THE SCIPPER PROJECT

Shipping Contributions to Inland Pollution Push for the Enforcement of Regulations

EUROPEAN COMMISSION
Horizon 2020 No. 814893

Deliverable No.	SCIPPER Project D 2.3	
Deliverable Title	Quality assurance of remote monitoring systems and harmonised reporting	
Dissemination Level	Public	
Written by	Johan Mellqvist, Jörg Beecken, Vladimir Conde, Daniëlle van Dinther, Jan Duyzer, Simone Griesel, Matti Irjala, Bettina Knudsen, Jon Knudsen, Marcel Moerman, Leonidas Ntziachristos and Andreas Weigelt	
Approved by	Erik Fridell and Ruud Verbeek	06/12/2022
Status	First Submission to EC portal	21/12/2022

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nr.814893

This report reflects only the author's view and CINEA is not responsible for any use that may be made of the information it contains

Executive Summary

ES 1 Introduction

In 2015, new rules from the International Maritime Organization (IMO), European Commission and US legislation require ships to run with maximum Fuel Sulphur Content (FSC) of 0.10 % S m/m, primarily in Northern European and US waters. National sulphur fuel restrictions also occur in part of Chinese and South Korean waters since January 1, 2022. In addition, since 2020, ships worldwide are only allowed to operate with a maximum FSC of 0.50 % S m/m outside the Sulphur Emission Control Areas (SECAs). The Mediterranean Sea will also become a SECA from 2025 on. From 2016 and 2021, respectively, the US coast and Northern Europe, also comprise NO_x emission control areas, requiring about 75 % NO_x reduction of ships keel laid from the corresponding dates and on and compared to the Tier II[1] requirement. Finally, there is presently a discussion within IMO on how to control particle emission in form of black carbon.

Port state control authorities check that vessels operate with compliant fuel in each area, by on-board fuel sampling for ships at berth. A supporting method to the legally required on-board inspection comprise different techniques of measurement of ship emissions remotely. This can be achieved using fixed measurement stations along popular shipping lanes, patrol vessels, and also measuring on manned or unmanned aircraft, i.e. drones. Such methods are used widely for FSC measurements in Europe (Belgium, Denmark, Finland, Germany, the Netherlands and Sweden). In addition, the European Maritime Safety Agency (EMSA) has a drone program conducting measurements using drones in different countries in Europe.

The presented work focuses on best practices in remote monitoring of FSC, using remote gas analysis. It also discusses whether similar methods can be applied to investigate emission levels of NO_x, and whether such methods could also be used for monitoring of particle emissions in the future. The objective was to assess the performance of various techniques for remote measurement of ship emissions, i.e. how well the measurements compare to real world data and whether the claimed measurement uncertainties are realistic. This assessment is based on results from several field campaigns, and it includes a discussion on how to harmonize the measurements with respect to measurements, data validation procedures, uncertainty calculations and reporting.

ES 2 Harmonization

The main objective of this study is to align the measurement procedures of different groups for remote sniffer measurements of FSC and future measurements of NO_x, based on discussions and analysis of the project field campaign results. Three types of techniques have been included, ranging from compact and light weight to more complex high sensitivity ones.

Recommendations are included on how to conduct remote measurements of gas-pollutant emissions from ships, including measurements, data validation procedures, uncertainty calculation and reporting. It is recommended that ship emission data should be reported with expanded uncertainty, corresponding to 95 % confidence intervals (CI 95 %). However, for triggering on-board inspections the standard uncertainty (CI 68 %) should instead be used, by subtracting this value from the measured results before assessing whether a ship is operating on compliant fuel or not.

The calculation of measurement uncertainty was done in different manners by the project partners and similarities and differences in the calculation procedures are highlighted in this report. This should provide a good basis for future standardization work. Two types of uncertainty assessment were applied, i.e. type A and type B. For type A, the uncertainty is calculated for each single measurement, based on the signal to noise ratio, other measured

[1] Global requirement for ship with keel lay date as of 2011

parameters and systematic uncertainties. It is also possible to assess the uncertainty through a combination of assessing the measurement quality and the typical uncertainty associated with such a quality (called type B). Independent of whether type A or type B uncertainty approach is used, it is in any case important to assess whether a ship measurement is valid or not. To do this, many groups use a scoring system for the measurement quality such as a number (0 - 10) or given in text (poor, medium, high). The quality scoring is empirically derived, and it is based on the used sensor technique, weather conditions, signal-to-noise of the various measurements and traceability (calibrations, linearity checks, cross interference calibration) and whether the ships are uniquely identified. The quality scoring should not be part of the reporting since it is given in arbitrary units and hence not comparable between different operators and instruments. Negative values can occur when the measured gas concentrations are near the detection limit; these values should be reported as the minimum detectable FSC based on the SO_2 concentration detection limit (3σ). The corresponding recommendation for NO_x emissions should be applied. Regular comparisons between the different measurement groups are recommended, possibly as intercomparison measurement campaigns or with Round-Robin tests.

ES 3 Validation campaigns

A suite of sensor systems for remote measurements of FSC, fuel-mass specific emissions of NO_x and particles were used and compared in four campaigns, which included areas with different FSC IMO thresholds (0.10 %, 0.50 % and 3.50 % S m/m). The systems included, high-sensitive and standard sniffers, used at fixed locations and mobile manned platforms and compact mini sniffers used on drones. Five project partners carried out these measurements, i.e. Aeromon and Explicit (mature mini sniffer), BSH (2 kinds of standard sniffer systems), Chalmers (experimental mini sniffer, standard and high-sensitive sniffer system, TNO (standard sniffer system).
The error of the remote FSC measurements has been assessed through side-by-side measurements and by comparison of the measurements to on-board taken fuel samples, analysed in certified laboratories. For the validation of fuel-mass specific emissions of NO_x only data from remote side-by-side measurements are available and the ensemble average of these measurements has therefore been used as the assumed "true" value that has been compared to. It is assumed that the measurement error, which is the actual deviation from the true value, corresponds to the measurement uncertainty (with the same statistical distribution), i.e. the predicted deviation from the measured value that should be reported by the instruments.

