Answer to Referee #1

We would like to thank Referee #1 for his/her positive and constructive comments and suggestions. We have studied comments carefully and made corrections, which we hope meet with approval. Comments and responses are listed as follows. In order to facilitate the reference to the questions and proposed changes, we use the following color coding:

Color coding:

Referee comment

Our answer

Proposed change in manuscript

This is an interesting paper on an important issue. Sulphur compliance monitoring at sea. The method using drones is rather new and the paper shows details that could help other researchers in this area. Especially the verification measurements presented in the paper and the experiences with the on-board verification attempts are important. Worldwide readers will be interested in the results of the measurements but also in the experiences with on-board inspections. The paper reads well and although it is only a small contribution, I think it deserves publication paying attention to some of the points below. I would like to see a more detailed and perhaps more quantitative treatment of the way the measured data are handled and converted to FSC and when they are rejected. I have only a few small comments on the English (see below).

Thank you for the comments, we are very encouraged.

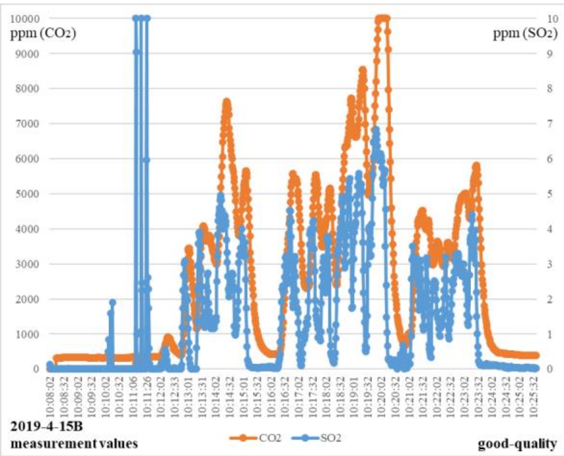
Other more general comments:

- The qualifications poor and good are recognizable. We use that as well in our monitoring. Yet I would like to challenge the authors to come up with a more objective assessment of the quality of each measurement or at least some description as to why some measurements are considered poor. It could be difficult to find objective measures, but it seems needed before the method will become a true enforcement tool.

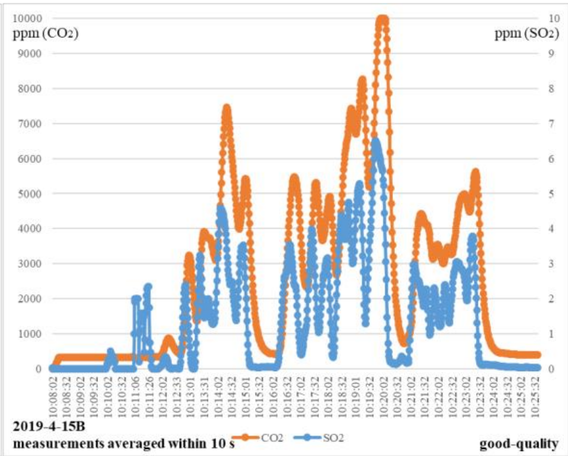
How to objectively and correctly evaluate the quality of plume data is indeed a very important scientific issue, which is also the key to the effective application of the measurement results for maritime law enforcement. I think it would be ideal to design a computational model with input values for plume measurements, output values for FSC results and corresponding confidence levels (or a score that represents the quality of plume data).

This model requires two aspects of work: 1. Adequate an adequate data sets, including gas measurements and the true value of FSC. However, in many cases the true value of the FSC is not available, especially sailing ships. 2. Design and implement the model based on the data sets.

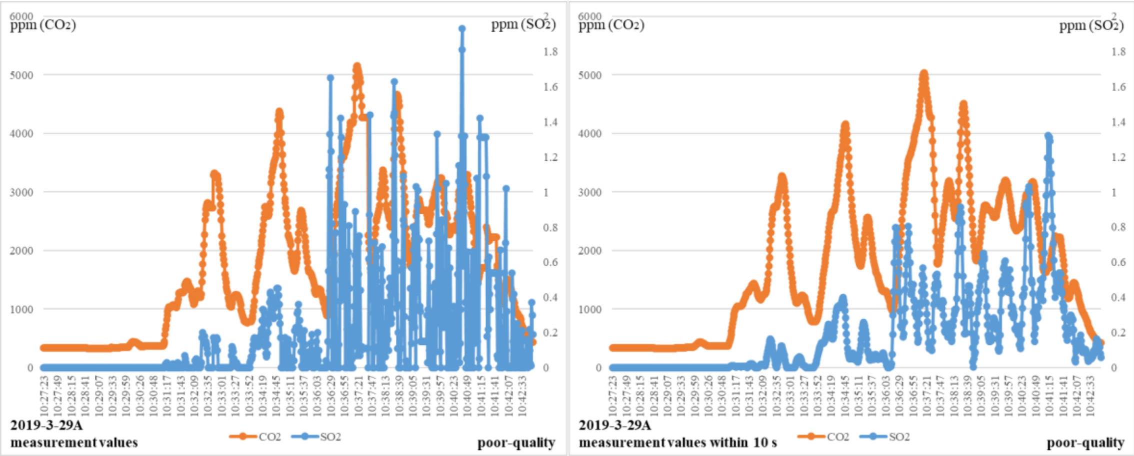
Therefore, there needs a lot of work to implement this model. In fact, that's what we're working on right now. In our current application, however, we assess data quality primarily on the basis of experience. To illustrate how to evaluate data quality, I have supplemented the manuscript with typical good- and poor- data.



(a)



(b)



(c) (d)

Figure 5. Typical measurement data for SO_2 and CO_2 concentrations, and their corresponding average values within 10 s. (a) and (b) good-quality data from plume ID 2019-4-15B. (c) and (d) poor-quality data from plume ID 2019-3-29A. There are some errors in the measurements from 10:11:06 to 10:12:02 in (a), which may have been caused by sensor uncertainty. These data were ruled out and did not affect the calculation results.

- The S-content is now derived from the "fluctuating" signal presented in figure 5. It is not entirely clear why the 10 s averaged data would lead to a better result. I can hardly see the difference. Please show why this is better.

I have added description and discussion in the manuscript as bellow.

The average gas concentration within 10 s was chosen for the FSC calculations; however, this does not mean that 9 s or 11 s could not have been selected. To demonstrate this, a comparison calculation was carried out using both 9 s and 11 s, which showed that these led to very little differences in the results. However, it is necessary to ensure that the gradient of the gas measurements is stable within the sampling time (the interval length of the integral). Moreover, the interval length cannot be too short (e.g., 2 s) or too long (e.g., 20 s). If the time is too short, it is difficult to determine whether the measurements are stable and undisturbed over time. Similarly, if the time is too long, it is also difficult to ensure that all of the measurements in the integral interval are stable and undisturbed. In addition, during the flight of the UAV in this study, the time available for measuring the plume was ~5 minutes. As both the ship and the UAV were moving at this time, it was virtually impossible to ensure that the UAV was flying consistently within the plume and obtaining stable measurements. Accordingly, 10 s is also a relatively appropriate value for the measurement process.

– It should be noted that the signal is not noisy but simply reflects the incomplete mixing of the exhaust gas with clean air. The peaks represent the air that is exhausted from the funnel i.e. only in the peaks you will find the ratio between SO2 and CO2 that is a direct measure of the fuel composition. The highest peak is probably the best choice but could still be a result of mixing of clean air with exhaust air and lead to bias in the result. Or is this negligible? I would welcome a discussion showing that the peak height is a good measure.

According to my understanding, the key problem is how to ensure that the gas being measured is evenly mixed. Or, the selected peaks of SO_2 and CO_2 are measured from the complete mixing of the exhaust gas with clean air. And then the ratio between SO_2 and CO_2 can be used to calculate the fuel composition. Assuming that the gas is mixed complete, the variation trend of SO_2 and CO_2 measurements should be the same (given that the corresponding time of the sensor is not consistent, there may be some deviation), and this trend can be easily identified in the peak area.
But not all peaks can be used to calculate the FSC. Therefore, in the previous study, we developed a selection process. The measurements from incomplete mixing have been ruled out. Meanwhile, the measurements with errors have also been ruled out. The maximum values are likely to have been measured in the center of the ship's plume. At that location, the measurement value is relatively stable, and the probability of interference from other factors is lower. At the same time, the higher the peak value, the greater the proportion of exhaust gas, so the impact from the incomplete mixing of

the exhaust gas with clean air is smaller.

To sum up, the obvious and stable maximum peak appears in the measured value over 10 s periods as the calculated value is a more appropriate choice. There are, of course, cases where multiple similar peaks can occur simultaneously. At this time, their calculations may be very similar, in which case, the results obtained by the calculation of the highest peak should have high credibility. However, it is difficult to describe the problem quantitatively and further research is needed. The ideal model that I mentioned above need a lot of work to do. I have added description and discussion in the manuscript as bellow.

Nevertheless, there is also some uncertainty associated with choosing the peak values. After ruling out the peak values across the full range as well as those corresponding to dramatic changes, the global maximum values were selected as the peak values to calculate the FSC. The maximum values probably correspond to the measurements taken in the center of the ship's plume. At that location, the measurement values were relatively stable, and the probability of interference from other factors was lower. Furthermore, the higher the peak value is, the greater the proportion of exhaust gas is; hence, the impact from the incomplete mixing of the exhaust gas with clean air is relatively small.

In summary, the obvious and stable maximum values are selected as peak values to calculate the FSC. There are, of course, situations where multiple similar peaks can occur simultaneously. In this case, their calculated FSCs may be very similar, and the results obtained by the calculation of the highest peak should have high credibility, for instance, the measurements of plume 2019-4-15B.

I also added the selection of peak values of Figure 5 in the Table 2 as bellow.

Table 2: All peak values and their corresponding FSC results. The background values of plume 2019-4-15B were 0 ppm and 310 ppm for SO_2 and CO_2, respectively. The background values of plume 2019-3-29A were 0 ppm and 329 ppm for SO_2 and CO_2, respectively. The remarks indicate the reason for choosing or not choosing the peak. It can be seen that the peak value of plume 2019-4-15B was more obvious and that the results obtained by multiple alternative peaks were similar. The peak of plume 2019-3-29A was less obvious and there were fewer alternative peaks. This was also the basis for distinguishing data as being of a "good"/"poor" quality. The FSC result of selected peak values are marked as "√".

Plume ID	Time point	Peak value of SO_2 and CO_2 (ppm)	Estimated value of FSC (% (m/m))	True value of FSC (% (m/m))	Remark
2019-4-15B	10:12:52	2.406, 2020	0.326	0.168	Reject; less obvious peak values
	10.13.23	3.235, 2372	0.364		
	10.14.07	4.594, 4665	0.245		Non-maximum peaks of alternative peak values
	10.14.57	3.529, 4872	0.179		
	10.16.39	3.549, 4444	0.199		
	10:17:27	3.989, 3911	0.257		
	10:18:01	3.159, 4607	0.171		
	10:18:47	4.757, 6895	0.168		
	10:19:11	5.287, 7634	0.167 (√)		Maximum peak of the alternative peak value
	10:19:46	6.515, 8100	0.194		Reject; measurements exceeded the range
2019-3-29A	10:34:41	0.399, 3880	0.026	0.035	Reject, less obvious peak values
	10:35:19	0.258, 2011	0.036		Non-maximum peaks of the alternative peak values
	10:37:15	0.567, 4994	0.028		Reject; less obvious peak values
	10:38:27	0.913, 4022	0.057 (√)		Maximum peak of the alternative peak value
	10:40:37	1.031, 2996	0.090		Reject; error in the measurement data
	10:41:13	1.321, 1700	0.224		

According to my understanding, the above comments are all about the 2.4 Calculation and 2.5 Uncertainties. I have rewritten these parts.

- Why not convert table 1 and table 2 into x-y graphs? Perhaps if combined?

Please note that the estimated (UAV) and true (sampled fuel) values of the FSC from 11 berthing ships are list in table 1. But there are only estimated (UAV) values of the FSC from 27 sailing ships are list in table 2. It seems that table 2 cannot convert into x-y graphs and the table 1 and table 2 cannot be combined.

- Our enforcements contacts tell us that ship owners will normally use fuel with a Sulphur content just below the limit. If I look at all the individual samples, I don't see that. Is there an uncertainty that is missed or are the vessels changing from one fuel to the other at the time of the sampling?

I think it's different in different areas, especially when the regulations are different.
In the DECA of China, the regulations are as follow:
Starting January 1, 2017, the FSC cannot exceed 0.5% (m/m) during berthing, excluding the first hour after arrival and the last hour before departure.
Starting January 1, 2018, the FSC cannot exceed 0.5% (m/m) during berthing.
Starting January 1, 2019, the FSC cannot exceed 0.5% (m/m) for both sailing and berthing ships.
Overall, our FSC monitoring results are shown in Figure 6. It shows that the FSCs of the sailing ships were considerably higher than those of the berthing ships. The FSCs of berthing ships are measured in the year of 2018 and 2019. The FSCs of sailing ships are measured in the year of 2019. On July 15, 2019, it was the first time that a sailing ship had been caught for having failed the FSC regulations in China.
I believe that with the further implementation of the policy, the ship owners will normally use fuel with a Sulphur content just below the limit in the DECA of China.

More specific comments: Abstract: Line 12 Emissions of CO2 and SO2 are not measured if I am correct. S content (S%) is measured.

Yes, this sentence "measure the sulfur dioxide and carbon dioxide emissions from sailing ships" is not appropriate. I guess the S content (S%) is calculated from the concentrations of SO_2 and CO_2. This sentence has been modified as follow.

