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

The title is sloppy: "High-precision monitoring of compliance with fuel sulfur content through UAV measurements of ship emissions" Should be rather: "High-precision monitoring of compliance with fuel sulfur content regulations/limitation/standards through UAV measurements of ship emissions"

Considering the content of the manuscript, we guess the word "regulations" is more appropriate. At the same time, Referee #2 suggest leaving out the word "High-precision" because "in ECA areas, with a limit of 0.1%, an uncertainty of 0.03% is not very good". In addition, with the consideration that the precision will improve further, with the progress of technology and method. "High-precision" seems not appropriate, we changed the title as:

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

1. Introduction

The literature review is poor.

The suggestion is very important. We have rewritten this part carefully. The major changes are listed below.

Page1, line20: Year of citation is 2005. Add current data here since anthropocentric SO_2 emission might change significantly during 14 years.

This paragraph has been rewritten.

Estimations show that ships contribute 4-9% of global SO_2 emissions and 15% of NO_x (Eyring et al., 2010). According to the United Nations Conference on Trade And Development (UNCTAD, 2017), the volume of the world's seaborne trade grew by 66% between 2000 and 2015. As global commerce expands, ocean-going ships consume more fuels, generally low-quality residual fuels containing high concentrations of sulfur and heavy metals (Lack et al., 2011). From the viewpoint of spatial distribution, the highest emissions of SO_2 per unit area occur in the eastern and southern China seas, sea areas in south-eastern and southern Asia, Red Sea, Mediterranean Sea, North Atlantic

near the European coast, Gulf of Mexico and Caribbean Sea, and along the western coast of North America. (Johansson et al., 2017). Ship-emitted pollutants influence air quality, human health, and climate. They not only affect the air quality in coastal areas but even influence the inland areas hundreds of kilometers away from the emission sources (Liu et al., 2016).

Reference:
[1] 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.
[2] 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.
[3] Lack, D. A., Cappa, C. D., Langridge, J., Bahreini, R., Buffaloe, G., Brock, C., Cerully, K., Coffman, D., Hayden, K., Holloway, J., Lerner, B., Massoli, P., Li, S.-M., McLaren, R., Middle-brook, A. M., Moore, R., Nenes, A., Nuaaman, I., Onasch, T. B., Peischl, J., Perring, A., Quinn, P. K., Ryerson, T., Schwartz, J. P., Spackman, R., Wofsy, S. C., Worsnop, D., Xiang, B., and Williams, E.: Impact of Fuel Quality Regulation and Speed Reductions on Shipping Emissions: Implications for Climate and Air Quality, Environ. Sci. Technol., 45, 9052-9060, https://doi.org/10.1021/es2013424, 2011.
[4] Johansson, L., Jalkanen, J. P., and Kukkonen, J.: Global assessment of shipping emissions in 2015 on a high spatial and temporal resolution, Atmos. Environ., 167, 403-415, https://doi.org/10.1016/j.atmosenv.2017.08.042, 2017.
[5] 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, Nat. Clim. Change., 6, 1037-1041, 10.1038/nclimate3083, 2016.

Pg1, ln25: "some regulation went into effect" - needed to be rephased.

These sentences have been rewritten.

In 2005, some regulations went into effect after being received by appropriate laws of the signatory states (at the European level it was received with the directives 1999/32/EC, 1999, and 2005/33/EC, 2005), and introduces limits to marine fuel sulfur content and engine performance to reduce SO_x and NO_x emissions. Further amendments to Annex VI were adopted in 2008 and entered into force in 2010.

Reference:
[1] Directive 1999/32/EC: Official Journal of the European Union, L 121, p. 13, 26 April 1999.
[2] Directive 2005/33/EC: Official Journal of the European Union, L 191, p. 59, 22 July 2005.

Pg2, ln4: "To cope with..." ???

This sentence has been rewritten.

In order to reduce the air pollution caused by ship emissions, the Atmospheric Pollution Prevention and Control Law of the People's Republic of China was promulgated in 2015 (Standing Committee of the National People's Congress, 2015).

Reference:
[1] Standing Committee of the National People's Congress, Atmospheric Pollution Prevention and Control Law of the People's Republic of China, 2015.