In campaign 1 (C1), in Marseille 2019, remote measurements were carried out for ships operating with FSC up to 3.5 % S m/m, in the waters of Marseille and Fos-sur-Mer. Side-by-side measurements of 13 ships were carried out from a small and fast measurement vessel using a variant of a standard sniffer and a drone-based mini sniffer system. Note, that in this case there were no fuel samples available, and the true FSC was not known. The used standard sniffer was an airborne variant which was susceptible to organic compounds, in contrast to the standard sniffers on the fixed stations, and this increased the uncertainty. In campaign 2 (C2), in Kiel 2021, side-by-side measurements were carried out by a drone-based system and from two fixed sniffer stations, 1 km apart, one equipped with high-sensitive FSC sniffer and the other with a standard FSC sniffer system. Here, the Stena Germanica ferry, with a known FSC fuel analysis of 0.095 % S m/m, was monitored by remote measurements during 19 individual passages. In campaign 3 (C3), in Wedel 2020, side-by-side measurements using eight different systems, with all available techniques mentioned above, were carried out and compared to onboard fuel analysis of 60 samples from 34 vessels operating with FSC up to 0.1 % S m/m. In campaign 4 (C4), in Marseille, side-by-side measurements with a high-sensitive and a standard sniffer was carried out on ships operating with FSC up to 0.5 % S m/m in the waters of Marseille and Fos-sur-Mer. In all mentioned campaigns above, NOx emissions were monitored by standard NO_x sniffers and drone-based mini sniffer ones. Also PM, using several types sensors, were monitored in several of the campaigns and intercompared in C3.

ES 3.1 Fuel Sulphur Content

It is estimated that the FSC in by-passing ships can be detected with remote measurements with an estimated uncertainty (CI 95 %) varying between 0.03 % and 0.14 % S m/m for ships operating at the FSC level of 0.1 % S m/m,

depending on system, as demonstrated in the field validation campaigns. At higher FSC levels at 1 % S m/m the uncertainty (CI 95 %) appears to increase to about 0.2 % S m/m. This was derived from the measurements errors that were estimated from the difference between the measurements and the "true" values from fuel analysis, bunker receipts or ensemble averages of side-by-side measurements.

In Table ES1 the measurements errors for the individual instruments in campaign C2 and C3 are shown. These were obtained from the difference between measurements on-board fuel analysis. Note that the measurement statistics is rather poor for the fixed systems in C2, corresponding to multiple measurements of a single passenger ferry. For C3 there is considerably more data with a variation in the type of ships and therefore this result is considered to be more reliable. For C1 in Marseille the difference between the measurements and ensemble averages of side-by-side measurements was used as the measurements error. In this campaign the Aeromon (mature) drone-based mini sniffer was compared against a variant of a standard sniffer (made for airborne measurements) at FSC ranges between 0.1 % to 3.5 % S m/m. The comparison showed overall good agreement between the measurements in terms of magnitude and correlation and if the average of the two systems is assumed as the "true" emission, then the expanded uncertainty (CI 95 %) for both instruments is 0.20 % S m/m.

Table ES 1. The obtained measurement errors for the remote measurements performed by the individual instruments in campaign C2 and C3. The errors were obtained from the difference between measurements and on-board fuel analysis.

	Random meas. error (2σ) % S m/m	**Bias** % S m/m	**Total meas. error (CI 95 %)** % S m/m	**Comment**	**Group**
High sensitive sniffer	0.01	-0.03	0.036	C2, 1 fuel sample Single ship, 3 occasions	Chalmers
High sensitive sniffer	0.04	-0.02	0.05	C3, fuel samples 39 ships	Chalmers
Standard sniffer	0.03	-0.067	0.08	C2, 1 fuel sample Single ship, 4 occasions	BSH
Standard sniffers	0.05-0.08	-0.02 to -0.08	0.07 to 0.09	C3, fuel samples 14-39 ships	Chalmers, BSH, TNO
Mature mini sniffer on drone	0.025	0.015	0.03	C2, 1 fuel sample Single ship, 11 occasions	Explicit
Mature mini sniffer on drone	0.08	~0	0.08	C3, fuel samples 14 ships	Explicit
Experimental mini sniffer on drone	0.13	-0.03	0.14	C3, fuel samples 14 ships	Chalmers

The reported uncertainties of the standard and high-sensitive sniffer systems varied between 0.01 % to 0.02 % S m/m and they matched the actual measurement error poorly in C3 (20-30 % of all cases) and not at all in C2, due to negative measurement biases that was observed. For the Explicit drone-based mini sniffer the reported uncertainties were around 0.03 % S m/m and they matched the measurements error in 2 out of 3 of the cases in C3 for 14 individual ships and in 9 out of 10 cases in C2.

An apparent dependence between measurement error and relative humidity for the high-sensitive system and several of the standard sniffer systems were observed in C3. For the high-sensitive system, using only FSC data acquired below RH 80 % the negative bias basically disappeared, and the total uncertainty (CI 95 %) decreased from 0.044 % to 0.028 % S m/m. However, this effect was not observed in C2 and the apparent relative humidity effect hence needs to be studied further.