The present study adopts a monitoring method involving an unmanned aerial vehicle (UAV) that takes off from a patrol boat to measure the concentrations of SO_2 and CO_2 within the plumes of sailing ships.

Line 13: I don't think the costs of this method are presented or discussed explicitly in this paper. The cost of a vessel capable of operating in open sea seems neglected. In our country that is not low cost. This is also rather costly.

Yes, this is rather costly. The patrol boats of the maritime department should cruise in the area regularly (once a week). To keep costs down, we measure the ships plume at the same time. Therefore, the cost is lower compared to aircraft and applicable to maritime department. I have added the discussion in the manuscript.

The method proposed in this study can be used to monitor ship emissions at a comparatively low cost to understand the FSCs of sailing ships in open waters. Although the cost of using patrol boats is not negligible, it is convenient and lower cost for maritime authorities compared with small aircraft.

Line 17 According to the monitoring results: I suggest changing to: Based upon the online monitoring results ....
OK, this sentence has been modified.

Line 69: low cost but doesn't include the cost of sailing

As mentioned above, I have added discussion in the manuscript.

Line 90 precision of 5 % at full range (is 10 ppm) or 0.5 ppm. Is that correct? Please mention. And how high is that compared to the observed values? What is the Sulphur content of the example presented in figure 5? Please add.

Yes, it's relative to the range (0.5 ppm is 5% of 10 ppm full range). According to the comments of Referee #2, I have supplemented the detailed parameter information of the UAS. In the original manuscript, the range of SO_2 sensor and accuracy of CO_2 sensor were wrong. I have modified and cross-checked to make sure the product information was right.

There are different range models of sensors, such as 5 ppm, 10 ppm, 100 ppm and so on. Precision generally depends on the size of the range. For example, precision of 0.25 ppm for 5 ppm, 0.5 ppm for 10 ppm, and 5 ppm for 100 ppm. It needs to make sure that observed values do not exceed the range in most cases, and the precision should not be too low. Therefore, observed values of SO_2 are generally in the range of 0-10 ppm. CO_2 are 400-5000 ppm range.
As mentioned above, I have added Table 2.

Line 90 etc. Could some details or results of the calibration procedure be presented as well.

I have added the details in the manuscript.

These sensor characteristics were provided by the instrument manufacturer and were ensured to be within the tolerances by calibration. The zero and full scales are usually calibrated by a standard mixed gas when the equipment is used on a daily basis. The major parameters of the UAS are listed in Table 1.

Line 107 "measure the concentration of SO2 and CO2". Change to: Measure the concentration of SO2 and CO2 in the plume

OK, this sentence has been modified.

Line 124 EF has no unit or? I wonder why equation 1 is mentioned. I think it is a bit confusing.

EF has unit, g_{SO2}/kg_{fuel}, in g emitted per kg fuel. I have added it in manuscript.
Equation 1 is a description of the measurement principle. We use the parameter of 10s to derive equation 2. If don't discuss equation 1, just list equation 2. It seems too sudden to the reader.

Line 132: What is sampling rate? Electronically? And why are the SO2 and CO2 sensors not synchronized? In the graphs it looks like a delay. And you could just shift them a little. Why is that not done? And why is the 10 s averaged data better. I don't see that. Figure 5 Please use equal y-scales in the right and left panel. This is confusing.

The sampling rate is 1s.
This is due to inconsistent response times of different sensors. In the vast majority of cases, the response time of SO_2 are faster, and the CO_2 looks a little bit behind SO_2.
Of course, we can adjust the timing artificially. For example, adjust the peak time of CO_2 to the peak time of SO_2.
I've done a comparative calculation. The data set is of the estimated (UAV) and true (sampled fuel) values of the FSC

from 11 berthing ships (using second-generation pod). On the whole, the difference is not great and the accuracy was slightly reduced. Therefore, we chose the method described in the manuscript to do the calculation.
The average gas concentration over 10 s was chosen for the FSC calculations; however, this does not mean that 9 s or 11 s could not have been selected. To demonstrate this, a comparison calculation was carried out using both 9 s and 11 s, which showed that these led to very little differences in the results. However, it is necessary to ensure that the gradient of the gas measurements is stable over the sampling time (the interval length of the integral). Moreover, the interval length cannot be too short (e.g., 2 s) or too long (e.g., 20 s). If the time is too short, it is difficult to determine whether the measurements are stable and undisturbed over time. Similarly, if the time is too long, it is also difficult to ensure that all of the measurements in the integral interval are stable and undisturbed. In addition, during the flight of the UAV in this study, the time available for measuring the plume was ~5 minutes. As both the ship and the UAV were moving at this time, it was virtually impossible to ensure that the UAV was flying consistently within the plume and obtaining stable measurements. Accordingly, 10 s is also a relatively appropriate value for the measurement process.
Therefore, 10 s is an appropriate parameter.
I have added the details in the manuscript as mentioned above.

OK, I have added equal y-scales in the right and left panel.

Line 137: The 10 sec averages hardly differ from 1 sec data. How could that happen?

Does this refer to the data in Figure 5?
The data in Figure 5 (b) is relatively smooth, making it easy to select peak values.
The data in Figure 5 is of good-quality and the difference is not significant. This is even more pronounced if the data is of poor-quality. I've added good- and poor- typical figures as mentioned above.

Line 140 What is meant by the calculated function? Not just the average measurement value?

It is "calculate the average of the data within 10 s". This sentence has been modified.

Line 160: What would be objective criteria to tell whether it's a poor-quality plume?

In the previous study, we developed a selection process. If most of the data in the measurement dataset is ruled out, then it is a poor-quality plume. If multiple peaks are preserved and their FSC result are similar, then it is a good-quality plume. But, strictly speaking, we do not have a completely objective evaluation criterion. This requires the computational model I mentioned above. I have added tow typical example in the Figure 5 and Table 2.

Line 167 Is this correct 20% (m/m)? I would expect 20% uncertainty with no units. So I suggest to leave (m/m) out.

Yes, this sentence has been modified.

Line 168 and 169. Unclear what is meant here (after therefore)? It could be interesting for the reader interested in enforcement to mention (even in the abstract) how many of the Non Compliants were detected and how many were missed (i.e. ships that were not identified as Non Compliants but had an FSC above the limit.

Please note that boarding inspection is a complicated process. The target ship cannot stop immediately in the channel for inspection and have to sail to the anchorage. Patrol boats need to follow to the anchorage for boarding inspection. In the process, to avoid punishment, crew will take various measures to drain the high-sulfur fuel in the main engine fuel oil

pipeline.
The whole process of boarding inspection requires more than half a day and the work of more than a dozen law enforcement officers (such as sailors, driver, inspectors, VTS watchmen).
Therefore, in order to ensure that can accurately tracked down the offending ship. Law enforcement officers of the Pudong maritime safety administration only intercepted four sailing ships for which the UAV FSC results were of a good-quality and all exceeded 2% (m/m). The chemical FSC results of the four ships were: 0.534% (m/m), 0.744% (m/m), 0.813% (m/m), and 1.991% (m/m).
If the UAV FSC results is just over 0.5%, and then we board the ship for inspection. I believe the Non Compliants and Missed will happen. However, the factor of switching fuel cannot be ignored. Intercept and boarding a sailing ship are not just an experiment.

Line 189: perhaps the word optimistic is not the right word. Perhaps it should be: The uncertainty in the assessment is not small but the results so far, do not lead to optimism with respect to the FSC used by ships sailing in the area.

OK, this sentence has been modified.

Line 210: It Is only a small sample isn't it, but still convincing looking at figure 6. Please mention that.

OK, I have added the this in the manuscript.

Page 15: Number of digits in the given numbers are large (such as 40913 ppm, I suggest changing that to 4.1 %)

OK, these words have been modified.

Page 15 Why are the results of the sampling by the maritime authority not given in the table.

OK, I have added this.

English:
Line 15 ships → vessels
The word "ships" or "vessels" seems all appropriate. For consistency, "ships" is used all over the text.
Line 46 Several studies have suggested monitoring methods or similar. Otherwise I understood wrong but then the sentence is not very clear.
I have changed the sentence "Several studies have suggested monitoring ship emissions to estimate the FSC of the target ship" as "Several studies have suggested estimating FSC by measuring ship plumes".
Line 45 and 66 supervise or supervision→ enforce and enforcement; Line 54 and 65 navigation → navigating (?); Line 60: airplace→ aircraft; Line 65 inaccurate → non representative; Line 67 Suggestion: leave therefore out Line 85 extracts gas? → draws air; Line 108: approximately a few hundred meters. This is double: a few hundred meters is already an expression showing that it is an approximate value. Suggest leaving the word approximately out.; Page 13 legends to figure 4: the enlarged UAV is shown in the top left corner (and it is in the right corner: a detail)
The above English grammar problems have been revised. Thank you very much for your earnest help in pointing out the English problems.
NB: I have two versions of the text. Also, one with an appendix including two figures: Figure A2 The Chinese text could be difficult to read for non-Chinese readers. Perhaps add some explanation of the Chinese text.
As the suggestion of the Associate Editor, "remove the figures from the appendix as they do not contribute to the

discussion of the measurement techniques”. I have removed the Figure A1 and A2.

Answer to Referee #2

We would like to thank Referee #2 for his/her positive and constructive comments and suggestions. We have studied comments carefully and made corrections, which we hope meet with approval. Comments and responses are listed as follows. In order to facilitate the reference to the questions and proposed changes, we use the following color coding:

Color coding:

Referee comment

Our answer

Proposed change in manuscript

The paper by Zhou et al. reports on fuel sulfur content (FSC) compliance monitoring of sailing ships with unmanned aerial vehicles (UAV). Measurements were carried out in the Yangtze River Delta close to Shanghai in China, which is selected as an domestic emission control area (DECA) in China having a FSC limit of 0.5% (m/m). Since measurements of the FSC and therefore compliance monitoring from sailing ships are sparse, the topic is of interest not only for the scientific community. The manuscript is in general clearly written and I recommend it for publication in AMT. However, to better demonstrate the quality of the instrumentation and the methods used in this study more details should be given in the paper.

Thank you for the comments, we are very encouraged.

Instrumentation:

- There are no details on the custom sensors for SO2 and CO2 given. At least a link to a data sheet (in English) or better a table with specifications is needed.

OK, I have added a detail table (Table 1).

- I am wondering about the short response time of the modules (T90< 1s) which is much better than every electrochemical sensor I know. Looking to figure 5 this response time is very unlikely. SO2 and CO2 measurements of the same plume are out of phase (at least 10 to 15 s) having also completely different gradients.

We don't make sensors, and sensors were purchased from Shenzhen Singoan Electronic Technology Co., Ltd., China. I looked up the relevant materials (Mellqvist et al., 2017), the t_{90} of CO_2 sensor based on NDIR (LI-COR 7200) is 0.1s. I also consulted other relevant literature (Alföldy et al., 2013, Beecken et al., 2014, Balzani Lööv et al., 2015), the t_{90} of CO_2 sensor is about < 1-5 s. The CO_2 sensor used by us is also base on NDIR principle. It seems reasonable that the CO_2 response time is less than 1 s. However, it can be clearly seen from Figure 5 that the response time of SO_2 is significantly faster than that of CO_2. Does this mean that the t_{90} of SO_2 sensor is also <1s?

To ensure the accuracy of parameter information, I contacted one technical support of Shenzhen Singoan Electronic Technology Co., Ltd., China. I double confirm the relevant parameter information of the sensor. I got the detailed information sheet of these two sensors. The response time (defined as T_{90}) is <30 s. "The t_{90} represents the time taken to reach 90% of the stable response following a step change in the sample concentration". I guess "a step change" does not mean from 0 to full range. If it is from 0 to full range, the response time (T_{90}) is <30 s. The information is list as follow.

Figure picture of the sensor

The type of SO_2 sensor is SGA-700A-SO_2. Link: http://www.singoan.com/article/detail/id/6233.htm

The type of CO_2 sensor is SGA-700A-CO_2. Link: http://www.singoan.com/SGA_400_700_CO2.htm

I fill the main information into Table 1 as bellow. In the original manuscript, the range of SO_2 sensor and accuracy of CO_2 sensor are wrong. I have modified and cross-checked to make sure the information was consistent as the product information.

Table 1: Parameters of the UAS

	Parameter	Value
UAV	Symmetrical motor wheelbase	1133 mm
	Size	1668 mm × 1518 mm × 727 mm
	Weight	9.5 kg
	Recommended maximum take-off weight	15.5 kg
	Hovering accuracy(P-GPS)	Vertical: ±0.5 m, Horizontal: ±1.5 m
	Maximum rotational angular velocity	Pitch axis: 300°/s, Heading axis: 150°/s
	Maximum pitch Angle	25°
	Maximum rising speed	5 m/s
	Maximum rate of descent	3 m/s
	Maximum sustained wind speed	8 m/s
	Maximum horizontal flight speed	65 km/h (no wind environment)
	Hover time	Non-loaded: 32 min; load 6 kg: 16 min
SO_2 sensor	Type	SGA-700A-SO2
	Principle	Electrochemistry
	Measuring range	0–10 ppm
	Diameter and height	33.5 mm; 31 mm
	Weight	30 g
	Accuracy	≤ ±3 % (0.3 ppm)
	Linear error	≤ ±2 % (0.2 ppm)
	Repeatability	≤ ±2 % (0.2 ppm)
	Power consumption	≤ 50 mA
	Response time (T_{90})	≤ 30 s
CO_2 sensor	Type	SGA-700A-CO2
	Principle	Non-Dispersive InfraRed
	Measuring range	0–10000 ppm
	Diameter and height	33.5 mm; 31 mm
	Weight	30 g
	Accuracy	≤ ±3 % (300 ppm)
	Linear error	≤ ±2 % (200 ppm)
	Repeatability	≤ ±2 % (200 ppm)
	Power Consumption	≤ 100 mA
	Response time (T_{90})	≤ 30 s

Alföldy, B., Lööv, J. B., Lagler, F., Mellqvist, J., Berg, N., Beecken, J., Weststrate, H., Duyzer, J., Bencs, L., Horemans, B., Cavalli, F., Putaud, J.-P., Janssens-Maenhout, G., Csordás, A. P., Van Grieken, R., Borowiak, A., and Hjorth, J.: Measurements of air pollution emission factors for marine transportation in SECA, Atmos. Meas. Tech., 6, 1777–1791, https://doi.org/10.5194/amt-6-1777-2013, 2013.