The overview of techniques (Pg2, ln15-ln26) is incorrect. The optical methods (LIDAR, UV cam, DOAS) can measure only the SO_2 emission rate. For emission factor calculation simultaneous CO_2 emission rate measurement is needed preferable on the same parcel of the plume. It can be implemented by open path FTIR technique but it is quite challenging. Another solution could be to model the CO_2 emission rate based on the ship's technical properties and sailing characteristics. Then the SO_2 or S emission factor (EF) as well as the fuel sulphur content (FSC) can be calculated.

In the manuscript, we only discuss that the optical methods can be used to measure ship emissions (do not indicate that they could measure the CO_2). We think this part was not clear enough, so we have rewritten it combined with the suggestion.

Optical methods analyze variations in light properties after interactions with the exhaust plume, and the local wind field before determining the SO_2 emission rate is observed. The simultaneous measurement of CO_2 and SO_2 emissions on a routine basis is unrealistic at present. Thus, the amount of fuel burned at the time of measurement is unknown and has to be estimated via modeling for calculating the FSC. For instance, the model STEAM (ship traffic emission assessment model), developed by the Finnish Meteorological Institute (Jalkanen et al., 2009) was used in the research for estimating FSC by Balzani Lööv et al. (2014). In addition, using the ratio of SO_2 and NO_2 measured via DOAS in the ship plume can be used as an indicator of FSC (Johan, R et al. 2017, Cheng, Y et al, 2019).

Reference:
[1] 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., vander Hoff, G. R., and Borowiak, A.: Field test of available methods to measure remotely SO_x and NO_x emissions from ships, Atmos. Meas. Tech., 7, 2597–2613, doi:10.5194/amt-7-2597-2014, 2014.
[2] Cheng, Y., Wang, S., Zhu, J., Guo, Y., Zhang, R., Liu, Y., Zhang, Y., Yu, Q., Ma, W.,

and Zhou, B.: Surveillance of SO_2 and NO_2 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-2019-369, in review, 2019.
[3] Johan, R., Conde, V., Beecken, Jörg and Ekholm, J.: Certification of an aircraft and airborne surveillance of fuel sulfur content in ships at the SECA border, CompMon (https://compmon.eu/), 2017.
[4] Jalkanen, J.-P., Brink, A., Kalli, J., Pettersson, H., Kukkonen, J.,and Stipa, T.: A modelling system for the exhaust emissions of marine traffic and its application in the Baltic Sea area, Atmos. Chem. Phys., 9, 9209–9223, doi:10.5194/acp-9-9209-2009, 2009.

Pg2, ln25: What is the effect of NO_x sensor on FSC measurement?

The SO_2 analyzer (fluorescence) response has cross sensitivity to NO. The supplementary explanation is given in the manuscript.

The "sniffing" method is based on simultaneous measurement of elevated SO_2 and CO_2 concentrations in the exhaust plume from the target ship and comparing them with the background. The measurement of CO_2 allows for relating the measurement of SO_2 to the amount of fuel burned at a given time, thus enabling the calculation of FSC directly. The concentration of SO_2 in plumes was generally measured using UV fluorescence or electrochemical sensors, and CO_2 was measured using a non-dispersive infrared analyzer (NDIR) or cavity ring down spectrometer (CRDS). The advantage of the "sniffing" method is that it offers more accuracy estimation for FSC. However, the instrument must be placed in the plume exhausted by the target ship. In some studies (Van Roy and Scheldeman, 2016a, 2016b), the "sniffing" method offers a measurement accuracy between 0.1–0.2% (m/m) FSC, which can be further increased up to 0.05–0.1% (m/m) FSC if combined with an additional NO_x sensor. This is because the response of SO_2 analyzers (fluorescence) has cross sensitivity to NO. Deviations are not the same at different FSC levels, with an estimated relative uncertainty of 20% (m/m) for ships with 1% (m/m) FSC and a relative uncertainty of 50–100% at 0.1% (m/m) FSC. Balzani Lööv et al. (2014) obtained the following FSC measurements based on the "sniffer" principle: 0.86±0.23% (m/m) from land, 1.2±0.15% (m/m) from an on-board stack, and 1.13±0.18% (m/m) from a mobile platform. There was a 6% relative uncertainty for an FSC of 1% (m/m) but a 60% relative uncertainty for an FSC of 0.1% (m/m).