It was found that there were significant differences in calibration gas concentration accuracies from the same gas manufacturer, i.e. up to 40 % compared to the cylinder certificates, at the low SO_2 levels used (90 ppb to 400 ppb in air). It is therefore recommended to quality assure the bottled concentrations by verification in reference

laboratories with certified equipment and processes, by blind round-robin tests between different groups and by cross-comparison of new cylinders to other, already verified cylinders. It also recommended to (a) study the influence of using dry calibration gas, since the SO_2 sensor is potentially affected by humidity and (b) to further study the NO cross sensitivity of the SO_2 monitor.

ES 3.2 NO_x emissions

It appears that the fuel-mass specific emission of NO_x (g/kg_{fuel}) of by-passing ships can be remotely detected with an estimated uncertainty (CI 95 %) of 8 - 17 g/kg_{fuel} (17 % - 40 %). The reported uncertainties agree reasonably well with observed measurement errors in the field campaigns with some exceptions. These uncertainties were derived from the difference between individual measurements and the "true" values from the ensemble average of all the individual side-by-side measurements.

In more detail, in C1 the measurements of a standard fixed sniffer (Chalmers) and a mature drone-based mini sniffer (Aeromon), had a systematic difference of 5 g/kg_{fuel} (13 %) with 14 g/kg_{fuel} 1σ variability (17 %). This corresponds to an estimated measurement error (CI 95 %) of 7.5 g/kg_{fuel} (17 %), for both systems. The reported uncertainties by the two systems explain the differences in 90 % of the cases. In C2, 19 ships were measured using two, shore based, standard sniffers (Chalmers and BSH) and a mature drone-based mini sniffer (Explicit) and there is an agreement better than 10 % between the data sets. The reported uncertainties by the systems explain the differences in 85 % of the cases. The comparison shows systematic differences and random 1σ variability of -4.8±4.7 g/kg_{fuel} (-8.5±11) %, 4.0±4.9 g/kg_{fuel} (-6.9±11) % and -2±4.3 g/kg_{fuel} (-3.3 ±8.5) % for the NO_x measurements by Chalmers, BSH and Explicit, respectively. This corresponds to a total measurements error (CI 95 %) of 10.6 g/kg_{fuel} (23 %) for the fixed systems and 8.8 g/kg_{fuel} (17 %) for the drone-based mini sniffer. Noteworthy, is that the reported calculated uncertainties, are larger for all systems (20 %) than the apparent measurements error. In C3, several hundred ships of varying size were measured using several fixed sniffers. The reported uncertainties matched the apparent measurement errors in 60 to 70 % of the cases, with exception for one system which only matched in 25 % of the cases, due to underestimation of the uncertainties. The comparison shows systematic differences and random 1σ variability of 5.5±8.4 g/kg_{fuel} (12±19) %, -5.1 ± 8.4 g/kg_{fuel} (-14±20) % and -1.2 ± 8.8 g/kg_{fuel} (1.6±18) % for the NO_x measurements by Chalmers, BSH and TNO, respectively, relative to the ensemble average. This corresponds to a total measurement error (CI 95 %) of approx. 17 g/kg_{fuel} (40 %) for all groups. The measurement error is almost twice as large as obtained in campaign C1 and C2. This is not understood but could potentially be caused by the fact that some of the data were evaluated automatically and that there is potentially a larger spread in ship types and engine load, than in the other campaigns, since the ships were sailing with, and against tide.

The IMO NO_x technical code requires that NO_x emissions of marine diesel engines above 130 kW output power are constrained. The regulation corresponds to varying engine emission limits depending on keel laying date (Tier) and engine and the emission value limit to be controlled, is the brake specific emission ($g{\cdot}NO_x/kWh$,) calculated as a weighted average of multiple engine loads, and this value is heavily weighted towards engine loads above 50 %. The fuel-mass specific emission of NO_x obtained by the remote measurements therefore needs to be converted to the brake specific emission by multiplying with the specific fuel oil consumption (SFOC), for instance obtained from IMO 4th greenhouse study. This value depends on ship type and the engine load and typically varies between 160 and 250 g_{fuel} $(kWh)^{-1}$, with generally increasing values at lower engine loads. The data from C3 obtained on river Elbe, and complementary data by several of the participating groups, shows that the fuel-mass specific NO_x emission, and the brake specific one, varies relatively little above 50 % engine loads for the measured Tier I and Tier II ships. Since the emission value limit in the NO_x technical code is heavily weighted towards higher engine loads this means that the remote measurements can be used to assess and control this value. For engine loads below 50 % more care must be taken when assessing whether a ship complies with the IMO rules for its specific Tier. For Tier III the already existing implementation, when doing type approval of engines, of a Not-To-Exceed limit (NTE) of no more than 50 % above the applicable NO_x emission limit (max 3.4 g/kWh) at <u>any</u> load point means even low loads (<50 %) have a regulatory 'cap' which may be monitored via remote measurements.

However, knowing the SFOC value for each ship is crucial for the factor calculation and an accurate emissions analysis will depend on individual SFOC data confirmed by the shipping companies. This data point is currently not readily available but could be envisaged reported to the authorities going forward for the purpose of NO_x compliance monitoring.

