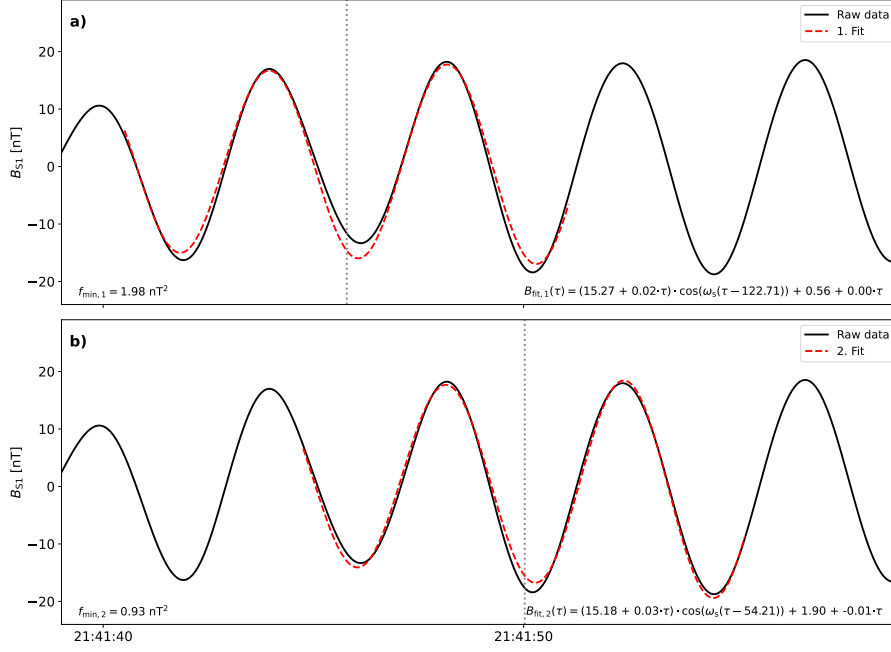


Figure 3: Two consecutive fits applied to the modified raw data (black) to calculate B_{pSA} . The dashed red lines represent the applied fit which is only defined where drawn. The fit in panel b) is shifted by one spin period compared to the one in a). In addition, at the bottom we present the f_{min} values and the fitting parameter for each fit.

possible to solve the considered problem - then the equation (2) has all input quantities unknown.

We derive the angle α by comparing our recovered projected spin axis contribution B_{pSA} with the model field near perigee. We assume that this provides a sufficiently accurate estimate, since the model field is a good approximation of the actual field near perigee. To verify this, we compared our estimated α values with those from the calibration matrix over a three-year period from 2017 to 2020. This is shown in Fig. 4.

As can be seen from the histogram, the angles differ only minimally, with median values of 0.045° and 0.058° for α_{S1} and α_{S1} , respectively. We agree that we need to be more precise in our wording and revised the relevant sections of the manuscripts accordingly.

We revised the subsection 3.2 explaining how we calculate α values in lines 168 to 195. We also added the comparison to α values from the calibration file to show that we get reliable estimates from the model field. This is done in lines 185 to 187.

Please increase the size of the images.

We increased the size of the images.

We updated the size of all figures.

Please add information regarding the data's affiliation to the period and probe in the description of Figures 5 and 6 (as well as to the other figures).

The information which spacecraft is used at what time and in which data mode is now given for all figures. In the process of updating figures due to other referee comments, we also updated the captions of almost all figures. Now, they should offer sufficient information.

We updated the captions of Figure 1, 2, 4, 5, 6, 7, 8 and A1.

Conclusions: Please, design the optimal geometric arrangement of the sensor axes and justify it.

We agree that this statement should not be made without adequate justification. As referee #1 also noted, we must first analyze the implications for the calibration method currently in use, as well as the

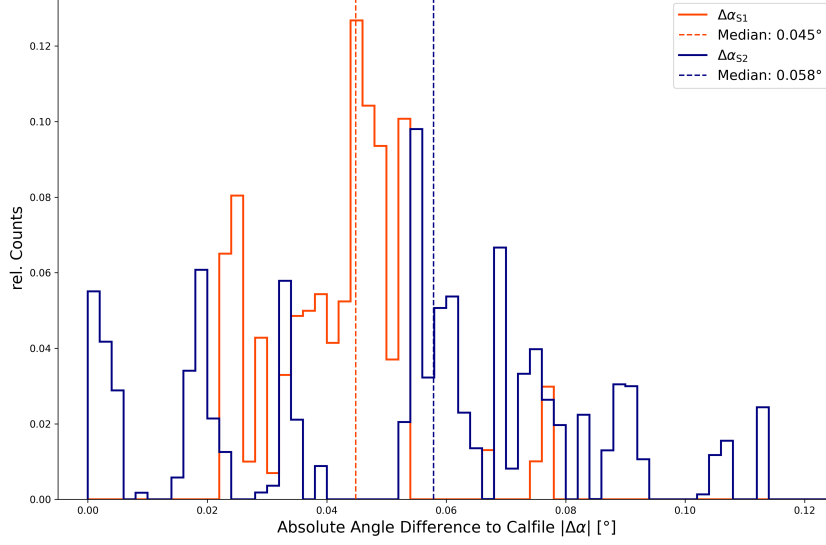


Figure 4: Difference between α values from comparison of model field and B_{pSA} and values from the calibration matrix.

errors and accuracy of the parameters. Since this would go outside the scope of this work, we softened the statement.

The revised paragraph is in lines 373 to 381.

Minor corrections:

p.1, l.3: ...including a set of fluxgate magnetometers that measure...

p.1, l.21: ...predefined rate denoted as T_{spin} ... please correct. Provide the value.

p.2, l.36: ...six orientation angles... please, explain or correct.

1. We kindly disagree, since there is only fluxgate magnetometer on each spacecraft.
2. We provided a range since T_{spin} changed during the duration of the mission.
3. We briefly mention the orientation angles and will refer to the section 2.1 where we explain the geometry in more detail.

We added a range for T_{spin} in line 19 and added a brief explanation and reference to Sec. 2.1 in lines 36 to 37.

References

Auster, H., Glassmeier, K., Magnes, W., and W. Baumjohann, O. A., Constantinescu, D., Fischer, D., Fornacon, K., Georgescu, E., Harvey, P., Hillenmaier, O., Kroth, R., Ludlam, M., Narita, Y., Nakamura, R., Okrafka, K., Plaschke, F., Richter, I., Schwarzl, H., Stoll, B., Valavanoglou, A., and Wiedemann, M. (2008). The themis fluxgate magnetometer. *Space Sci. Rev.*, 141:235–264.