Review: Atmos. Meas. Tech Discuss. (amt-2019-29)

High-precision Monitoring of compliance with fuel sulfur content through UAV measurement of ship emissions.

By Fan Zhou et al.

General comments

Title: Without trying to be negative I would suggest leaving out the words *High-precision*.

Abstract: As in the title I would suggest leaving out *high precision* in the last sentence. I would also mention the range of sulfur contents that were encountered in the study i.e. how many non-conformities were encountered. And I would like to mention more explicitly that the deviation of the estimated value for +FSC is less than 0.03% (m/m) at a level of 0.04 % to 0.24 % FSC. Note that in ECA areas, with a limit of 0.1%, an uncertainty of 0.03% is not very good. I would also suggest mentioning that in all cases the estimated FSC was always *lower* than the actual FSC derived from samples taken on board. This is an important aspect with a strong impact on the usefulness of the method in SECA areas with a 0.1% limit value.

Paper: This could be a very useful paper with lots of detail. Especially the level of detail is useful since this is an area with a lot of development and sharing of these new results could very helpful to other scientists. I provide some comments that could help to make the paper a bit clearer in some areas. See my specific comments below.

Figure 1: I am not familiar with UAVs and in a first glance I thought the black box mentioned in the text was the large flight case black box below the drone.

Pag 3 line 16. Not everybody may be familiar with the word *Pod*

Pag4 last sentence: electrochemistry method. *Electrochemical method?*

Pape 5 line 12-13. These sentences are rather unclear. What is meant with 180 working hours apart? Each 180 working hours? It is not entirely clear what the actual accuracy is if it is 1% full scale.

Pag 7 line 16: *correction* should be *corrected.* Gradually establishing a quality management system…. Is rather vague what is meant. Please rephrase.

Pag 7 line 22. Here 200 ppm is mentioned where in other places in the text 0.03 % (300 ppm) is mentioned. This should be explained or there should rather be only one number.

Same place: the deviations mentioned in Balzani et al (2014) were determined at FSC of 1%. It is not clear whether these deviations are still the same at 0.1% FSC. They could be lower at 0.1% S content. The authors should mention that or provide more information (which would be useful).

Page 7 last paragraph. To me it is not clear how errors in determination of the peak height is propagated in the total error and it is not clear how this is done. The error of 300 ppm is (it seems) related to the comparison with the on-board samples. And not from error propagation analysis as far as I can tell. It would be nice to show the error propagation numbers as well and see how well these two approaches match. In general, I think that the uncertainty discussion could be more quantitative.

Page 9 line 16: this makes the *FSC value relatively larger than that of CO_2*. It is not clear what is meant here.

Page 9 line 6: were *synchronized* is rather vague. Please explain

Page 9 In general, the data treatment is unclear to me. Why are peak values taken to compare SO_2 and CO_2? Or is it the surface area? The S-content may be derived from any set of concentrations. Taking the pea area ins just of way of averaging. It seems to me now that the peak position and its height is depending on the performances of the sensors (especially response time) and the accidental position in the plume. This could lead to uncertainties especially if the peak height only is used. This should be explained better. Especially the “approach” could be elaborated more. Sometimes I am in doubt whether peak means the highest point in the concentration or the peak area.

Page 9 in general: what exactly is “selected”. This should be made clear. Now it seems a bit arbitrary. Of course, full range values are not used. But what are *dramatic changes*? Would be useful to explain.

Page 9 line 21: 300 ppm at what level??

Page 10: Figure 5. Sometimes background values of SO2 are 400 ppb? That is very high. Why not subtract the background? Also in plume 6 the background seems to fluctuate which could hinder the intepretation

Page 11 table. Why is not a graph provided? Such as *true value* (x-axis) against *estimated value* (y-axis). Then also a correlation coefficient could be calculated. Also a good measure of quality.

In general: The results section could improve in clarity if some structure was used: data treatment; FSC etc. For example, the issues with sampling rate etc. (page 1 top) are perhaps important but mixed here with the results. To increase clarity this could be treated separately

Conclusions

High precision is not reasonable to state in view of the rather large underestimations.

Page 12: in Conclusions something might be said on the effect of SO_3 and SO_4

Specific comments:

I am not a native speaker, but the English seems fine with me in general. Some specific text could be altered:

- on ships the “chimney” is often called the “funnel”
- “ship” is normally “vessel”.
- *Culled* is not a word that is often used
- Pag 8 line 3: English: none of the monitored ships were fitted with exhaust cleaning equipment