ES 3.3 Particle emission

Fuel-mass specific particulate emissions from ships sailing on the river Elbe were derived by side-by-side remote measurements using a suite of different sensors in campaign 3. In contrast to the gaseous species, the particle measurements are difficult to compare since different instruments measure different properties of the particles (such as different ranges of particle sizes and properties like electro-mobility or optical properties of particles). Nevertheless, a satisfactory agreement was found with differences in the derived particle number- and mass concentration ranging from far below 10 % to maximum 35 %. In general, it was found that 85 % of the emitted particles from vessels had a diameter between 10 and 80 nm and 70 % of the emitted particle mass comes from particles with a size between 40 and 200 nm. It is therefore suggested for possible future regulations on the reduction of particle emissions from seagoing vessels, that the emission of very small particles should be considered (down to 10 nm and 40 nm for particle number- and mass concentration, respectively). These regulations should set limits such as particle number or particulate mass per kg of fuel burned, to make potential violations easier to detect by remote measurement than for NO_x. To monitor particle emissions, sufficiently sensitive and fast instruments must be used to cover the above given size ranges. Total particle emission factors were found to be in the range from $0.8 \cdot 10^{16}$ to $1.5 \cdot 10^{16}$ particles/kg_{fuel} on average, but with significant different emission behaviours for different kind of vessels. However, further studies are needed to validate the found conclusions and to investigate effectiveness of the recommended further actions.

List of abbreviations

σ	standard deviation
aer.uas	UAS based sniffer system operated by Aeromon
AIS	Automatic Identification System
BH-12	Sniffer analyser built and operated by Aeromon
BSH	German Federal Maritime and Hydrographic Agency
bsh.ap	Sniffer system operated by BSH in Wedel (Airpointer)
bsh.hor	Sniffer system operated by BSH in Wedel (Horiba)
bsh.mms	Mobile measurement sniffer system operated by BSH
bsh.std	Standard Sniffer system operated by BSH
BSH-MMS	Mobile Measurement Station of the BSH
cha.uas	UAS based sniffer system operated by Chalmers
cha.las	Sniffer system operated by Chalmers based on laser-spectrometry
cha.std	Standard Sniffer system operated by Chalmers
CI	Confidence Interval
CPC	Condensation Particle Counter
DOAS	
d_p	Particle Diameter
EC	ElectroChemical
EEPS	Engine Exhaust Particle Sizer
EMSA	European Maritime Safety Agency
EPC	Environmental Particle Counter
EU	European Union
exp.uas	UAS based sniffer system operated by Explicit
FMPS	Fast Mobility Particle Sizer
FSC	Fuel Sulphur Content
FTIR	
GHG	Green House Gas
GUM	Good Laboratory Praxis
LASX-II	Airborne-particle spectrometer
MDO	Marine Diesel Oil
NDIR	Non dispersive Infra Red
NTE	Not-To-Exceed
OPS	Optical Particle Sizer
PMX	Particulate Matter, where x denotes the diameter (in µm) up to which the particles are considered
RH	Relative Humidity
RPAS	Remotely Piloted Aircraft System
SFOC	Specific Fuel Oil Consumption
S/N	Signal to Noise ratio
TNO	the Netherlands Organisation for applied scientific research
tno.std	Sniffer system operated by TNO
UAV	Unmanned Aerial Vehicle
UAS	Unmanned Aerial System
UFP	Ultrafine Particles
UTC	Coordinated Universal time

Contents

1 Introduction 14
2 Remote Measurement techniques 16
2.1 Sensors for measuring ship pollutant concentrations 16
2.1.1 Standard system 17
2.1.2 High-sensitive system 18
2.1.3 Mini sniffer system 18
2.2 Remote measurements of FSC and fuel specific NOx emissions using sniffer systems 18
2.3 Implementation of standard sensors for fixed ship emission monitoring 20
2.3.1 BSH standard sniffer system 20
2.3.2 Chalmers standard sniffer system 23
2.3.3 TNO standard sniffer system 28
2.4 Implementation of high-sensitive laser instrument for ship emission monitoring 29
2.4.1 Calibration 29
2.4.2 Cross interference 30
2.4.3 Quality criteria 31
2.4.4 Calculation of uncertainty 31
2.5 Implementation of mini sniffer sensors for ship emission monitoring from drones and helicopters 31
2.5.1 Explicit 31
2.5.2 Aeromon 33
2.5.3 Chalmers 36
2.6 Other methods applied within project 41
3 Intercomparisons and validation measurements 42
3.1 Campaign activities 42
3.1.1 Marseille campaign, September 2019 (C1) 42
3.1.2 Wedel campaign, September 2020 (C3) 42
3.1.3 Stena Germanica and onshore Kiel Campaign, September 2021 (C2) 43
3.1.4 Campaign 4: Marseille, July 2021 (C4) 44
3.2 Key results for FSC and NO_x 45
3.2.1 FSC measurements 45
3.2.2 NO_x emissions 51
4 Current status of particle measurement techniques for ship emission monitoring 57
4.1 General description 57
4.2 Instrumentation 57
4.3 Preliminary results on particle emissions 59
4.4 Further Comparison of measured particle number- and mass emission factors from ship plumes measured during C3 in Wedel 62
5 Recommendations for harmonized ship emission monitoring 65
5.1 Measurements 65
5.1.1 Physical system 65

5.1.2 Performance requirements 65
5.1.3 Quality control 67
5.1.4 Measurement methodology and calculation of FSC and fuel specific emission factors 68
5.2 Data validation procedures 68
5.3 Calculation of uncertainty 69
5.3.1 Basic introduction 69
5.3.2 Implementation 70
5.4 Reporting 72
6 Conclusions 73
6.1.1 Harmonization 73
6.1.2 Field validation FSC 73
6.1.3 Field validation NO_x 74
6.1.4 Particles 76
7 Acknowledgements 77
8 References 78