Reference:
[1] 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., vander 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, doi:10.5194/amt-7-2597-2014, 2014.

[2] Van Roy, W. and Scheldeman, K.: Results MARPOL Annex VI Monitoring Report Belgian Sniffer Campaign 2016, CompMon (https://compmon.eu/), 2016a.
[3] Van Roy, W. and Scheldeman, K.: Best Practices Airborne MARPOL Annex VI Monitoring, CompMon (https://compmon.eu/), 2016b.

In addition, several UAV applications have been done before that must be mentioned here.

After reviewing relevant literatures, we found that there are some UAVs used to measure greenhouse gases and volcanic eruptions. Only one paper has been found on the measurement of ship emissions. We also have carried on the supplementary discussion in the manuscript.

Ship emission measurements can be divided into land-based (Kattner et al., 2015, Yang et al., 2016), marine-based (Cappa et al., 2014), airborne-based (Beecken et al., 2014, Aliabadi et al., 2016), satellite-based (Ding et al., 2018) and Unmanned Aerial Vehicle (UAV)-based (Villa et al., 2019) according to different platforms.
…
UAV-based measurements have gradually increased in the research regarding the atmosphere (Mori et al., 2016, Malaver Rojas et al., 2015). However, to date, there are relatively few applications of these measurements in ship emissions.

Reference:
[1] Aliabadi, A. A., Thomas, J. L., Herber, A. B., Staebler, R. M., Leaitch, W. R., Schulz, H., Law, K. S., Marelle, L., Burkart, J., Willis, M. D., Bozem, H., Hoor, P. M., Köllner, F., Schneider, J., Levasseur, M., and Abbatt, J. P. D.: Ship emissions measurement in the Arctic by plume intercepts of the Canadian Coast Guard icebreaker Amundsen from the Polar 6 aircraft platform, Atmos. Chem. Phys., 16, 7899–7916, doi:10.5194/acp-16-7899-2016, 2016.
[2] Beecken, J., Mellqvist, J., Salo, K., Ekholm, J., and Jalkanen, J.P.: Airborne emission measurements of SO_2, NO_x and particles from individual ships using a sniffer technique, Atmos. Meas. Tech., 7, 1957–1968, doi:10.5194/amt-7-1957-2014, 2014.
[3] Cappa, C. D., Williams, E. J., Lack, D. A., Buffaloe, G. M., Coffman, D., Hayden, K. L., Herndon, S. C., Lerner, B. M., Li, S.M., Massoli, P., McLaren, R., Nuaaman, I., Onasch, T. B., and Quinn, P. K.: A case study into the measurement of ship emissions from plume intercepts of the NOAA ship Miller Freeman, Atmos. Chem. Phys., 14, 1337–1352, doi:10.5194/acp-14-1337-2014, 2014.
[4] Ding, J., van der A, R. J., Mijling, B., Jalkanen,J.-P.,Johansson,L.,and Levelt,P.F.: Maritime NOx emissions over Chinese seas derived from satellite observations. Geophysical Research Letters, 45, 2031–2037, doi:10.1002/2017GL076788, 2018.
[5] 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.
[6] Villa, T. F., Brown, R. A., Jayaratne, E. R., Gonzalez, L. F., Morawska, L., and Ristovski, Z. D.: Characterization of the particle emission from a ship operating at sea using an unmanned aerial vehicle, Atmos. Meas. Tech., 12, 691-702, https://doi.org/10.5194/amt-12-691-2019, 2019.
[7] 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.

As a summary; a new and more precise Introduction is needed.

We have tried our best to rewrite the introduction.

2. Measurement:
Avoid mini-sniffer term. Sniffing technique supposes an airflow through the analyzer. In this context the CO_2 analyzer can sniff but the SO_2 sensor definitely cant. Better to use sensing or electrochemical sensing terms.