Balzani Lööv, J. M., Alfoldy, B., Gast, L. F. L., Hjorth, J., Lagler, F., Mellqvist, J., Beecken, J., Berg, N., Duyzer, J., Westrate, H., Swart, D. P. J., Berkhout, A. J. C., Jalkanen, J.-P., Prata, A. J., van der Hoff, G. R., and Borowiak, A.: Field test of available methods to measure remotely SOx and NOx emissions from ships, Atmos. Meas. Tech., 7, 2597–2613, https://doi.org/10.5194/amt-7-2597-2014, 2014.

Mellqvist, J., Conde, V., Beecken, J., and Ekholm, J.: Certification of an aircraft and airborne surveillance of fuel sulfur content in ships at the SECA border, CompMon (https://compmon.eu/, last access: 6 November 2018), 2017b.

Beecken, J., Mellqvist, J., Salo, K., Ekholm, J., and Jalknen, J.-P.: Airborne emission measurements of SO2, NOx and particles from individual ships using a sniffer technique, Atmos. Meas. Tech., 7, 1957–1968, http://dx.doi.org/10.5194/amt-7-1957-2014, 2014.

- Please give more information on the method water vapor was filtered out. What about contamination with particles/soot ...?

The pod is a convenient, lightweight device. It does not have the same complex gas filtration as shore-based equipment. A hose filter valve was used to filter out the water vapor, particles and soot. Its figure is as follow.

Figure Different type of the hose filter valve, the length is about 4-20mm

- Have the authors investigated cross-sensitivities to e.g. NO and NO2?

For the sake of lightweight and convenience, the second-generation pod is only equipped with SO_2 and CO_2 sensors. In the research of Mellqvist et al. (2017), they proposed a treatment for cross-sensitivities SO_2 and NO_2. But their sensor is based on the principle of fluorescence (Thermo 43i-TLE), and our sensor is based on electrochemistry. I have looked up the relevant materials, the cross-sensitivities of SO_2 and NO_2 do exist when using the electrochemical sensor. I guess more experiments and researches are needed to eliminate the effects of cross-induction.

Mellqvist, J., Conde, V., Beecken, J., and Ekholm, J.: Certification of an aircraft and airborne surveillance of fuel sulfur content in ships at the SECA border, CompMon (https://compmon.eu/, last access: 6 November 2018), 2017.

- Measurements were carried out close to the funnel. Therefore, temperatures of the air sucked into the system are highly variable. How this is accounted for?

The UAS has a gas pump, gas circuit, filter which can cool the gas to a certain extent.
I know that the equipment used in relevant research work has a constant temperature and pressure detecting environment. The weight of the equipment is usually tens of kilograms. But there are limits to the weight and size of the equipment that UAV can carry.
The conversion relation of units is as follow:

$$\mathbf{X}\,(mg/m^3) = \frac{\mathbf{M}}{22.4} * \mathbf{Y}\,(ppm) * \left[\frac{273}{273+\mathbf{T}}\right] * (\mathbf{Ba}/101325)$$

M is molecular weight of the gas, T is the temperature, and Ba is air pressure.
I guess it has to redesign the pod and laboratory experiments to reduce the impact of this factor.

- Give information on the calibration methods used in this study to ensure long-term data quality. Since sensors used in this study are completely different to those used in Zhou et al. 2019 a simple plot showing the outcome of the sensors when using standard gas mixtures (e.g. 5 and 0 ppm SO2) would be nice.

The calibration process is the same, I have added description.

The details of calibration are as follows: We bought the sensors and then designed and built the pod. Then we send the pod to a third-party inspection agency for certification. Content of verification is mainly about the accuracy. We only know the results and we don't have detailed data. I know that you may interest in the sensors. But the information available to me is limited. I have attached the validation report at the end (the original is in Chinese).

- The UAV used in this study is the same as in Zhou et al., 2019. What does it mean for the off-shore measurements? What is the operation time under typical weather conditions having e.g. a wind speed of 5 m/s? What is the maximum reasonable distance to a sailing ship? Please add a table with specifications of the whole UAV system.

The UAV is a product of DJI and is a relatively common UAV model. We have successfully applied it offshore measurements. But there are still have limitations, it is more affected by the weather compare with other measurement platform. For safety reasons, we can't use it when it rains or when the wind is high. Nevertheless, this is basically the best and most suitable civilian UAV that we can find on the market.
About 15-20 min.
This question is hard to say. We did not measure the distance by instruments. At sea, the human sense of distance is very weak. My personal feeling is about 500 to 3000 meters.
I have added the Table 1.

Methods and uncertainties:
- The authors used peak values of SO2 and CO2 after applying a running mean of 10s to the measurements to calculate the FSC. Looking to Figure 5 b it is not clear to me, how this could give reasonable results. As already mentioned above the gradients (and therefore real response times of the sensors) look completely different even for averaged values. At least one example proving and illustrating this method is needed.

The data in Figure 5 is of good-quality and the difference is not significant. This is even more pronounced if the data is of poor-quality. I've added good and poor typical figures and related discussion. To illustrate how to evaluate data quality, I have supplemented the manuscript with typical good- and poor- data, and the relevant description as follow.

The continuous measurement data for two typical plumes (2019-4-15B and 2019-3-29A) are exhibited in Fig. 5. The data for plume 2019-4-15B (Fig. 5a) were considered to be of a "good" quality, whereas those for plume 2019-3-29A (Fig. 5c) were considered to be of a "poor" quality. Data were determined to be of a good-quality when obvious, easily distinguished peak values were observed, whereas less obvious peaks that still corresponded to a result were considered as poor-quality data. The selection of peak values leads to uncertainty because when the area ratio is selected for the calculation, the starting and ending time points of the area are still associated with substantial uncertainty. Figure 5b and 5d depict the average concentrations of the SO_2 and CO_2 measurements (in Fig. 5a and 5c, respectively) for 10 s periods. The peak value of each average concentration was selected for the calculation. This process is equivalent to selecting the area ratio of SO_2 to CO_2 within 10 s for the calculation, as shown in Eq. (2).

$$FSC[\%] = 0.232\frac{\int(SO_{2,peak}-SO_{2,bkg})dt[ppb]/10}{\int(CO_{2,peak}-CO_{2,bkg})dt[ppm]/10}[\%] + R \approx 0.232\frac{AVG(SO_{2,peak})-AVG(SO_{2,bkg})}{AVG(CO_{2,peak})-AVG(CO_{2,bkg})}[\%], \quad (2)$$

where $AVG(\cdot)$ is the calculated function for the average measurement value within 10 s; hence, the data in this study are the average values of measurements in 10 s. When the UAV took off from the patrol boat and flew high into the air, the SO_2 and CO_2 concentrations were relatively low. The background values were obtained at this stage as the minimum SO_2 and CO_2 concentrations. As the UAV flew into the plume, the measured concentrations of SO_2 and CO_2 increased. The obvious, stable maximum values in the observations of the average measurement values should be selected as the peak values. It can be seen that using the average values of measurements within 10 s makes it easier to select the peak values, especially with respect to poor-quality data. However, as there can still be several options for peak values, the data treatment methods reported by Zhou et al. (2019) were incorporated in this study to select the most appropriate peak values. In Fig. 5b, the time point of selected peak values is at 10:19:11. The measurement values from 10:19:57 to 10:20:15 were not used because the CO_2 concentration covered the full range. In Fig. 5d, the time point of the selected peak values is at 10:38:27. The measurement values from 10:39:57 to 10:41:41 were not used because we ruled out data exhibiting either dramatic changes or errors in continuous observations. The details for selecting the peak values are listed in Table 2.

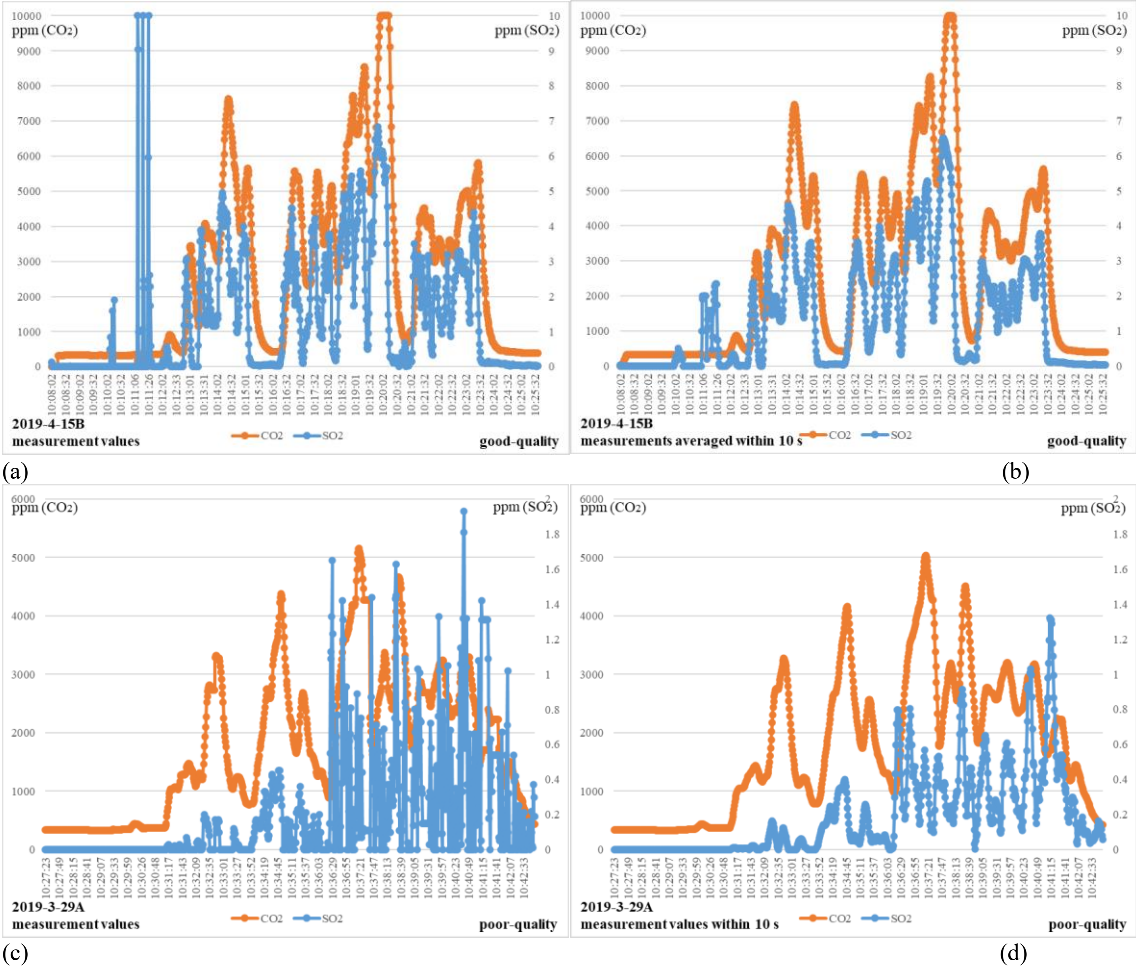


Figure 5. Typical measurement data for SO_2 and CO_2 concentrations, and their corresponding average values within 10 s. (a) and (b) good-quality data from plume ID 2019-4-15B. (c) and (d) poor-quality data from plume ID 2019-3-29A. There are some errors in the measurements from 10:11:06 to 10:12:02 in (a), which may have been caused by sensor uncertainty. These data were ruled out and did not affect the calculation results.

2.5 Uncertainties

In previous research (Zhou et al., 2019), the main uncertainties of UAV measurements were summarized as sensor uncertainty, measurement uncertainty, calculation uncertainty, and exhaust uncertainty. The instrument calibration method, UAV flight procedures, and data treatment methods were designed to reduce these uncertainties. However, some uncertainties remain, as discussed below.

The average gas concentration within 10 s was chosen for the FSC calculations; however, this does not mean that 9 s or 11 s could not have been selected. To demonstrate this, a comparison calculation was carried out using both 9 s and 11 s, which showed that these led to very little differences in the results. However, it is necessary to ensure that the gradient of the gas measurements is stable within the sampling time (the interval length of the integral). Moreover, the interval length cannot be too short (e.g., 2 s) or too long (e.g., 20 s). If the time is too short, it is difficult to determine whether the measurements are stable and undisturbed over time. Similarly, if the time is too long, it is also difficult to ensure that all of the measurements in the integral interval are stable and undisturbed. In addition, during the flight of the UAV in this study, the time available for measuring the plume was ~5 minutes. As both the ship and the UAV were moving at this time, it was virtually impossible to ensure that the UAV was flying consistently within the plume and obtaining stable measurements. Accordingly, 10 s is also a relatively appropriate value for the measurement process.