Figures

Figure 1. The European SECA area which requires fuel with FSC of 0.10% or less (for non-scrubber ships). From 2021 this area is also a NO_x emission control area, requiring Tier III (approx. 90 % NO_x abatement) for all ships from this year and onward. The area extends all the way to 5°W in the English Channel. 14
Figure 2. The IMO MARPOL Annex VI NO_x emission limits for ship engines. The emission levels depend on ship build year (keel laying date), divided into three tiers, and engine type. The engine type is defined according to the corresponding rated engine speed. The emission corresponds to g NO_x per axial power in kWh, corresponding to a weighted average of several engine loads. Tier III applies in ECAs in north America (ships built from 2016) and Northern Europe (ships built from 2021). 20
Figure 3. Estimated span calibrations coefficients of CO_2, SO_2 and NOx for the Chalmers standard system throughout the period corresponding to the Wedel field campaign. 24
Figure 4. Frequency distribution of the relative variation in calibration factors of CO_2, SO_2 and NOx from Figure 3 for the Chalmers standard system. 24
Figure 5. Measured plume and uncertainty parameters 27
Figure 6. Example of a database query corresponding to FSC measurements at the Älvsborg site in Göteborg. 27
Figure 7. Example of an animation generated by the Chalmers Systems corresponding to a non-compliance vessel measured at the Älvsborg site in Göteborg. 28
Figure 8. Estimated span calibrations coefficients of CO_2 and SO_2 for the Chalmers Laser system throughout the period corresponding to the Wedel field campaign 30
Figure 9. Frequency distribution of the relative variation in calibration factors of CO_2, and SO_2 from Figure 8 for the Chalmers Laser system 30
Figure 10. The flight approach used by Explicit. 32
Figure 11. Aeromon BH-12 mini sniffer 33
Figure 12. Daily calibration level check is always performed through the sample line in use. 34
Figure 13. Data flow schematics example for reporting. 35
Figure 14. Ship plume direction (Berg et. al. 2012) 36
Figure 15. Vessel measurement example showing CO_2 (above atmospheric background), SO_2, NO_2 and NO signals. 36
Figure 16. Mini Sniffer developed by Chalmers University 37
Figure 17. Schematics of the Chalmers drone sniffer box. 37
Figure 18. Difference between SO_2 readings from electrochemical sensor as compared to the standard sniffer system. 38
Figure 19. SO_2 cross interference of the NO_2 38
Figure 20. Example of the IGPS software display. A) Position display of the mini sniffer and targeted vessel. B) Real time plume measurements. 40
Figure 21. Example of column measurements of SO_2 and NO_2 in the plume of different ships during C4 using zenith sky DOAS by Chalmers. The measurements were carried out from a measurement vessel. The y-axis In the left graph corresponds to the column in mg m^{-2} 41
Figure 22. SO_2 over NO_2 column ratios from single ship plumes and comparison to FSC from the high-sensitive laser sniffer system. The numbers in parenthesis correspond to the length of the ship and the distance to the ship in meters. 41
Figure 23. The measurement area in campaign 1 and 4. The underlying satellite picture is taken from google earth. 42
Figure 24. Setup of the side-by-side inter comparison systems at the BSH measurement site in Wedel. Left to right: Chalmers (white container and blue VW transporter car), BSH mobile measurement station (white trailer), TNO (artistic trailer), Explicit (hovering drone), BSH (elevated platform at building). 43
Figure 25. Location of remote measurement carried out by the project partners Explicit, BSH and Chalmers during SCIPPER C2 campaign in the Kiel Fjord (Germany). The underlying satellite picture is taken from google maps (https://google.de/maps/). 44
Figure 26. Comparison of sniffer (Chalmers) and drone (Aeromon) measurements of different ships operating on the waters of Marseille and Fos-sur-Mer during September 2019 (C1). 45
Figure 27. Summary of the measurement error, corresponding to absolute difference between remote FSC measurements and laboratory-analysed fuel samples (main engine), for about 40 measured ship plumes. Here bsh.hor, bsh.ap and bsh.mms represent the three standard sniffers operated by BSH and cha.std, cha.las and cha.uas correspond to the results from the standard-, laser- and the mini sniffer systems operated by Chalmers. The TNO