The term of "mini-sniffer" in the manuscripts has been revised.

This section is also incomplete. What sensors were used? Manufacturers, types, characteristics? The description of the sniffing technique (Pg5) is poor and incomplete. For example, the proper handling of the water vapor interference with CO_2 measurements is crucial. How did the authors handle that?

The sensors are commercially available. This information has been supplemented.

In the measurement process, the ship exhaust is pumped into the pod by the gas pump. After the filter removes the water vapor, the sensors react and the communication module sends the measurement results to the receiving end. The sensors included instrumentation for both SO_2 and CO_2 measurements. These sensors were purchased from HH Feuerungstechnik GmbH, Germany.

In addition, we supplement the description in figure 1.

Figure 1. Image of the modified UAV platform. The black box installed under the UAV is a pod which was designed and customized by us. It carries a gas pump (to collect the ship's exhaust), gas circuit, a filter (to remove water vapor), sensors for SO_2 and CO_2, a small motor (to provide energy for pumping), a camera, and communication modules.

In the experiment, we used the MATRICE 600 UAV (SZ DJI Technology Co., Ltd.), and modified it. We designed and customized a special pod, which was installed underneath the UAV, to carry sensors, communication circuit boards, gas circuit systems, and other modules, as shown in Fig.1.

Ln13: "calibrated 3 month or 180 working hours apart" ??? - why is this big difference?

This sentence has been rewritten.

Sensor calibration is required when the equipment is used daily. The time interval for sensor calibration is three months or when the accumulated working time of the sensor exceeds 180 h. If either of these conditions is met, calibration will be carried out.

I miss evidences of laboratory experiments where the sensors were calibrated and tested, effects of environmental factors (temperature, humidity) were investigated as well as interference of other components (water vapor) was checked.

In the laboratory experiment, we mainly test the stability and safety of the whole UAV system as well as communication modules. At the same time, it also allows the UAV operator to practice how to operate the UAV for sampling close to the smoke stack. There is a risk of getting too close, the operator needs to practice.
We used commercial sensors. We chose the type of sensors according to the need of experiment. We have added some details (sources of sensors, technical parameters, calibration, filters, etc) about the whole pod as mentioned above. We hope this will make the readers more familiar with our work.

In the experiment, we used the MATRICE 600 UAV (SZ DJI Technology Co., Ltd.), and modified it. We designed and customized a special pod, which was installed underneath the UAV, to carry sensors, communication circuit boards, gas circuit systems, and other modules, as shown in Fig.1. After the successful assembly of the UAV platform, we first carried out preliminary experiments in the automatic engine room laboratory of Shanghai Maritime University. Through the preliminary test, we verified the stability and security of the whole UAV system. At the same time, it also allowed the UAV operator to practice how to operate the UAV for sampling close to the smoke stack.

4. Results

The authors claimed that they measured 20 ships. Why only 6 plumes were presented here? It is not clear how the authors accepted or discarded results. What were the main steps of the consideration?

At the time of initial submission, we listed 12 plumes. However, similar plumes do not seem to need to be listed multiple times. Therefore, we chose the typical six plumes and discussed in detail the process of selecting the peak and background values. The data of these six plumes are only suitable for our discussion method, and have no other particularities. In addition, we supplemented the results of all 23 monitoring experiments in the manuscript.