Nevertheless, there is also some uncertainty associated with choosing the peak values. After ruling out the peak values across the full range as well as those corresponding to dramatic changes, the global maximum values were selected as the peak values to calculate the FSC. The maximum values probably correspond to the measurements taken in the center of the ship's plume. At that location, the measurement values were relatively stable, and the probability of interference from other factors was lower. Furthermore, the higher the peak value is, the greater the proportion of exhaust gas is; hence, the impact from the incomplete mixing of the exhaust gas with clean air is relatively small.

In summary, the obvious and stable maximum values are selected as peak values to calculate the FSC. There are, of course, situations where multiple similar peaks can occur simultaneously. In this case, their calculated FSCs may be very similar, and the results obtained by the calculation of the highest peak should have high credibility, for instance, the measurements of plume 2019-4-15B.

Table 2: All peak values and their corresponding FSC results. The background values of plume 2019-4-15B were 0 ppm and 310 ppm for SO_2 and CO_2, respectively. The background values of plume 2019-3-29A were 0 ppm and 329 ppm for SO_2 and CO_2, respectively. The remarks indicate the reason for choosing or not choosing the peak. It can be seen that the peak value of plume 2019-4-15B was more obvious and that the results obtained by multiple alternative peaks were similar. The peak of plume 2019-3-29A was less obvious and there were fewer alternative peaks. This was also the basis for distinguishing data as being of a "good"/"poor" quality. The FSC result of selected peak values are marked as "√".

Plume ID	Time point	Peak value of SO_2 and CO_2 (ppm)	Estimated value of FSC (% (m/m))	True value of FSC (% (m/m))	Remark
2019-4-15B	10:12:52	2.406, 2020	0.326	0.168	Reject; less obvious peak values
	10.13.23	3.235, 2372	0.364		
	10.14.07	4.594, 4665	0.245		Non-maximum peaks of alternative peak values
	10.14.57	3.529, 4872	0.179		
	10.16.39	3.549, 4444	0.199		
	10:17:27	3.989, 3911	0.257		
	10:18:01	3.159, 4607	0.171		
	10:18:47	4.757, 6895	0.168		
	10:19:11	5.287, 7634	0.167 (√)		Maximum peak of the alternative peak value
	10:19:46	6.515, 8100	0.194		Reject; measurements exceeded the range
2019-3-29A	10:34:41	0.399, 3880	0.026	0.035	Reject, less obvious peak values
	10:35:19	0.258, 2011	0.036		Non-maximum peaks of the alternative peak values
	10:37:15	0.567, 4994	0.028		Reject; less obvious peak values
	10:38:27	0.913, 4022	0.057 (√)		Maximum peak of the alternative peak value
	10:40:37	1.031, 2996	0.090		Reject; error in the measurement data
	10:41:13	1.321, 1700	0.224		

I have rewritten these parts of 2.4 Calculation and 2.5 Uncertainties to illustrate the problem and related problem.

- Simply taking into account the given accuracies for the sensors (5 and 3 % on measuring ranges respectively) an error of at least 0.03 % (m/m) can be calculated. Within this calculation no other errors e.g. due to the measurement procedure are included. Therefore, the reported total uncertainty of 0.03 for low FSC levels sounds quite optimistic to me.

According to the literature available, the main method to measure the ship plume are land-based and airborne-based method. UAV measurements are indeed more accurate than these approaches. I guess this is mainly because UAV measurements are taken at close range. However, please note that this accuracy is the measurement result of the berthing ships. One is unable to obtain samples of fuel from sailing ships normally. We attempted to measure more than 40 ship plumes in open water; however, only 27 of them resulted in good- or poor-quality data, i.e., usable data. The success rate is not very high.

- Please give details or an illustration what is meant with good- and poor-quality data.

As mentioned above, I have added good- and poor-quality.

Minor corrections:
- 2.1. Instrumentation: I guess, dimensions of the pod are given in cm.

Yes, you are right, it is cm. Thank you for pointing out this error.

- Figure 5: More details needed, which values are used for the calculation of the FSC (see above)

As mentioned above, I have added.

- Tables 1 and 2: Please always give FSC Values in % (m/m) using only the number of meaningful digits. "True value" implies that the analysis of fuel samples have no error which is of course not the case. Please refer to e.g. fuel sample analysis and give the typical error for this method (should be roughly 0.01 % (m/m))

Yes, I have modified the unit.
Maritime authorities send fuel samples to third-party testing institutions, and the test results can be used as a basis for law enforcement. I don't have the information about the error of their detection methods. But I guess the accuracy of direct fuel detection is definitely far higher than the FSC result estimated from gas measurement. It is appropriate to take it as "True value".

- Table 2: Add values for the fuel sample analysis when available.

OK, I have added.

The following is the third-party inspection report of the pod.

上海市计量测试技术研究院
华东国家计量测试中心
中国上海测试中心

检 测 报 告

Test Report

委托者 Customer	上海安馨信息科技有限公司
委托者地址 Address of customer	上海临港新城海基六路218弄13号楼3楼
样品名称 Name of sample	船舶尾气检测吊舱
制造厂 Manufacturer	上海安馨信息科技有限公司
型号/规格 Model/Specification	AX-YD
样品编号 No. of sample	1145003

（机构检测专用章）

批准人/职务 Approved by / Functions	郝玉红 质量主管
核验员 Checked by	核验员标记
检测员 Tested by	检定员标记

检测日期 Date for test　2019 年 Year　03 月 Month　11 日 Day

地址：上海市张衡路1500号(总部) Address No.1500 Zhangheng Road，Shanghai(headquarters)　电话：021-38839800 Tel.　传真：021-50798390 Fax　邮编：201203 PostCode
客户咨询电话：800-820-5172 Inquire line　投诉电话：021-50798262 Tel. for complaint

国家法定计量检定机构计量授权证书号(中心/院):(国)法计(2017)01039号/(2017)01019号
The number of the Certificate of Metrological Authorization to The Legal Metrological Verification Institution is No.（2017）01039/ No.（2017）01019

本次检测所依据的技术规范（代号、名称）：
Reference documents for the test (code 、name)

参照JJG 635-2011《一氧化碳、二氧化碳红外气体分析器检定规程》
参照JJG 551-2003《二氧化硫气体检测仪检定规程》

本次检测所使用的主要测量仪器：
Main measuring instruments used in this test

名称 Name	型号规格 Model	编号 Number	测量范围 Measurement range	不确定度或准确度等级或最大允许误差 Uncertainty/Accuracy Class/Maximum Permissible Error	证书编号/有效期限 Certificate No./Due date
氮中二氧化碳气体标准物质	/	770357/201803	4.97×10^{-2}mol/mol	U_{rel}=2.0%（k=2）	770357/2019-03-29
氮中二氧化硫标准气	/	L00502040	100×10^{-6}mol/mol	U_{rel}=2.0%（k=2）	L00502040/2019-03-29
精密气体稀释仪	MGB1000	06132	稀释比：1~1000	±3%	2019I30-30-1727259001/2020-02-14
/	/	/	/	/	/

检测地点及环境条件：
Location and environmental condition for the test

地点： 张衡路1500号理化东楼125室
Location

温度： 21℃
Ambient temperature

湿度： 63%RH
Relative humidity

其它： /
Others

备注： /
Note:

本报告提供的结果仅对本次被测的样品有效。
The data are valid only for the sample(s).

检测结果/说明：
Results of test and additional explanation

委托日期	2019.03.07	样品状态描述	正常	受样方式	客户送样

检测项目	标准气体浓度（μmol/mol）	仪器显示值（μmol/mol）
CO_2	10.0	9.84
	900	924
SO_2	1.0	0.96
	8.0	8.21

检测结果内容结束

Monitoring compliance with fuel sulfur content regulations of sailing ships by unmanned aerial vehicle (UAV) measurements of ship emissions in open water

Fan Zhou[1], Liwei Hou[2], Rui Zhong[1], Wei Chen[3], Xunpeng Ni[3], Shengda Pan[1], Ming Zhao[1,4], Bowen An[1]

[1]College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China
[2]College of Ocean Science and Engineering, Shanghai Maritime University, Shanghai 201306, China
[3]Pudong Maritime Safety Administration of the People's Republic of China, Shanghai 200137, China
[4]Key Laboratory of Intelligent Infrared Perception, Chinese Academy of Sciences, Shanghai 200083, China

Correspondence to: Fan Zhou (fanzhou_cv@163.com)

Abstract. Due to technical and cost limitations, the monitoring of emissions from ships sailing in open water within the ship emission control areas (ECAs) is relatively rare. The present study adopts a monitoring method involving~~that uses~~ an unmanned aerial vehicle (UAV) that takes off from a patrol boat to measure the concentrations of SO_2 and CO_2 within the plumes of~~sulfur dioxide and carbon dioxide emissions from~~ sailing ships. Our method aims to provide a low-cost, remote approach for estimating the fuel sulfur content (FSC) of sailing ships in open water, which overcomes the limitations of ground-based and small aircraft methods. The selected monitoring area was the Yangtze River estuary, a domestic ECA with an FSC limit of 0.5% (m/m) implemented by the Chinese government. A total of 27 sailing ships were monitored, 14 of which were found to have an FSC of > 0.5% (m/m). Moreover, the FSCs of the sailing ships were found to be higher than those of berthing ships in the study area. Based upon the online monitoring results~~According to the monitoring results~~, four of the monitored ships were intercepted by the maritime law enforcement, and fuel samples were collected and analyzed in a laboratory; the results confirmed that all four FSCs were > 0.5% (m/m). Among them, one offending ship was tracked down on July 15, 2019, which was the first time that a sailing ship had been caught for having failed the FSC regulations in China. Overall, the present study provides scientific support for evaluating the effectiveness of ECA policies, and recommends that emissions from sailing ships should be monitored more often in the open water in the future.

1. Introduction

With the rapid development of the shipping industry (UNCTAD, 2017) over the past decades, air pollution caused by ship emissions has received an increasing amount of attention (Eyring et al., 2010; Wan et al., 2016). The pollutant gases emitted by ships not only affect the global climate (Huebert, 1999; Corbet, 2016), but also local air quality and can harm people's health. (Yang et al., 2016; Wang et al., 2019). Shipping accounted for 15%, 13%, and 3% of the annual global anthropogenic emissions of NOx, SOx, and CO_2 from 2007 to 2012, respectively (Smith et al., 2014). In Europe, estimated ship emissions were responsible for 3.0 million tons of NOx, 1.2 million tons of SOx, and 0.2 million tons of fine particulate matter (PM2.5) in 2011 (Jalkanen et al., 2016). In East Asia, shipping emissions accounted for 16% of global shipping CO_2 in 2013, whereas they only accounted for 4–7% during 2002–2005 (Liu et al., 2016).

To reduce the negative impacts of ship emissions, the International Maritime Organization (IMO) regulates emissions through the International Convention for the Prevention of Pollution from Ships and its Annex VI (MARPOL, 1997). The air-pollution limits for shipping were adopted in 1997, but only came into force in 2005. The global cap for the fuel sulfur content (FSC) of seagoing ships was set at 3.5% (m/m) in 2012, and was reduced to 0.5% (m/m) in 2020. To date, four emission control areas (ECAs) (the Baltic Sea, the North Sea, the United States Caribbean, and the North American and United States Caribbean Sea) have been set up, and the corresponding FSC limit for seagoing ships in these areas was set at 0.1% (m/m) in 2015 (IMO, 2017).

The IMO has not yet set up ECAs in East Asia, which includes the world's ten largest container ports, for example, Shanghai, Ningbo-Zhoushan, and Shenzhen ports. To limit the air pollution caused by ship emissions, the Chinese government established three domestic emission control areas (DECAs) in 2015: the Yangtze River delta, the Pearl River delta, and the Bohai Sea. DECAs was expanded to cover a wider area since 2020, and include most of the coastal ports, the Yangtze River main line, and the Xijiang River main line. The FSC limit for sailing and berthing ships in the DECAs has been set at 0.5% (m/m) since January 1, 2019.

A key problem regarding the implementation of the policy of the ECAs is the question of how to enforce~~supervise~~ the FSC of ships. Several studies have suggested estimating FSC by measuring ship plumes~~monitoring ship emissions to estimate the FSC of the target ship~~ (Berg et al., 2012; Balzani Lööv et al., 2014). At present, the main method to monitor the emissions of surrounding ships is to place monitoring equipment either on the wharf, shore, port area, or bridge (i.e., ground-based methods) (Alföldy et al., 2013; Pirjola et al., 2013; Beecken et al., 2015; Kattner et al., 2015; Mellqvist et al., 2017a; Cheng et al., 2019; Zhang et al., 2019). Although ground-based methods can provide continuous monitoring, the results obtained depend on the wind speed, wind direction, and the relative position of a ship to the monitoring equipment. Additionally, the boundaries of the ECAs that are designated by the IMO are 200 nautical miles from the coast (Viana et al., 2015); hence, ground-based methods are not able to monitor the fuel that is used on the open sea in ECAs because sailing ships are too far from the shore or bridges.

Therefore, some researchers have used sensors that are carried by small aircrafts to monitor navigating~~navigation~~ ships within ECAs (Berg et al., 2012; Beecken et al., 2014). However, because this kind of monitoring method is costly, the monitoring of navigation ships is relatively rare. Beecken et al. (2015) observed 434 plumes during ground-based measurements and 32 plumes from a helicopter. Balzani Lööv et al. (2014) took 475 measurements using “sniffing” instruments from ground-based measurements, whereas only 25 measurements were obtained using this method from mobile platforms. In the study undertaken by Mellqvist et al. (2017b), 114 individual ships were measured effectively during 27 flight hours at a cost of approximately 470 Euro per ship, which was for the aircraft~~airplace~~ cost and did not included the ferry, operator, or instrument rental costs. Therefore, the high cost of flying precludes extensive monitoring of ship emissions.