standard sniffer results and Explicit mini sniffer results are indicated with tno.std and exp.uas, respectively. The fuel samples used for this comparison were taken from the fuel system of the main engine. 46
Figure 28. The measurement error versus the reported uncertainty which is reported by the respective systems for measurements obtained during C3 in Wedel. Different colours indicate the respective results of the different systems. The data points within the unshaded area indicate results which match the laboratory found FSC values within the reported uncertainty. Results in the grey-shaded area are either under- or overestimating the FSC when taking into account the individual measurement uncertainty. 47
Figure 29. Absolute deviation of measured FSC in by-passing ships by the different remote systems to the laboratory analysed fuel samples. 48
Figure 30. Comparison of the on-board and Kiel shore-based measurements of the FSC from Stena Germanica during SCIPPER campaign C2 in August and September 2021. Three systems were operated side-by-side, i.e. a high-sensitive fixed sniffer (cha), a standard fixed sniffer (bsh) and UAV-based mini sniffer system (exp) 49
Figure 31. Side-by-side measurements in Kiel during campaign C2 in August and September 2021. Three systems were operated side-by-side, i.e. a high-sensitive fixed sniffer (cha), a standard fixed sniffer (bsh) and UAV-based mini sniffer system (exp). 50
Figure 32. Measurements of 22 ships outside Marseille in 2021 during C4 using the Chalmers laser and standard sniffer instrument. 50
Figure 33. Comparison between Chalmers sniffer and Aeromon drone during C1 in Marseille in September 2019. Measurements were carried out from a small and fast vessel outside the water of Marseille and Fos-sur-Mer. 51
Figure 34. Comparison of NO_x emission factors (g/kg_{fuel}) for individual ships from fixed sniffers by Chalmers and BSH and drone-based mini sniffer by Explicit during C2 in Kiel. The fixed sites were positioned on the eastern shore of the ship channel with 1 km separation. 52
Figure 35. NO_x EF data (g/kg_{fuel}) for individual ship emission measurements by different groups versus the ensemble average. Data with two or three coincident measurements have been used. 53
Figure 36. Statistics (median in red and 5th, 25th ,75th, 95th percentiles in blue and black) of the absolute and relative deviation of NOx emission factors measured individually by Chalmers, BSH and TNO for the same ship, against the corresponding ensemble average. Data were measured in Wedel 2020 during C3 and only data with three coincident measurements were used. The numbers given on top correspond to average value and the standard deviation. 53
Figure 37. Multiple measurements during C3 by several instruments of the fuel-mass specific NO_x emission rates for a ship (Suction Dredger) that passed the Wedel site many times. The upper part of the figure shows statistics (median in red and 5th, 25th ,75th, 95th percentiles in blue and black) of the absolute and relative deviation of NO_x emission factors (g/kg_{fuel}) measured individually by Chalmers, BSH and TNO for the same ship, against the corresponding ensemble average. The numbers given on top correspond to average value and the standard deviation. The lower figure shows the measured emission factors of NO_x by the different systems versus time, and estimated measurements uncertainty. 54
Figure 38. The absolute deviation of the NO_x EF measurements by the individual groups in C3 relative to the ensemble mean is plotted versus the estimated uncertainty. 54
Figure 39. The ship's emissions of NO_x versus load, measured in the Elbe river approx. 10 km from the entrance to Hamburg harbour. a) Break specific mass emission (g/kWh) b) fuel-mass specific emission(g/kg). The data correspond to measurements by Chalmers standard sniffer for Ships that weigh more than 75 ktons during C3. 56
Figure 40. Combined particle size spectrometer (TSI Fast Mobility Particle Sizer + TSI Optical Particle Sizer) installed into a climate-controlled shelter. Sample air is dried with a Nafion counterflow dryer. The PM-10 inlet prevents particles larger than 10 µm to enter the system. 58
Figure 41. Comparison of particle mass concentrations measured in ship plumes with different combined particle size spectrometers operated by BSH (TSI Fast Mobility Particle Sizer + TSI Optical Particle Counter) and Chalmers University (TSI Engine Exhaust Particle Sizer + TSI Optical Particle Counter). The left graph shows the comparison of the two identical systems operated by BSH during SCIPPER C5. The right graph compares one of the BSH systems with the very similar one (EEPS instead of FMPS) operated by Chalmers also during C3 in Wedel, but without Nafion counterflow dryer 59
Figure 42. Time series of the number concentration of particle (a), particle size distribution (b) and soot concentration (c), also known as black carbon, BC) on 29th of September 2020 measured in Wedel, Germany during the C3 campaign 60
Figure 43. Cumulative fraction (in %) according to particle diameter for particle number and mass. 61
Figure 44. Integral particle emission factors for different ship types, measured between September 2020 and April 2022 at the two BSH measurement sites in Wedel (Hamburg approach) and Kiel (approach to Kiel and Kiel Canal). 62
Figure 45. Scatterplot of the particle mass emissions measured during the C3 campaign in Wedel, Germany in a size range of 0.09 and 0.3 µm. The BSH-AP system is taken as a reference on the x-axis and on the y-axis the BSH-MMS

(left top), TNO-LASX II system (right top) and Chalmers (left bottom) are plotted. The plots include the one-to-one line and the corresponding regression statistics. 63
Figure 46. Scatterplot of the emission of the particle number measured during the C3 campaign in Wedel, Germany in a size range of 0.09 and 0.3 µm. The BSH-AP system is taken as a reference on the x-axis and on the y-axis the BSH-MMS (left top), TNO-LASX II system (right top), and the Chalmers-EEPS (left bottom) are plotted. The plots include the one-to-one line and the corresponding regression statistics. 64

Tables

Table 1. The instruments employed for remote gas measurements of ship emissions in the SCIPPER project. The platform are denoted with a number correspond to (1) Drone, (2) shipborne, (3) fixed land based fixed. The campaigns correspond to C1 to C4. The different instruments used by each partner are denoted according to the list of abbreviations. 16
Table 2. Standard deviation of relative variation in calibration factor with daily calibration frequency during the C3 Wedel measurement campaign using the standard system. 25
Table 3. Example of the quality criteria for Chalmers applied for fixed measurements using standard system at the Great Belt site. . 25
Table 4. Standard deviation of relative variation in calibration factor with daily calibration frequency during the C3 Wedel measurement campaign using the high-sensitive laser system. 29
Table 5. Summarizing table of absolute deviations of FSC from laboratory analysed fuel sample for all measured plumes per system during C3 in Wedel. The fuel samples used for this comparison were taken from the fuel system of the main engine. For clarifying the abbreviations for system see Figure 27. 47
Table 6. The number of NO_x measurements for the deviation of the NO_x EF measurements of the individual groups relative to the ensemble mean is categorized according to the estimated uncertainty. 55
Table 7. Quality criteria that are used by the individual groups on fixed stations (BSH, Chalmers, TNO) and drones (Aeromon, Chalmers, Explicit) in the data validation procedures. Change names to instrument. 68
Table 8. Summary of Implementation of the uncertainty analysis of each group. 71
Table 9. Suggested parameters to report related to remote FSC measurements. 72