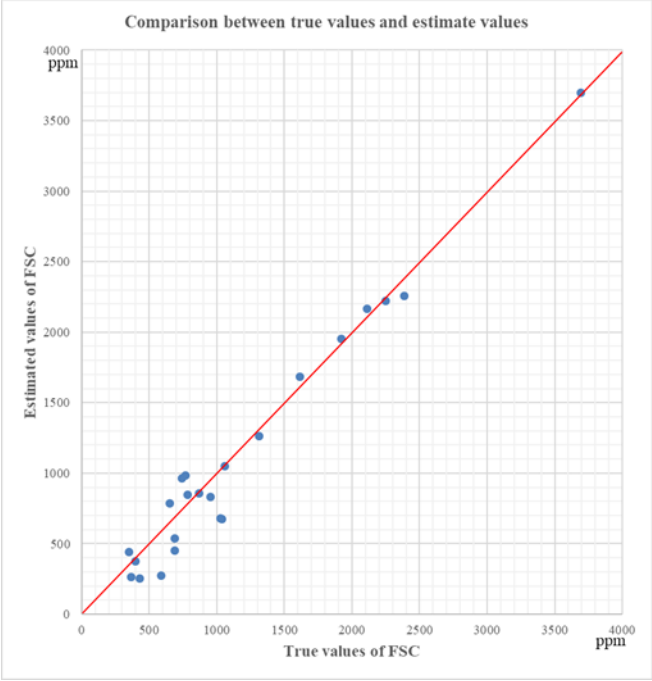

Figure 6. Comparison between the true values of FSC (x-axis) against the estimated values of FSC (y-axis) of 23 times measurement.

As shown in Fig 6, the FSC in our experiments was mainly at a level of 0.035% (m/m) to 0.24% (m/m) (only one measurement of 0.37% (m/m), not enough for reference). The deviation of the estimated FSC value calculated using the proposed method was within 0.03% (m/m), although there was some uncertainty. Considering the uncertainties listed in section 3.3, the proposed method provides accurate results. Overall, the estimated FSC is smaller than the true value in many cases. This is because 1–19% of the sulfur in the fuel is emitted in other forms, possibly SO_3 or SO_4.

On the other hands, the main strength of the work that the authors compared their plume measurements with the chemical analysis of the fuel. The authors did not mention the biggest challenge of the technique, namely how can we synchronize the time variations of two different measurements (SO_2 and CO_2) in order to calculate their ratio. In case of broader plume parcels (and thus longer measurement time) during conventional sniffing technique the uncertainty of the integrals are negligible, so more or less exact ratio can be calculated. On the other hands, the sensing technique provides narrow peaks of both components where the uncertainty of the integral is significant. How can the authors describe the differences in the time variation of the two components (see the figures on Pg10)?

Yes, this is indeed a key technical problem we encountered. We have added the following explanation:

The response time of both sensors is less than 1s. Even if the sampling rates of the two sensors are set to be consistent, the two sensors cannot be completely synchronized. This makes it difficult to calculate the ratio of SO_2 and CO_2. Our approach is that the sensor sends the average measurement value of the last 10 s to the receiver at an interval of 10 s. Therefore, the interval of integration in Eq. (1) is 10 s. We determined that taking the mean of measurements directly or at shorter intervals leads to too many narrow peaks in one measurement process. This makes it difficult to select the peak value, and the calculation results are unstable. At the same time, the interval should not be set too long, which will make the crest very inconspicuous or too flat. Therefore, we selected 10 s as the empirical parameter value after several experiments.
Also, in the description of result:
After the measurement of plume 5, the communication module was fault when we wanted to adjust sampling rate. We consequently replaced the communication protocol "HTTP protocol" with the "TCP/IP protocol". The main changes involved adjusting the data sampling rate from 10 to 2 s to make it easier to find the peak value (the sensor sends the average measurement value of the last 10 s to the receiver at an interval of 2 s), and the sensors were consequently recalibrated by standard mixture gas.

In Plume 5 and Plume 6 the CO_2 sensor was saturated at different concentrations (1900 vs 5000 ppm). What was the reason?
The CO_2 sensor has a range of 5000ppm. There appears saturation in plume 6 but not in plume 5. It looks like saturated, but it's not saturated. We checked the data as fellow:

time	SO_2	CO_2
11:18:14	774	1122.8
11:18:25	993	1616
11:18:35	938	1895.2
11:18:46	938	1895.2
11:18:57	1049	1896.8
11:19:08	1406	1893.6
11:19:18	1348	1894.8
11:19:29	1078	1896
11:19:40	765	1886
11:19:51	635	1516
11:20:01	603	1066.8
11:20:12	576	811.6

This kind of situation is rare. It is difficult to draw conclusions at this time. We guess that this may be due to sensor uncertainty. In any case, the data in this period were not used as peak values of the plumes as present in the manuscripts.

In the end, we thank the Referee #1 for his/her positive and constructive comments.