As a result of the aforementioned factors, there is less monitoring of ships on the open sea in ECAs. This is despite the fact that numerous studies (Pirjola et al., 2014; Kattner et al., 2015; Zhang et al., 2019) have shown that the FSC of ships were significantly reduced by the implementation of the ECA policy. However, most of these studies did not involve the monitoring of ships on the open water, which could lead to non-representative ~~inaccurate~~ assessments for the implementation of policies. At the same time, the lack of open sea monitoring results in a blind area for maritime enforcement~~supervision~~ and is not conducive to the implementation of ship ECA policy by maritime authority~~departments~~. ~~Therefore, the present study used an unmanned aerial vehicle (UAV) to monitor the emissions of sailing ships on the open sea in the Yangtze River estuary DECA. In comparison to the small aircraft method, this monitoring method can be used to monitor ship emissions at a lower cost to understand the emissions of sailing ships in open waters.~~The present study used an unmanned aerial vehicle (UAV) to monitor the FSC of sailing ships on the open sea in the Yangtze River estuary DECA. The method proposed in this study can be used to monitor ship emissions at a comparatively low cost to understand the FSCs of sailing ships in open waters. Although the cost of using patrol boats is not negligible, it is convenient and lower cost for maritime authorities compared with small aircraft.

2 Experimental methods

The research undertaken in the present study forms part of the project "Monitoring and inspecting ship exhaust emissions in the Shanghai free-trade zone" (MISEE). In this project, an unmanned aircraft system (UAS) was designed and developed, and mainly included a pod for measuring the exhaust gas from ships and a UAV to carry the pod. In previous research (Zhou et al., 2019), the plumes of 23 berthing ships were measured using the first-generation pod. The deviation of the estimated FSC obtained by the UAS was < 0.03% (m/m) for an FSC of between 0.035% (m/m) and 0.24% (m/m).

In the present monitoring for sailing ships, we developed the second-generation pod by optimizing the structure and layout of the first-generation pod to achieve a lighter weight and smaller volume. A short overview of the instrumentation is provided in Section 2.1. We measured the plumes of 11 berthing ships to verify the accuracy of the second-generation pod, and the plumes of 27 sailing ships to estimate the FSC.

2.1 Instrumentation

The UAS that was used for monitoring the FSC of sailing ships is shown in Fig 1. The UAV was a MATRICE 600 PRO (SZ DJI Technology Co., Ltd., Shenzhen, China). This type of UAV cannot be used on rainy days or when the wind speed is higher than 8 m/s. The white box installed underneath the UAV in Fig. 1 is the aforementioned second-generation pod for measuring the exhaust gas. When the UAV approaches a ship's plume, the gas pump in the pod ~~extracts~~ draws ~~gas~~ air using the gas probe. The water vapor, particles, and soot in the gas are subsequently removed by a hose filter valve.~~The water vapor in the gas is subsequently removed by the filter.~~ The sensors detect the gas and measurement information is sent out by communication modules. The pod has dimensions of 20 ~~mm~~ cm × 12 ~~mm~~ cm × 9 ~~mm~~ cm and weighs 900 g.

~~The sensors used were able to measure both SO_2 and CO_2, and were purchased from Shenzhen Singoan Electronic Technology Co., Ltd., China. The SO_2 sensor is based on the electrochemical method, and has a measuring range of 0–10 ppm, an accuracy of ± 5% (full range), and a response time (t_{90}) of < 1 s. The CO_2 sensor is based on the non-dispersive infrared analyzer method, and has a measuring range of 0–5000 ppm, an accuracy of ± 3% (full range), and a t_{90} of < 1 s. The t_{90} represents the time taken to reach 90% of the stable response following a step change in the sample concentration. These sensor characteristics were provided by the instrument manufacturer and were ensured to be within the tolerances by our own examination and calibration.~~The sensors used were able to measure both SO_2 and CO_2, and were purchased from Shenzhen Singoan Electronic Technology Co., Ltd., China. The SO_2 sensor is based on the electrochemical method, and has a measuring range of 0–10 ppm, an accuracy of ± 3% (0.3 ppm), and a response time (T_{90}) of ≤ 30 s. The CO_2 sensor is based on the non-dispersive infrared analyzer method, and has a measuring range of 0–10000 ppm, an accuracy of ± 3% (300 ppm), and a T_{90} of ≤ 30 s. The T_{90} represents the time taken to reach 90% of the stable response following a full range change in the sample concentration. These sensor characteristics were provided by the instrument manufacturer and were ensured to be within the tolerances by calibration. The zero and full scales are usually calibrated by a standard mixed gas when the equipment is used on a daily basis. The major parameters of the UAS are listed in Table 1.

2.2 Monitoring region

As illustrated in Fig. 2, the monitoring region was the channel of the Yangtze River estuary, near the Waigaoqiao port area to the north of Shanghai. The Yangtse River is the first (third) longest river in China (the world). Shanghai is one of the most prosperous cities in the world, and at the end of 2017 that city had a permanent resident population of approximately 24 million people (Shanghai Municipal Bureau of Statistics, 2017). The Waigaoqiao port area is only 20 km away from the city center, and the air pollution caused by ship emissions directly affects the urban air environment and the health of residents (Wang et al., 2019; Feng et al., 2019). The experimental area of the MISEE project is mainly within the Waigaoqiao port and the Yangtze River estuary.

2.3 Measurement method

During the experiment, the operator took a patrol boat to the channel and then selected a target ship at random. After identifying the target ship for monitoring, the patrol boat would accelerate to a distance to the left or right ahead of the vessel. The patrol boat would then stop and the UAV was operated to takeoff from its deck, and would then fly towards the plume of the target ship and measure the concentrations of SO_2 and CO_2 in the plume (Fig. 3). The distance between the patrol boat and the target ship was approximately a few hundred meters.

During the measurements, the operator adjusted the position of the UAV to ensure that it was in the ship's plume. Real-time measurements of SO_2 and CO_2 were made such that the pod could effectively detect the plume. Generally, it was necessary for the UAV to follow the ship's funnel mouth for approximately 5 minutes, as illustrated in Fig. 4. The target ship continued to move during the measurements; hence, it was followed by the patrol boat in order to avoid the UAV moving too far away from the operator. When the operator was sure that valid data had been collected, the patrol boat stopped and the UAV returned and landed back on the deck of the patrol boat.

2.4 Calculation

The FSC in this study was obtained directly by sampling the gas concentrations in the ship plumes using the UAS. The enhancements of SO_2 and CO_2 in measurements that were affected by exhaust gases were calculated, and the ratio of these SO_2 and CO_2 peaks was used to calculate the FSC (Eqs. 1 and 2). This method has been widely used to calculate the FSC in related studies (Alföldy et al., 2013; Pirjola et al., 2014; Balzani Lööv et al., 2014; Beecken et al., 2014; Beecken et al., 2015; Kattner et al., 2015; Zhou et al., 2019). In the calculation, the molecular weights of carbon and sulfur are 12 g mol^{-1} and 32 g mol^{-1}, respectively, and the carbon mass percent in the fuel is 87 ± 1.5% (Cooper et al., 2003). By assuming that 100% of the carbon content of the fuel is emitted as CO_2, and sulfur is emitted as SO_2 and other forms, the FSC mass percent can be determined using Eq. (1):

$$FSC[\%] = \frac{S[kg]}{fuel[kg]} = \frac{SO_2[ppm]\cdot A(S)}{CO_2[ppm]\cdot A(C)} \cdot 87[\%] + R = 0.232\frac{\int(SO_{2,peak} - SO_{2,bkg})dt[ppb]}{\int(CO_{2,peak} - CO_{2,bkg})dt[ppm]}[\%] + R = \frac{1}{20}EF[g_{SO_2}/kg_{fuel}] + R, \quad (1)$$

where R represents the sulfur content that is emitted in forms other than SO_2 because preliminary studies have shown that 1–19% of the sulfur in fuel is emitted in other forms, possibly SO_3 or SO_4 (Schlager et al., 2006; Alföldy et al., 2013; Balzani Lööv et al., 2014). EF is the emission factor and bkg is the abbreviation of background. In Eq. (1), if the sensors measuring SO_2 and CO_2 have approximately the same t_{90} response time and can be set to be synchronized, the peak concentrations of SO_2 and CO_2 can be used to calculate the FSC, ; otherwise, integrals need to be used. In our research, the sampling rate of the SO_2 and CO_2 sensors was 1 s, and integrals were used because the two sensors could not be completely synchronized. integrals were used because the two sensors could not be completely synchronized. In addition, the t_{90} of the sensors were not the same, although they were both < 1 s. A typical calculation is as follows.

The continuous measurement data for two typical plumes (2019-4-15B and 2019-3-29A) are exhibited in Fig. 5. The data for plume 2019-4-15B (Fig. 5a) were considered to be of a “good” quality, whereas those for plume 2019-3-29A (Fig. 5c) were considered to be of a “poor” quality. Data were determined to be of a good-quality when obvious, easily distinguished peak values were observed, whereas less obvious peaks that still corresponded to a result were considered as poor-quality data. The selection of peak values leads to uncertainty because when the area ratio is selected for the calculation, the starting and ending time points of the area are still associated with substantial uncertainty. Figure 5b and 5d depict the average concentrations of the SO_2 and CO_2 measurements (in Fig. 5a and 5c, respectively) for 10 s periods. The peak value of each average concentration was selected for the calculation. This process is equivalent to selecting the area ratio of SO_2 to CO_2 within 10 s for the calculation, as shown in Eq. (2).

$$FSC[\%] = 0.232\frac{\int(SO_{2,peak} - SO_{2,bkg})dt[ppb]/10}{\int(CO_{2,peak} - CO_{2,bkg})dt[ppm]/10}[\%] + R \approx 0.232\frac{AVG(SO_{2,peak}) - AVG(SO_{2,bkg})}{AVG(CO_{2,peak}) - AVG(CO_{2,bkg})}[\%], \quad (2)$$

where $AVG(\cdot)$ is the calculated function for the average measurement value within 10 s; hence, the data in this study are the average values of measurements in 10 s. When the UAV took off from the patrol boat and flew high into the air, the SO_2 and CO_2 concentrations were relatively low. The background values were obtained at this stage as the minimum SO_2 and CO_2 concentrations. As the UAV flew into the plume, the measured concentrations of SO_2 and CO_2 increased. The obvious, stable maximum values in the observations of the average measurement values should be selected as the peak values. It can be seen that using the average values of measurements within 10 s makes it easier to select the peak values, especially with respect to poor-quality data. However, as there can still be several options for peak values, the data treatment methods reported by Zhou et al. (2019) were incorporated in this study to select the most appropriate peak values. In Fig. 5b, the time point of selected peak values is at 10:19:11. The measurement values from 10:19:57 to 10:20:15 were not used because the CO_2 concentration covered the full range. In Fig. 5d, the time point of the selected peak values is at 10:38:27. The measurement values from 10:39:57 to 10:41:41 were not used because we ruled out data exhibiting either dramatic changes or errors in continuous observations. The details for selecting the peak values are listed in Table 2.

The sampling rate of the SO_2 and CO_2 sensors was 1 s. The continuous measurements obtained during one UAV flight are exhibited in Fig. 5a. The graph shows that the time point of the SO_2 peak is not synchronous with CO_2, which was mainly caused by different sensor response times. The selection of peak values is highly uncertain. When the area ratio is selected for the calculation, the selection of the starting and ending time points of the area is still associated with great uncertainty.

Figure 5b depicts the averages of the SO_2 and CO_2 measurements (in Fig. 5a) for 10 s periods. The peak value of the average was selected for the calculation. This process is equivalent to selecting the area ratio of SO_2 to CO_2 within 10 s for the calculation, as shown in Eq. (2).

$$FSC[\%] = 0.232\frac{\int(SO_{2,peak}-SO_{2,bkg})dt[ppb]/10}{\int(CO_{2,peak}-CO_{2,bkg})dt[ppm]/10}[\%] + R \approx 0.232\frac{AVG(SO_{2,peak})-AVG(SO_{2,bkg})}{AVG(CO_{2,peak})-AVG(CO_{2,bkg})}[\%], \quad (2)$$

where AVG $(\cdot)$ is the calculated function for the average measurement value. Therefore, the data in this study are the average values of measurements in 10 s. When the UAV took off from the patrol boat and flew high into the air, the SO_2 and CO_2 concentrations were relatively low. The background concentrations were obtained at this stage as the minimum SO_2 and CO_2 concentrations. As shown in Fig. 5b, the time frame was between 10:11:41 and 10:12:06. As the UAV flew into the plume, the measured concentrations of SO_2 and CO_2 increased. The stable, obvious, and maxima values in the observations of the average measurement values were selected as the peak values (time point 10:14:11).

2.5 Uncertainties

In previous research (Zhou et al., 2019), the main uncertainties of UAV measurements were summarized as sensor uncertainty, measurement uncertainty, calculation uncertainty, and exhaust uncertainty. The instrument calibration method, UAV flight procedures, and data treatment methods were designed to reduce these uncertainties. However, some uncertainties remain, as discussed below.