1 Introduction

In 2015 new rules from the International Maritime Organization (IMO), European directives (EU 1999; 2012) and US legislation require ships to run with maximum fuel Sulphur content (FSC) of 0.10 % S m/m in Northern European (Figure 1) and US waters. Scrubber ships usually need to monitor their SO_2/CO_2 emission and report exceedances. Since 2020, ships worldwide are only allowed to operate with a maximum FSC of 0.50 % S m/m outside the Sulphur emission control areas. Ships built from year 2000 and 2011, respectively should follow Tier I and Tier II NOx emissions From 2021, Northern Europe is a NO_x emission control area, requiring Tier III (more than 80 % NO_x reduction) for all ships built (keel laid) from this year and onward, compared to ships built 2000-2010. There is presently a discussion within IMO how to control particle emission (black carbon).
To verify that ships comply with the FSC legislation, on-board fuel sampling is carried out by port state control authorities for ships at berth. According to the European Sulphur directive, for each member state it is required to physically inspect every year 10 % of the ships calling on a harbour in the country and 40% of these inspected vessels should be additionally checked with fuel sampling. Hence, a fuel sample is taken from only a minority of the ships (4 %), and none while the ships are sailing in open waters (2015/253/EC Article 3). When using advanced monitoring strategies such as remote measurements or onboard monitoring with XRF (X-Ray Fluorescence), it is allowed according to the European Sulphur directive, to reduce the number of fuel sample analyses by up to 50 %. To verify NO_x compliance, on-board inspection cannot be carried out by measurements, and port state control authorities instead check that the ships carry the appropriate documentation with respect to type approvals of engines and service and logging routines.

Figure 1. The European SECA area which requires fuel with FSC of 0.10% or less (for non-scrubber ships). From 2021 this area is also a NO_x emission control area, requiring Tier III (approx. 75 % NO_x abatement compared to Tier II) for all ships from this year and onward. The area extends all the way to 5°W in the English Channel.

There are different types of remote systems for monitoring ship emissions. Stationary systems near shipping lanes are used to analyse the chemical composition of exhaust plume of ships passing by. These systems can run continuously and autonomously, positioned up to a few kilometres downwind of the ship lanes on for instance, bridges and harbours. The same instrumentation as for the stationary systems can also be used from mobile platforms such as harbour patrol vessels and fixed wing and rotor aircraft to monitor ships under way in open sea. More compact and lightweight mini sniffer systems are also in operation for compliance monitoring using drones or helicopters in both, near-shore and off-shore sea regions. The mini sniffer systems work in higher concentration ranges and hence typical sampling distances are in the order of 25 to 100 m from the exhaust gas stacks.
Various remote techniques have been developed in the last two decades in both national and EU projects. In 2009 there was an EU financed project (Berg, 2012; Alföldy, 2013; Balzani-Lööv, 2014) where side-by-side measurements were carried out in Rotterdam and Genua using sniffer and optical systems operated from fixed stations, aircraft and drones. These measurements were further used and developed in the EU CEF project CompMon (Mellqvist, 2017) and the Baltic Sea projects InnoShip (Beecken, 2014) and EnviSum (Repka et al, 2019). The different systems and operators in Europe are described in more detail in the SCIPPER D2.1 report (Beecken et al, 2019). However, the

parallel development of measurement and monitoring strategies in different countries has led to a lack of internationally uniform and comparable measurements. Although the measurements follow the same basic principle, they differ in the way uncertainties are calculated and measurements are reported to inspection authorities.
This report focuses on best practice in remote compliance monitoring of FSC. It also discusses whether or not the same methods can be applied to investigate if ships comply with the NO_x IMO limits and future particle measurements. The report describes results from several measurement campaigns within SCIPPER in which the available techniques have been tested against each other and against fuel samples and onboard measurements. Furthermore, it describes in detail how different operators calculate their uncertainties and makes suggestions on how the data should be reported in a unified way.

2 Remote Measurement techniques

2.1 Sensors for measuring ship pollutant concentrations

The instruments used for remote measurements of gases in the SCIPPER campaigns are described in this section. More details can be found in the SCIPPER D2.1 report (Beecken et al. 2019). Several types of platforms are employed, i.e. drones, manned aircraft, measurement vessels as well as fixed- and mobile land-based sites. The sniffer systems can be divided into three categories: standard, high-sensitive and mini sniffer systems. The mini sniffer systems that are utilized from drones or helicopters can be less sensitive and slower in their response than required for instruments on the other platforms, since they are employed close to the exhaust gas stacks of the ships can remain for a longer time period in the highly concentrated part of the plume. In Table 1 technical parameters for these instruments are described together with information about platforms, operated instruments by the different SCIPPER partners and in which SCIPPER campaigns they were employed.

Table 1. The instruments employed for remote gas measurements of ship emissions in the SCIPPER project. The platforms are denoted with a number corresponding to (d) drone, (s) shipborne, (l) fixed land-based. The campaigns correspond to C1 to C4. The different instruments used by each partner are denoted according to the list of abbreviations.

Species	Method	Model	t90 of setup	Det limit	Cross sensitivity	Platform& campaign	partner. instr.type
CO_2	Cavity ring down laser spectrometer	Picarro G-2301m	1 s	0.2 ppm		s,l C1,C2, C3, C4	CHA.std
CO_2	Laser spectrometer	Aerodyne	1 s	0.2 ppm		d,s,l C1	CHA.las
CO_2	NDIR	LI-COR 7000, 7200	1 s 1 s	0.2 ppm		s,l C1,C2, C3,C4	TNO.std CHA.std
CO_2	NDIR	LI-COR 840A	1 s	0.2 ppm		s,l C2, C3	BSH.std
SO_2	Fluorescence	Thermo 43i-TLE	40 s*	3 ppb	1.5% NO VOC*	s,l C1,C2, C3, C4	CHA.std TNO.std
SO_2	Fluorescence	Airborne Thermo 43i-TLE with kicker removed and high flow pump	2 s	3 ppb	1.5% NO VOC*	s,l C1,C2, C3, C4	CHA.std
SO_2	Fluorescence	HORIBA APSA-370	40 s	0.5 ppb (at 5 s)	0.5-1.0% NO	l C2, C3	BSH.std
SO_2	Fluorescence	mlu-recordum airpointer	40 s	2 ppb (at 10 s)	0.7-1.5% NO	s,l C2, C3	BSH.std
NO_x	Laser spectrometer	Aerodyne	1 s	0.06		s,l C2,C3,C4	CHA.las
NO_x	Chemiluminescence	Thermo 42i-TL	1 s	1 ppb		s,l C1,C2, C3,C4	CHA.las TNO.std
NO_x	Fluorescence	HORIBA APNA-370	30 s	1 ppb (at 5 s)		l C2, C3	BSH.std
NO_x		mlu-recordum airpointer	30 s	1 ppb (at 10 s)		s,l C2, C3	BSH.std
SO_2	Electrochemical	Alpha sense SO_2-A4	20s	~15 ppb	-120% NO_2	d C2, C3	CHA.uas
SO_2	Electrochemical	Aeromon SO_2	<25 s	~20 ppb	< -100% NO2	d,s C1, C2, C4	AERO.uas

