Geoscientific Instrumentation Methods and Data Systems Discussions

Geosci. Instrum. Method. Data Syst. Discuss., author comment AC1
https://doi.org/10.5194/gi-2021-17-AC1, 2021

Reply on RC1

Filomena Catapano et al.

Author comment on "Swarm Langmuir probes' data quality validation and future improvements" by Filomena Catapano et al., Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2021-17-AC1, 2021

We thank the referee for his careful review, and for the useful and stimulating comments. We understood that some part of the text could be actually improved and we tried to follow his suggestions as discussed in the following.

Specific comments

1.In the introduction the authors list some of the most noteworthy of Swarm's accomplishments up to now. This does not really prepare the reader for the contribution that is presented in the manuscript, so I would consider this as largely superfluous material; the reference to the Swarm publication list would be sufficient – unless the authors would indicate, for the cited contributions, what role the Langmuir probes have played in these accomplishments. That would then raise the reader's awareness that these data have been proven to be of considerable scientific importance and thus that quality control is an absolute necessity.

1.Answer: We thank the referee for this useful comment. We have modified the introduction part emphasising the role of LP measurements in the different cited studies. Also, we added a sentence to highlight the importance of the quality control activities to maintain a high instrument performance and data quality.

2.I was puzzled by the explanation (around line 80) of the difference between low and high gain, even after consultation of Knudsen et al. (2017) – in any case I suggest to add that reference here explicitly. Later on, at the beginning of section 3.1.1, the authors give a more precise explanation. I strongly recommend to move that explanation here; this will considerably improve the readability of the text.

2.Answer: Following this suggestion we have moved the explanation in section 3.1.1 in this part of section 2, adding also the reference to Knudsen et al 2017.

3.The end of the paragraph on line 81 is a surprising one. The authors say that the use of TiN was questioned and that therefore one of the probes received an Au coating. But then they end the paragraph with a statement about the uncertainty regarding the preservation of the Au coating. I would expect the final statement to address whether the TiN coating is still preserved, because that was the initial concern.

3.Answer: Following this suggestion we have added a sentence that both probes don't

seem to suffer from any degradation because of oxidation. Both methods, nitration and gold-plating seem to work as far as we can tell.

4.I am curious to know how geomagnetic indices are used to process the LP data (line 94). Could the authors expand on this?

4.Answer: We are grateful to the referee for this comment because we realised that there was an error in the text. Actually, geomagnetic indices are not used in LP data processing. The sentence has been modified to reflect the actual input of auxiliary data in the LP PLASMA processor.

5.On line 100, the authors state that interpolating the 2 Hz data product at exact UTC leads to the 1 Hz sampled data. Is linear interpolation used? To what extent is linear interpolation justified as the signal likely contains faster time variations? I would expect that one would first perform a smoothing or filtering in order to remove time variability faster than 1 s before doing such an interpolation. Does one interpolate the densities and temperatures or their logarithms – which is more appropriate for quantities that vary over orders of magnitudes and that are always strictly positive? One could also interpolate the measured currents, and then do the processing with those interpolated values ...

5.Answer : In the ripple (or, equivalently, harmonic) mode the measurements for each plasma data "point" occurs over a period of about 400 ms. The ion density estimate is from only one of the three current/admittance measurements which still stretch over a period of about 100 ms. By correlation with the 50 Hz magnetic data the sample "points" of the ion density has been determined to be at 167 and 696 ms into the full second. Not surprisingly the correlation peak is quite broad. The interpolation of Ni and Te to the full second n is just linear between the n-1.696 and n.167 points.

The referee's suggestions are very reasonable and we might consider them as an improvement in a future update of the data processing. But we fear that any positive effects would be swamped with the uncertainties in the sampling timing and other data noise. The purpose of the interpolated data set is mainly convenience for users who need to correlate with the 1 s magnetic data.

We have replaced "exact UTC" by "full UTC second" and added "simple linear interpolation" to emphasize better that the method is quite approximating.

6.Section 3.1.1 gives a clear explanation of the old algorithm and of the new one. It presents an evaluation of the changes and indicates that this is an improvement because of the better correspondence with ground-based measurements. However, I miss a fundamental point: Why do the low and high gain measurements differ? As the authors have indicated, the difference is a resistor in the measurement circuit. Is there a physical underpinning of the measured difference? I believe significant effort should be invested into this. After all, this change in electron temperature calculation is the main point of improvement in data quality upon which the whole paper is built, so it has to be well documented and justified.