The average gas concentration within 10 s was chosen for the FSC calculations; however, this does not mean that 9 s or 11 s could not have been selected. To demonstrate this, a comparison calculation was carried out using both 9 s and 11 s, which showed that these led to very little differences in the results. However, it is necessary to ensure that the gradient of the gas measurements is stable within the sampling time (the interval length of the integral). Moreover, the interval length cannot be too short (e.g., 2 s) or too long (e.g., 20 s). If the time is too short, it is difficult to determine whether the measurements are stable and undisturbed over time. Similarly, if the time is too long, it is also difficult to ensure that all of the measurements in the integral interval are stable and undisturbed. In addition, during the flight of the UAV in this study, the time available for measuring the plume was ~5 minutes. As both the ship and the UAV were moving at this time, it was virtually impossible to ensure that the UAV was flying consistently within the plume and obtaining stable measurements. Accordingly, 10 s is also a relatively appropriate value for the measurement process.

Nevertheless, there is also some uncertainty associated with choosing the peak values. After ruling out the peak values across the full range as well as those corresponding to dramatic changes, the global maximum values were selected as the peak values to calculate the FSC. The maximum values probably correspond to the measurements taken in the center of the ship's plume. At that location, the measurement values were relatively stable, and the probability of interference from other factors was lower. Furthermore, the higher the peak value is, the greater the proportion of exhaust gas is; hence, the impact from the incomplete mixing of the exhaust gas with clean air is relatively small.

In summary, the obvious and stable maximum values are selected as peak values to calculate the FSC. There are, of course, situations where multiple similar peaks can occur simultaneously. In this case, their calculated FSCs may be very similar, and the results obtained by the calculation of the highest peak should have high credibility, for instance, the measurements of plume 2019-4-15B.

~~In previous research, the main uncertainties of UAV measurements were summarized as sensor uncertainty, measurement uncertainty, calculation uncertainty, and exhaust uncertainty. The instrument calibration method, UAV flight procedures, and data treatment methods reported by Zhou et al. (2019) were incorporated in this study to reduce these uncertainties. In addition, a distinction between good- and poor-quality data was made and some plume data were rejected. Good-quality data for a plume was reflected by obvious, easily distinguished peak values, whereas less obvious peaks that still corresponded to a result were considered as poor-quality data. When results could not be obtained, the plumes were rejected.~~

~~In this research, a~~Additional uncertainties were encountered during our monitoring of sailing ships because the UAV was usually hundreds of meters away from the operator. The location of a plume depended primarily on the following three aspects. 1. The position of most plumes with black smoke could be identified through the operator's visual judgment. 2. The real-time image shot by the camera can be used to assist in finding the ship's funnel mouth. 3. In the measurement process, the real-time measured concentration sent to the receiving equipment gradually increased, thus indicating that the UAV was approaching the center of the plume. However, the operator occasionally faced difficulties in accurately determining the ship's plume, which led to failed measurements. We attempted to measure more than 40 ship plumes in open water; however, only 27 of them resulted in good- or poor-quality data, i.e., usable data.

The deviation of the estimated FSC value obtained by the first-generation pod was < 0.03% (m/m) for an FSC level ranging from 0.035% (m/m) to 0.24% (m/m) (Zhou et al., 2019). The second-generation pod was also verified on berthing ships by using this method at a similar FSC level and the accuracy was approximately the same (see Section 3.1). These verifications of the deviation were based on the FSC measurement of berthing ships, which did not exceed the Chinese DECA FSC limit of 0.5% (m/m). However, some of the sailing ships did exceed this limit. It should be noted that the deviations for different FSC levels were not the same. Based on previous studies, the deviation of the FSC obtained from high-sulfur plume should be greater, for example, Van Roy and Scheldeman (2016a, b) estimated relative uncertainties of 20% ~~(m/m)~~ at a level of 1% (m/m) FSC and 50–100% at 0.1% (m/m) FSC. Therefore, the deviation of sailing ships may > 0.03% (m/m) when the FSC exceeds 0.5% (m/m). Nonetheless, our UAS was still able to accurately detect an FSC that obviously exceeded 0.5% (m/m).

3. Results

3.1 Berthing ships

Before monitoring the sailing ships, we first monitored 11 berthing ships between March and April 2019 in the Waigaoqiao port to verify the accuracy of the second-generation pod. Whilst one person operated the UAV to monitor one of the plumes, two maritime law enforcement officers boarded the corresponding ship to collect a fuel sample. Both processes took approximately 10–20 min. The fuel samples, which are considered to represent the true FSC values, were then sent for chemical analysis in a laboratory. The estimated (UAV) and true FSC values are listed in Table ~~1~~ 3 along with the identification number of each plume and the time and serial number. Table ~~1~~ 3 shows that the deviation did not generally exceed 0.03% (m/m) for an FSC level of between 0.03% (m/m) and 0.22% (m/m) (except for plume 2019-4-3B). Additionally, when the FSC of a target ship was low, for example, when light diesel fuel was used, the measured SO_2 concentrations were mostly zero. When this occurred, the FSC was generally < 0.02% (m/m), for example, as for plumes 2019-4-3A and 2019-4-12A.

3.2. Sailing ships and comparison with berthing ships

Between March and December 2019, effective monitoring of 27 sailing ships was undertaken using the UAV that took off from the patrol boat (Table ~~2~~4). The FSC of 23 berthing ships measured by the first-generation monitoring equipment and the FSC of 11 berthing ships (Table ~~1~~3) measured by the second-generation monitoring equipment in this study were taken as the FSC monitoring results for berthing ships. We compared the distribution of the FSCs of these 34 berthing ships with those of the 27 sailing ships. Figure 6 shows that the FSCs of the sailing ships were considerably higher than those of the berthing ships; the FSC of all 27 sailing ships exceeded 0.1% (m/m) and the FSC of 14 of these exceeded the Chinese DECA FCS limit of 0.5% (m/m), which included 5 exceedances of 2% (m/m). The uncertainty in the assessment is not small but the results so far, do not lead to optimism with respect to the FSC used by ships sailing in the area~~Despite the uncertainties described in Section 2.5, the current situation regarding the FSC of sailing ships has not been optimistic to date~~. The reason for this is that although berthing ships are sometimes boarded by maritime law enforcement officers for examination, an effective approach for monitoring the FSC of sailing ships in open water that leads to prosecution by China's maritime authorities has not existed prior to the present study.

According to the monitoring results, law enforcement officers of the Pudong maritime safety administration intercepted four sailing ships for which the UAV FSC results were of a good-quality and all exceeded 2% (m/m). The officers boarded these ships for inspection on July 15, August 14, August 20, and September 27, 2019, and took fuel samples, which were sent for chemical analysis in a laboratory. The FSC of all four fuels was also found to exceed 0.5% (m/m): 0.534% (m/m), 0.744% (m/m), 0.813% (m/m), and 1.991% (m/m) (in chronological order). The reason that all of these laboratory results did not exceed 2% related to the fact that ships cannot stop immediately in the channel for inspection and have to sail to the anchorage point; when the officers boarded the ships to take samples they found the crew taking various measures to drain the high-sulfur fuel in the main engine fuel oil pipeline. This means that the chemical analysis results of the sampled fuels were obviously lower than those of the UAV monitoring. Nevertheless, the four inspections successfully confirmed that the FSC of the fuels exceeded the standard for sailing ships. The inspection on July 15, 2019, was the first time that a sailing ship's FSC failed to meet Chinese regulations, and this aroused wide concern in the shipping community.

4. Conclusions

In this research, we used a UAV that took off from a patrol boat to monitor emissions from sailing ships in open water. Of the 27 sailing ships that were successfully monitored, 14 were found to have an FSC that exceeded 0.5% (m/m) and 5 exceeded 2% (m/m). Based on the monitoring results, law enforcement officers of the Pudong maritime safety administration caught the first case of excessive FSC for a sailing ship and confirmed three other cases. Additionally, the UAV monitoring results demonstrated that the FSC values of sailing ships in the surrounding waters of Waigaoqiao port were higher than those determined for berthing ships in the port. Although the sample size was relatively small, observation of Fig. 6 suggests that the data are still convincing.

Although a global cap on the FSC in marine fuel was set at 3.5% (m/m) in 2012 following the IMO regulation, this was reduced to 0.5% (m/m) in 2020 and has already been implemented in China. According to our monitoring results, the current situation for meeting the 0.5% limit is not optimistic. Successful compliance with this regulation by ship owners involves many challenges. We conclude that there is a need for further monitoring data on sailing ships in open water to ascertain the degree of exceedance and work toward compliance.

In addition, there are still some improvements to be made to the UAS. 4G transmission is the communication method for detecting information transmission; hence, in locations without a 4G signal (e.g., offshore), the receiving equipment cannot obtain real-time measurement results. Potential solutions include setting-up small base-stations on patrol boats or using satellite transmission. Although carrying an infrared camera on the UAV would make it easier to find the plume, this would require to replace the camera in Fig. 1 with an infrared camera and establish new data communication.

Data availability

Please address requests for data sets and materials to Fan Zhou (fanzhou_cv@163.com).

Author contribution

FZ designed the study and authored the article. FZ and LH analyzed the experimental data. RZ, WC and XN contributed to the experiments. SP contributed to setting instruments. LH, MZ and BA provided constructive comments on this research.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank Megan Anne for English language editing.

Financial support.

This research has been supported by the National Natural Science Foundation of China (grant No. 41701523), the Special Development Fund for China (Shanghai) Pilot Free Trade Zone, and Open Project Program of Key Laboratory of Intelligent Infrared Perception, Chinese Academy of Sciences.

References

Alföldy, B., Lööv, J. B., Lagler, F., Mellqvist, J., Berg, N., Beecken, J., Weststrate, H., Duyzer, J., Bencs, L., Horemans, B., Cavalli, F., Putaud, J.-P., Janssens-Maenhout, G., Csordás, A. P., Van Grieken, R., Borowiak, A., and Hjorth, J.: Measurements of air pollution emission factors for marine transportation in SECA, Atmos. Meas. Tech., 6, 1777–1791, https://doi.org/10.5194/amt-6-1777-2013, 2013.

Berg, N., Mellqvist, J., Jalkanen, J.-P., and Balzani, J.: Ship emissions of SO_2 and NO_2: DOAS measurements from airborne platforms, Atmos. Meas. Tech., 5, 1085–1098, https://doi.org/10.5194/amt-5-1085-2012, 2012.

Balzani Lööv, J. M., Alfoldy, B., Gast, L. F. L., Hjorth, J., Lagler, F., Mellqvist, J., Beecken, J., Berg, N., Duyzer, J., Westrate, H., Swart, D. P. J., Berkhout, A. J. C., Jalkanen, J.-P., Prata, A. J., van der Hoff, G. R., and Borowiak, A.: Field test of available methods to measure remotely SOx and NOx emissions from ships, Atmos. Meas. Tech., 7, 2597–2613, https://doi.org/10.5194/amt-7-2597-2014, 2014.

Beecken, J., Mellqvist, J., Salo, K., Ekholm, J., and Jalknen, J.-P.: Airborne emission measurements of SO2, NOx and particles from individual ships using a sniffer technique, Atmos. Meas. Tech., 7, 1957–1968, http://dx.doi.org/10.5194/amt-7-1957-2014, 2014.

Beecken, J., Mellqvist, J., Salo, K., Ekholm, J., Jalkanen, J.-P., Johansson, L., Litvinenko, V., Volodin, K., and Frank-Kamenetsky, D. A.: Emission factors of SO_2, NO_x and particles from ships in Neva Bay from ground-based and helicopter-borne measurements and AIS-based modeling, Atmos. Chem. Phys., 15, 5229–5241, https://doi.org/10.5194/acp-15-5229-2015, 2015.

Cheng, Y., Wang, S., Zhu, J., Guo, Y., Zhang, R., Liu, Y., Zhang, Y., Yu, Q., Ma, W., and Zhou, B.: Surveillance of SO2 and NO2 from ship emissions by MAX-DOAS measurements and implication to compliance of fuel sulfur content, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-l2019-369, in review, 2019.

Corbett, J.: Shipping emissions in East Asia, Nature Climate Change 6, 983–984, doi:10.1038/nclimate3091, 2016.

Eyring, V., Isaksen, I. S., Berntsen, T., Collins, W. J., Corbett, J. J., Endresen, O., Grainger, R. G., Moldanova, J., Schlager, H., and Stevenson, D. S.: Transport impacts on atmosphere and climate: Shipping, Atmos. Environ., 44, 4735–4771, https://doi.org/10.1016/j.atmosenv.2009.04.059, 2010.

Cooper, D. A.: Exhaust emissions from ships at berth, Atmos. Environ., 37, 3817–3830, https://doi.org/10.1016/s1352-2310(03)00446-1, 2003.

Feng, J., Zhang, Y., Li, S., Mao, J., Patton, A. P., Zhou, Y., Ma, W., Liu, C., Kan, H., Huang, C., An, J., Li, L., Shen, Y., Fu, Q., Wang, X., Liu, J., Wang, S., Ding, D., Cheng, J., Ge, W., Zhu, H., and Walker, K.: The influence of spatiality on shipping emissions, air quality and potential human exposure in the Yangtze River Delta/Shanghai, China, Atmos. Chem. Phys., 19, 6167–6183, https://doi.org/10.5194/acp-19-6167-2019, 2019.

Huebert, B. J.: Sulphur emissions from ships, Nature 400(6746), 713–714. doi:10.1038/23357, 1999.

International Maritime Organization (IMO), Emission Control Areas (ECAs) designated under MARPOL Annex VI. http://www.imo.org/en/OurWork/Environment/PollutionPrevention/AirPollution/Pages/Emission-Control-Areas-(ECAs)-designated-under-regulation-13-of-MARPOL-Annex-VI-(NOx-emission-control).aspx. Last access 10 December 2019. 2017.

