Monitoring of compliance with fuel sulfur content regulations through UAV measurements of ship emissions

Fan Zhou[1], Shengda Pan[1], Wei Chen[2], Xunpeng Ni[2], Bowen An[1]

Page 1 Line 20 Perhaps there are more recent data (references) on the contribution of shipping then those from 2010. Especially since there is mention of “rapid development”

Page 2 line 11 English: the FSC **content** rather than FSC **limit.** And perhaps it should not say **must** not exceed but **may** not exceed.

Page 2 line 26 English: check the sentence starting with Optical methods. Unclear.

Page 2 Line 29 English: ***for calculating*** change to: **to calculate**

Page 2 Line 32 English: **ship’s plume** rather than **ship plume**

Page 3 English: **more accuracy estimation** change to **more accurate estimation**

Page 3 lines 10 etc. It should be noted that the accuracy of the results of monitoring is a difficult issue and the accuracy estimates in literature may not always be comparable. Please make a note on that.

Page 3 the section starting line 15 with Ship emission measurements. It is perhaps not the best option to start here with the distinction between land based methods and others. I would suggest to start with this subject (i.e. land based), add fuel sampling and then address all issues on accuracy etc. that’s to me seems a more logic order.

Page 3 line 21 English: collect exhaust **gas.** This change is also needed in other parts of the text

Page 3 line 26 English: **taken from aircraft** and not airborne

Figure 1 legend: filter to remove water vapor? or is that a dryer (nafion)?

Figure 2 is a nice picture, but it is not very illustrative. There is more than one exhaust and there is no visible plume. Please consider using another picture.

Page 5 If the Matrice instrument is modified, perhaps something should be said on what has been done. Or say: small modifications

Page 6 I thought the SO2 sensor was rather new. Yet reference is made to Hodgson 1999. Is that correct? Please check.

Page 7 : at what distance from the ship is the background determined? This may vary but what are ranges.

Page 8 line 2**:** English: **flied** change to **flew**

Pag 2 line 3. Why not determine the background after the peak as well? And see how it changes the results (or not). Please discuss.

Page 8 discussion on sampling interval. It does not seem to be completely solved by setting all intervals to 10 second. This is an outcome of the experiments.

Page 8 discussion on what is called “exhaust uncertainty” is difficult. In my opinion this is a completely different uncertainty and should be discussed on a different location in the paper.

- We measure the composition of the exhaust gas and derive from that the S content in the fuel. There are errors: caused by inadequate sampling, calibration, errors subtraction of the background etc. An attempt should be made to estimate the total error in the measured fuel content.

That is one side of the comparison:

- on the other side there is fuel sampling with its errors (perhaps), and the representativity of the sample. Is the fuel sample representative for the the exact moment of the measurement with the UAV? Is there a chance that the ship may have switched fuel.? These are additional uncertainties (that play a role in European harbors: switching of fuels. But no need to go into that. Ok we assume that the sample is representative is the right sample.

Then the last question is:

- is all Sulphur converted to SO_2 or also to other species. In Balzani et al a number of 6 % Sulphur s mentioned that is not converted. A correct sniffer method would then automatically underestimate the fuel content by 6%. This was at a S content of 1%. It is unclear whether this percentage is still to be expected at 0.5 % Sulphur. Please make a note of that. The 6% is rater uncertain and probably depends on several parameters.

Therefore, I think that the exhaust uncertainty should not be discussed in the series of errors because it is not an error. It may cause unexpected differences, but it is not an error. A flawless monitoring method would discover this difference. Discuss this issue in the final comparison

Pag 9. Put the remark on the absence of ships with exhaust cleaning in the sample somewhere else. Perhaps in the conclusion.

Page 10 line 4 English: **ship smoke**: change to the **ships plume**

Page 10 selection method: number 3……No problem ruling out those cases, but I wonder how exhaust uncertainty could explain this. As was mentioned earlier in the text: this is because only 94 % of the Sulphur is oxidized to SO2.

Page 11 first sentence This will make the FSC relatively larger that that of CO2. Unclear what is meant.

Figure 5 This is the core of the paper. I have many comments, but will not all give them:

- Plume 1. There is an additional SO_2 source (and not CO_2 obviously) That is strange. Often SO2 comes from fossil fuel burning. So if the SO2 reses often CO2 rzes as well. Where does the SO2 come from?. Or is it a drift or cross sensitivity in the SO2 sensor? Please comment. Nevertheless, I think the interpretation is OK.
- Plume 2 Same on the SO_2 background. Poor quality expected
- Plume 6 seems good quality. But the maximum value is reached, and it is rejected (?)

General remarks:

- I would consider not using auto scaling of the Y axis to give the reader the chance to compare the different plumes
- The different peaks in the concentrations are related to the distance to the funnel. If you go a bit further away from the funnel perhaps the plumes are more mixed, and no individual peaks are observed and interpretation is easier. Please discuss.
- Although the level of detail is nice to inform the reader some structure could still help. Perhaps make a distinction between: **good quality** examples and **poor-quality** examples and **rejected** plumes, cases with **high Sulphur content**, case with **low** Sulphur content and then don't use autoscaling of the y axis. Also show how well the selection procedure works in the presented cases
- It is stated that the deviation is within 300 ppm is. But it is not entirely clear how the best value is selected. Please discuss the selection of the best value in relation to the error estimates on page 9 (equation 2 and the outcome of that)

Page 13 line 15. Please rephrase to: This could be due to Sulphur emitted in other forms. This is not proven or so in this paper.

- The choice of peaks that are selected in the end remains a bit arbitrary. Picking the highest perhaps leads to the best comparison but why? The authors should defend this choice. Or illustrate what happens when other choices are made
- Perhaps in the table the three peaks could be averaged, and a standard deviation in the calculated Sulphur content could be presented. Just to illustrate how large errors may be.
- And it would be very nice if the error calculated from equation 2 is given for each sample. This could show why difference between measurements by UAV and fuels samples are sometimes larger than other times
- If the background values for CO2 can be used it is perhaps better to carry out a recalibration afterwards. Presenting CO2 concentrations far below 400 is not wanted. Or use an arbitrary scale (?). pleas change.

Page 11 line 10. It is not clear what is meant with **interpolation ratio**

Pag 14 line 13. Significant contribution to the literature. I would refrain from such statements. That is for the readers to decide. Perhaps state that this research is part of a process towards reliable equipment to monitor FSC in ESCA

Page 15 point 2. Is it possible to sate w**hen** errors are large? This could be used to identify poor measurements. And it would be helpful.

I think the battery life is important but a rather detail and perhaps not for the conclusions but somewhere in the technical descriptions

Abstract:

- change the order of first two sentences.

- describe the principles of the method in a few words/lines

- Use **23** samples rather than **more than 20**

Suggestions

Please identify weak points in the current approach and provide suggestions fopr further research and improvements