6.Answer: We agree with the referee's comment. The two probes of each Swarm satellite are different with respect to

- the probe material: the left probe is made of TiN, the right probe of gold-plated Ti;
- the position on the spacecraft with a distance between probes of only about 30 cm;
- the electronic gain, which is configurable between high and low and has been, except for test periods, always opposite between probes.

By switching the gain between the probes we determined that the gain seems to be the

main cause of the differences, the others seemingly having only small effects. When the density is low, then a low gain leads to very small currents, which are difficult to measure and electronic noise and offsets become significant, especially for the ripple mode. The high gain configuration should be much more accurate at low density, but it overflows for high density. When designing the instrument the expectation was that the results from high and low gain configurations would approximately converge for a range of intermediate density. In spite of still ongoing investigations we do not have yet a clear explanation why this is normally not observed.

We prefer to not expand the discussion on this topic to avoid describing investigations which may not bring a real conclusion. We added a sentence to describe the status of the related investigations.

7.I would welcome a clear definition of "data quality" as it plays such a central role in this paper. My intuitive understanding seems to be at odds with that of the authors. For instance, I read on line 213: "A larger percentage of invalid measurements obviously indicates a poorer data quality." That is not evident to me. One could argue otherwise: "if one obtains a higher percentage of invalid measurements, one is apparently able to catch very well those situations where the measurement process fails, so that one can have more confidence in the remaining data." Indeed, if there is an ADC overflow, that measurement clearly is not reliable, but that does not immediately say anything about the quality of the measurements performed before or after.

7.Answer: The data quality is the goodness of the data product as output of a processing process. The data quality can be qualified by comparison with other dataset (in-situ or ground measurement), validation with numerical or empirical models, or derived by statistical data analysis of the product itself. In our definition of data quality, if the measurement is subject to a low level of errors or data contamination derived from statistical analysis or known issues, or/and has a high agreement with other dataset (model or spacecraft observations) then the quality of the data is considered good. We have added this definition in the manuscript.

8.The authors discuss the evolution of quality with the solar cycle. How certain are they that an apparent systematic trend with the cycle does not mask detector aging (such as cumulative damage to LP coating, consequences of nanodust impacts, etc.)? Wouldn't one need at least a complete solar cycle to evaluate this? The topic of detector aging is only briefly touched upon. I think it deserves more attention, as this would be one of the major aspects of instrument quality monitoring. Such a discussion could be part of a more extended discussion section (which is rather short at present).

8.Answer: We thank the referee for this useful and stimulating comment. This is indeed a very interesting analysis. We are using some parameters to monitor the detector performance and so far we do not have any evidence of a degradation with respect to the solar cycle. However this analysis certainly deserves more attention and we should expand the set of parameters that we are monitoring before coming to a clear conclusion. We reserve to further investigate this aspect and present the results in a future work,

9.A very interesting point is the comparison of LP and FP densities presented in Figs. 6 and 8. On line 235ff the authors say that the FP processor does not need any assumptions regarding ion composition, while the LP processor does. But then line 236 states that measurement differences between both are due to different assumptions – that contradicts the statement that FP uses no assumptions at all. This is an important point, because it suggests that the night side FP density measurements are considerably more reliable than the LP densities there. And consequently, in view of the observed relatively poor correlation, I do not understand the assessment on p. 245 "The Swarm LP measurements are of very good data quality" – shouldn't this be qualified somewhat, e.g.

restricted to the day side? The conclusions section does list this problem as one that will drive future attempts for improvement.

9.Answer: We agree with the referee that the sentence was misleading, thus we adjusted the sentence specifying that the difference in the FP and LP measurements may be due to the assumption made for LP processor. We removed part of the sentence at the beginning of section4, because the details of the goodness of data are summarised in the conclusion part.

Detailed issues

10.Abstract: The abstract reads well, explaining that the paper discusses the quality control approach. I suggest to add a sentence that states how good the data quality is, before then saying that there is an anomaly. Please also do not use the LP abbreviation in the abstract as it is not explained before.

10.Answer: We agree with the comments reported by the referee and we adjusted the abstract accordingly.