Jalkanen, J.-P., Johansson, L., and Kukkonen, J.: A comprehensive inventory of ship traffic exhaust emissions in the European sea areas in 2011, Atmos. Chem. Phys., 16, 71–84, https://doi.org/10.5194/acp-16-71-2016, 2016.

Kattner, L., Mathieu-Üffing, B., Burrows, J. P., Richter, A., Schmolke, S., Seyler, A., and Wittrock, F.: Monitoring compliance with sulfur content regulations of shipping fuel by in situ measurements of ship emissions, Atmos. Chem. Phys., 15, 10087–10092, doi:10.5194/acp-15-10087-2015, 2015.

Liu, H., Fu, M., Jin, X., Shang, Y., Shindell, D., Faluvegi, G., Shindell, C., and He, K.: Health and climate impacts of ocean-going vessels in East Asia, Nature Climate Change 6, 1037–1041, doi:10.1038/nclimate3083, 2016.

MARPOL: International Convention for the Prevention of Pollution from Ships, 1973 as modified by the Protocol of 1978–Annex VI: Prevention of Air Pollution from Ships, International Maritime Organization (IMO), 1997.

Mellqvist, J., Beecken, J., Conde, V., and Ekholm J.: Surveillance of Sulfur Emissions from Ships in Danish Waters. Chalmers University of Technology, Göteborg, Sweden. December 30, 2017a.

Mellqvist, J., Conde, V., Beecken, J., and Ekholm, J.: Certification of an aircraft and airborne surveillance of fuel sulfur content in ships at the SECA border, CompMon (https://compmon.eu/, last access: 6 November 2018), 2017b.

Pirjola, L., Pajunoja, A., Walden, J., Jalkanen, J.-P., Rönkkö, T., Kousa, A., and Koskentalo, T.: Mobile measurements of ship emissions in two harbour areas in Finland, Atmospheric Measurement Techniques 6(4), 7149–7184. doi:10.5194/amtd-6-7149-2013, 2013.

Schlager, H., Baumann, R., Lichtenstern, M., Petzold, A., Arnold, F., Speidel, M., Gurk, C., and Fischer, H.: Aircraft-based Trace Gas Measurements in a Primary European Ship Corridor, Proceedings TAC-Conference, 26–29 June, Oxford, 83–88, 2006.

Shanghai Municipal Bureau of Statistics. The statistic communique of Shanghai on the 2017 national economy and social development, 2017.

Smith, T. W. P. et al. Third IMO Greenhouse Gas Study 2014. International Maritime Organization, 2014.

UNCTAD: World seaborne trade by types of cargo and by group of economies, annual, United Nations Conference on Trade and Development, available at: https://unctadstat.unctad.org/wds/TableViewer/tableView.aspx?ReportId=32363, last access: 5 March, 2017.

Van Roy, W. and Scheldeman, K.: Results MARPOL Annex VI Monitoring Report Belgian Sniffer Campaign 2016, CompMon (https://compmon.eu/, last access: 6 November 2018), 2016a.

Van Roy, W. and Scheldeman, K.: Best Practices Airborne MARPOL Annex VI Monitoring, CompMon (https://compmon.eu/,last access: 6 November 2018), 2016b.

Viana, M., Fann, N., Tobías, A., Querol, X., Rojas-Rueda, D., Plaza, A., Aynos, G., Conde, J., Fernaìndez, L., and Fernández, C.: Environmental and health benefits from designating the marmara sea and the Turkish straits as an emission control area (ECA), Environ. Sci. Technol., 49, 3304–3313, https://doi.org/10.1021/es5049946, 2015.

Wan, Z., Zhu, M., Chen, S., and Sperling, D.: Pollution: Three steps to a green shipping industry, Nature 530, 275–277, doi:10.1038/530275a, 2016.

Wang, X., Shen, Y., Lin, Y., Pan, J., Zhang, Y., Louie, P. K. K., Li, M., and Fu, Q.: Atmospheric pollution from ships and its impact on local air quality at a port site in Shanghai, Atmos. Chem. Phys., 19(9), 6315–6330. doi:10.5194/acp-19-6315-2019, 2019.

Yang, M., Bell, T. G., Hopkins, F. E., and Smyth, T. J.: Attribution of atmospheric sulfur dioxide over the English Channel to dimethyl sulfide and changing ship emissions, Atmos. Chem. Phys., 16, 4771–4783, https://doi.org/10.5194/acp-16-4771-2016, 2016.

Zhang, Y., Deng, F., Man, H., Fu, M., Lv, Z., Xiao, Q., Jin, X., Liu, S., He, K., and Liu, H.: Compliance and port air quality features with respect to ship fuel switching regulation: a field observation campaign, SEISO-Bohai, Atmos. Chem. Phys., 19, 4899–4916, https://doi.org/10.5194/acp-19-4899-2019, 2019.

Zhou, F., Pan, S., Chen, W., Ni, X., and An, B.: Monitoring of compliance with fuel sulfur content regulations through unmanned aerial vehicle (UAV) measurements of ship emissions, Atmos. Meas. Tech., 12, 6113–6124, https://doi.org/10.5194/amt-12-6113-2019, 2019.

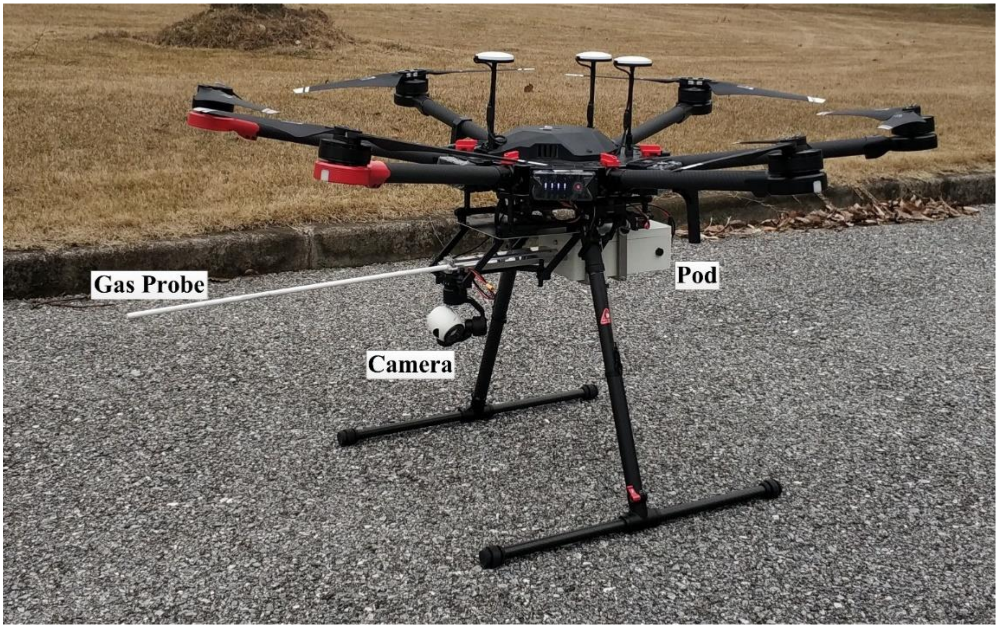


Figure 1. Image of the unmanned aircraft system. A gas probe, camera, and pod are installed under the unmanned aerial vehicle (UAV). The gas probe is used to collect the ship's exhaust gas, and the camera is used to assist in finding the ship's funnel mouth during flight. The pod is used to carry a gas pump, gas circuit, filter, small motor, sensors for SO_2 and CO_2, and communication modules.

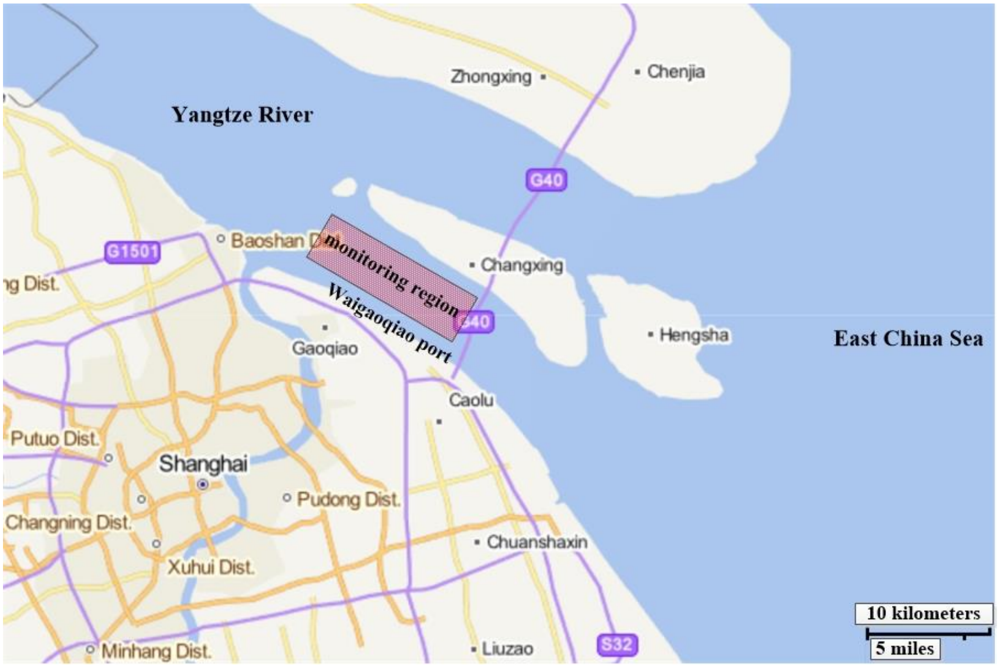
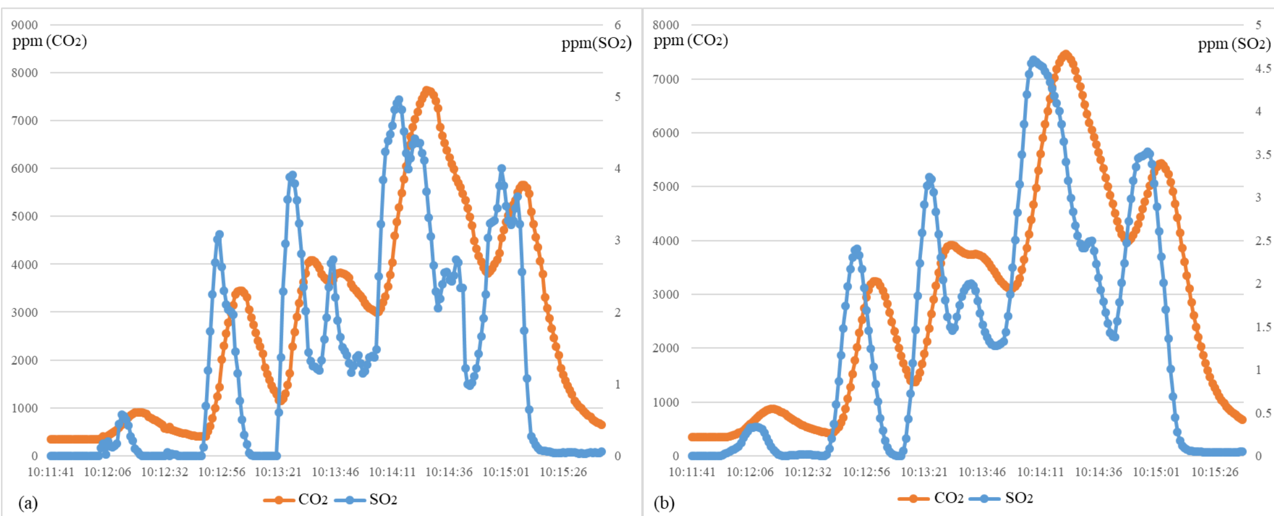


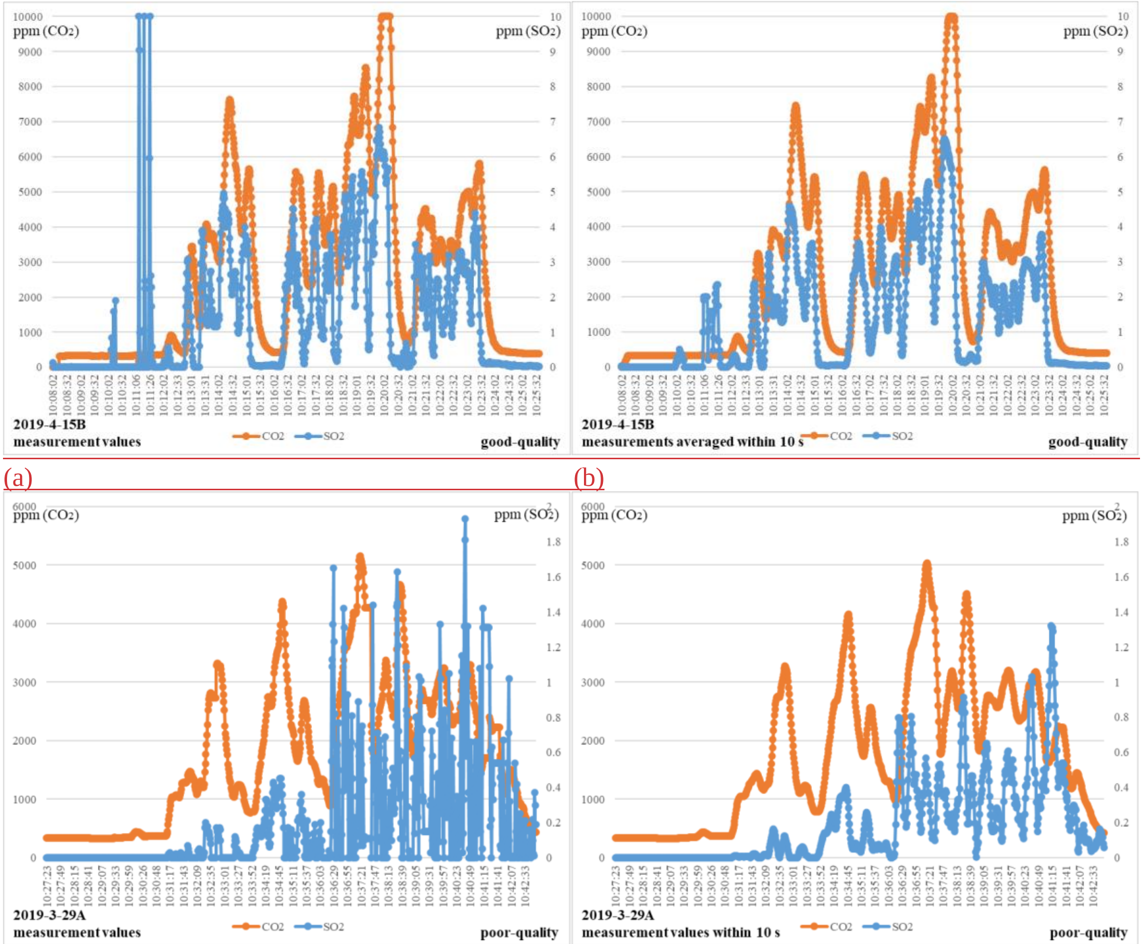
Figure 2. Monitoring regions in the channel of Yangtze River estuary, which belong to the DECAs of China. This area is to the north of Shanghai, on the southwest side of Changxing Island. The distance between the two sides is ~6–7 km. Ships leave the Yangtze River and sail into the East China Sea through this channel. Map data: @MapWorld (http://www.tianditu.gov.cn, last access: 5 March 2020).

Figure 3. Operator controlling the takeoff of the UAV from a patrol boat.

Figure 4. UAV (marked by a red circle) monitoring a ship's emissions in the open sea. An enlarged image of the UAV is shown at the top right. This picture was captured by another UAV.



~~**Figure 5. Typical calculations of data for the FSC, a) SO_2 and CO_2 measurement values, and b) SO_2 and CO_2 measurements averaged over 10 s periods.**~~



(a) (b) (c) (d)

Figure 5. Typical measurement data for SO_2 and CO_2 concentrations, and their corresponding average values within 10 s. (a) and (b) good-quality data from plume ID 2019-4-15B. (c) and (d) poor-quality data from plume ID 2019-3-29A. There are some errors in the measurements from 10:11:06 to 10:12:02 in (a), which may have been caused by sensor uncertainty. These data were ruled out and did not affect the calculation results.

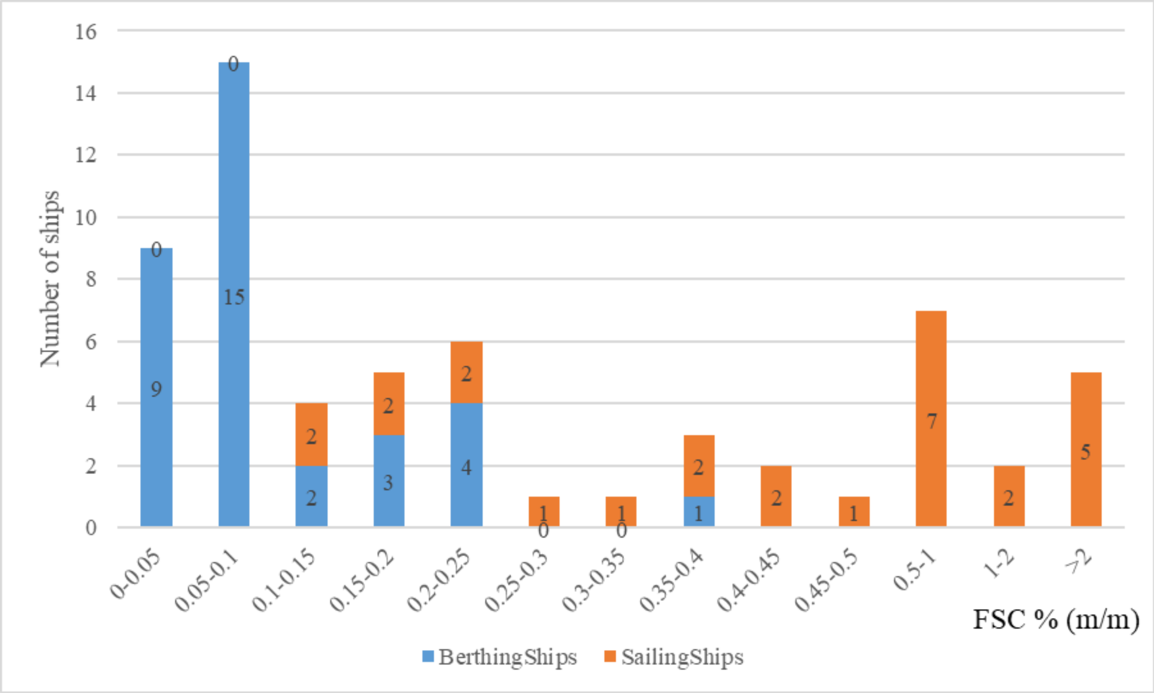


Figure 6. Comparison between the monitoring results of berthing ships and sailing ships.

Table 1: Parameters of the UAS

	Parameter	Value
UAV	Symmetrical motor wheelbase	1133 mm
	Size	1668 mm × 1518 mm × 727 mm
	Weight	9.5 kg
	Recommended maximum take-off weight	15.5 kg
	Hovering accuracy(P-GPS)	Vertical: ±0.5 m, Horizontal: ±1.5 m
	Maximum rotational angular velocity	Pitch axis: 300°/s, Heading axis: 150°/s
	Maximum pitch Angle	25°
	Maximum rising speed	5 m/s
	Maximum rate of descent	3 m/s
	Maximum sustained wind speed	8 m/s
	Maximum horizontal flight speed	65 km/h (no wind environment)
	Hover time	Non-loaded: 32 min; load 6 kg: 16 min
SO_2 sensor	Type	SGA-700A-SO2
	Principle	Electrochemistry
	Measuring range	0–10 ppm
	Diameter and height	33.5 mm; 31 mm
	Weight	30 g
	Accuracy	≤ ±3 % (0.3 ppm)
	Linear error	≤ ±2 % (0.2 ppm)
	Repeatability	≤ ±2 % (0.2 ppm)
	Power consumption	≤ 50 mA
	Response time (T_{90})	≤ 30 s
CO_2 sensor	Type	SGA-700A-CO2
	Principle	Non-Dispersive InfraRed
	Measuring range	0–10000 ppm
	Diameter and height	33.5 mm; 31 mm
	Weight	30 g
	Accuracy	≤ ±3 % (300 ppm)
	Linear error	≤ ±2 % (200 ppm)
	Repeatability	≤ ±2 % (200 ppm)
	Power Consumption	≤ 100 mA
	Response time (T_{90})	≤ 30 s

Table 2: All peak values and their corresponding FSC results. The background values of plume 2019-4-15B were 0 ppm and 310 ppm for SO_2 and CO_2, respectively. The background values of plume 2019-3-29A were 0 ppm and 329 ppm for SO_2 and CO_2, respectively. The remarks indicate the reason for choosing or not choosing the peak. It can be seen that the peak value of plume 2019-4-15B was more obvious and that the results obtained by multiple alternative peaks were similar. The peak of plume 2019-3-29A was less obvious and there were fewer alternative peaks. This was also the basis for distinguishing data as being of a "good"/"poor" quality. The FSC result of selected peak values are marked as "√".

Plume ID	Time point	Peak value of SO_2 and CO_2 (ppm)	Estimated value of FSC (% (m/m))	True value of FSC (% (m/m))	Remark
2019-4-15B	10:12:52	2.406, 2020	0.326	0.168	Reject; less obvious peak values
	10.13.23	3.235, 2372	0.364		
	10.14.07	4.594, 4665	0.245		Non-maximum peaks of alternative peak values
	10.14.57	3.529, 4872	0.179		
	10.16.39	3.549, 4444	0.199		
	10:17:27	3.989, 3911	0.257		
	10:18:01	3.159, 4607	0.171		
	10:18:47	4.757, 6895	0.168		
	10:19:11	5.287, 7634	0.167 (√)		Maximum peak of the alternative peak value
	10:19:46	6.515, 8100	0.194		Reject; measurements exceeded the range
2019-3-29A	10:34:41	0.399, 3880	0.026	0.035	Reject, less obvious peak values
	10:35:19	0.258, 2011	0.036		Non-maximum peaks of the alternative peak values
	10:37:15	0.567, 4994	0.028		Reject; less obvious peak values
	10:38:27	0.913, 4022	0.057 (√)		Maximum peak of the alternative peak value
	10:40:37	1.031, 2996	0.090		Reject; error in the measurement data
	10:41:13	1.321, 1700	0.224		

Table 3: Comparison and verification of the estimated (UAV) and true (sampled fuel) values of the FSC from 11 berthing ships.

ID	Estimated value of FSC (% (m/m))	True value of FSC (% (m/m))	Deviation (% (m/m))	Quality
2019-3-18A	0.217	0.222	-0.005	Good
2019-3-22A	0.069	0.099	-0.030	Good
2019-3-22B	0.046	0.042	0.004	Good
2019-3-29A	0.057	0.035	0.022	Poor
2019-4-1A	0.090	0.079	0.011	Good
2019-4-3A	<0.020	0.013	N	Poor
2019-4-3B	0.057	0.092	-0.035	Good
2019-4-12A	<0.020	0.004	N	Poor
2019-4-12B	0.092	0.080	0.012	Good
2019-4-15A	0.053	0.044	0.009	Good
2019-4-15B	0.167	0.168	-0.001	Good

Table 4: Estimated (UAV) values of the FSC from 27 sailing ships. "*" indicates that the ship was boarded by the maritime authority for inspection, and the value shown in parentheses is the result of the chemical examination of the fuel.

ID	Estimated value of FSC (% (m/m))	Quality	ID	Estimated value of FSC (% (m/m))	Quality
2019-7-12A	0.781	Good	2019-8-22A	0.186	Good
2019-7-15A	0.646	Good	2019-8-22B	0.385	Poor
2019-7-15B*	3.369 (0.534)	Good	2019-8-22C	0.415	Good
2019-7-25A	0.580	Good	2019-8-22D	0.112	Poor
2019-7-25B	0.675	Good	2019-8-22E	0.104	Good
2019-8-14A*	2.672 (0.744)	Good	2019-8-22F	0.239	Poor
2019-8-15A	0.382	Good	2019-9-17A	0.202	Good
2019-8-15B	0.694	Poor	2019-9-17B	0.628	Poor
2019-8-16A	0.175	Poor	2019-9-27A	0.419	Poor
2019-8-16B	0.267	Poor	2019-9-27B*	3.450 (1.991)	Good
2019-8-16C	1.127	Good	2019-10-9A	2.116	Poor
2019-8-16D	0.700	Poor	2019-10-17A	0.481	Good
2019-8-20A	1.508	Poor	2019-10-24A	0.326	Good
2019-8-20B*	4.091 (0.813)	Good			

~~**Table 1: Comparison and verification of the estimated (UAV) and true (sampled fuel) values of the FSC from 11 berthing ships.**~~

ID	Estimated value of FSC (ppm)	True value of FSC (ppm)	Deviation (ppm)	Quality
2019-3-18A	2171	2221	-50	Good
2019-3-22A	690	985	-295	Good
2019-3-22B	464	424	40	Good
2019-3-29A	574	345	229	Poor
2019-4-1A	895	787	108	Good
2019-4-3A	<200	128	N	Poor
2019-4-3A	566	917	-351	Good
2019-4-12A	<200	40	N	Poor
2019-4-12B	924	798	126	Good
2019-4-15A	527	440	87	Good
2019-4-15B	1675	1682	-7	Good

Table 2: The estimated (UAV) values of the FSC from 27 sailing ships. "*" indicates that the ship was boarded by the maritime authority for inspection.

ID	Estimated value of FSC (ppm)	Quality	ID	Estimated value of FSC (ppm)	Quality
2019-7-12A	7805	Good	2019-8-22A	1862	Good
2019-7-15A	6464	Good	2019-8-22B	3851	Poor
2019-7-15B*	33687	Good	2019-8-22C	4153	Good
2019-7-25A	5803	Good	2019-8-22D	1122	Poor
2019-7-25B	6752	Good	2019-8-22E	1043	Good
2019-8-14A*	26723	Good	2019-8-22F	2386	Poor
2019-8-15A	3817	Good	2019-9-17A	2021	Good
2019-8-15B	6941	Poor	2019-9-17B	6276	Poor
2019-8-16A	1746	Poor	2019-9-27A	4191	Poor
2019-8-16B	2672	Poor	2019-9-27B*	34495	Good
2019-8-16C	11273	Good	2019-10-9A	21158	Poor

~~2019-8-16D~~	~~7008~~	~~Poor~~	~~2019-10-17A~~	~~4813~~	~~Good~~
~~2019-8-20A~~	~~15076~~	~~Poor~~	~~2019-10-24A~~	~~3258~~	~~Good~~
~~2019-8-20B*~~	~~40913~~	~~Good~~			