SO_2	Electrochemical	Explicit SO_2	~20 s	~20 ppb	-120% NO_2	d C2, C3	EXP.uas
NO	Electrochemical	Alpha sense NO-A4**	25 s	~10 ppb	5% NO_2	d C2, C3	CHA.uas
NO	Electrochemical	Aeromon NO	<20 s	100 ppb		d, s C1, C2, C4	AERO.uas
NO	Electrochemical	Explicit NO**	<25 s	120 ppb	<7% NO_2	d C2, C3	EXP.uas
NO_2	Electrochemical	Alpha sense NO2-A43F**	80 s	~15 ppb	5% NO	d C2, C3	CHA.uas
NO_2	Electrochemical	Aeromon NO_2	<35 s	~20 ppb	<10% NO	d, s C1, C2, C4	AERO.uas
NO_2	Electrochemical	Explicit NO_2 **	<80 s	~20 ppb	<5% NO	d C2, C3	EXP.uas
CO_2	NDIR	Smartgas F3-212205	1 s	8 ppm		d C2, C3	CHA.uas
CO_2	NDIR	Aeromon CO_2	<15 s	NA		d, s C1, C2, C4	AERO.uas
CO_2	NDIR	Explicit CO_2	20 s	NA		d C2, C3	EXP.uas

* Airborne version has 2 s response time, but it is then susceptible to VOCs.
** Response time corresponds to t_{90} to 2 ppm.
ΔCO_2 correspond to elevated concentrations above the ambient levels.

2.1.1 Standard system

Standard system refers to systems based on conventional air quality instruments, which are also used for air quality measurements by national agencies. These standard systems can be employed on fixed land-based stations as well as shipborne. In some cases, they have been modified by change of pumps or removal of diffusion tubes to increase the instrument response time. The standard CO_2 instrument is a Non-Dispersive InfraRed instrument (NDIR) which is an optical instrument that measures infrared absorption in two wavelength bands around 5 µm by using a broadband light source and bandpass filters. In these wavelength bands the species H_2O and CO_2 show strong, characteristic absorption features. The instrument includes two measurement cells, one sample cell and one reference cell containing a known concentration of CO and H_2O. The concentration in the former cell is obtained by calculating the light absorption due to CO_2 and H_2O by comparing the intensities in the two cells. The instrument 's response is non-linear and it is therefore calibrated by multiple gas concentrations when manufactured. Furthermore, when operated in the field in this project, the calibration curve has been corrected by a span gas calibration in the measurement range of interest. The H_2O obtained from the instrument is used by the LI-COR software to correct the CO_2 concentration values, since H_2O interferes weakly with CO_2.
CO_2 measurements, A *Cavity Ring Down laser spectrometer* is optionally used (O'Keefe and Deacon, 1988) for CO_2 measurements, replacing the Licor. This system works with semi-conductor lasers in the 1.5 µm range combined with a multireflection cell, in which long pathlengths (up to 15 km) can be obtained, and hence large sensitivity. This type of instrument is today the backbone in global CO_2 monitoring. The instrument has a linear concentration response curve, and it is very stable and regular calibration is therefore not required.
The standard SO_2 instrument is a UV fluorescence instrument, which is based on a pulsed UV lamp, and a bandpass filter, that excites the SO_2 in the measurement cell by light in the wavelength region 220-230 nm. This causes the SO_2 to fluoresce around 300 nm and this is detected by a photomultiplier tube and then recalculated to a SO_2 gas concentration. Also, other gases, such as NO and aromatic hydrocarbons reacts in similar manner and hence cause interference, Typically NO causes an interference of 0.5-1.5 % of the NO reading, hence 100 ppb NO will be interpreted as 0.5-1.5 ppb SO_2. There is also an interference with hydrocarbons from lubrication oil, especially for the airborne version for which the "kicker" is removed, i.e. a diffusion tube that absorbs aromatic VOCs before the gas enters the instrument. The SO_2 instrument has a linear concentration response.

The standard NO_x instrument is based on the chemiluminescence technique (Kley and McFarland, 1980) which measures the sum of NO and NO_2. This is conducted, first, by leading the gas through a heated molybdenum converter that converts NO_2 to NO. The NO is then measured by the chemiluminescence reaction in which ozone is reacted with NO to excited NO_2 (NO_2*), which emits light (luminescence) that can be measured by a photomultiplier, and whose intensity is proportional to the concentration. The response of the instrument is favoured by low pressure and is affected by humidity. The NOx instrument has a linear concentration response.

2.1.2 High-sensitive system

During the last two decades various laser-based sensors have emerged on the market. In the SCIPPER project two types of such instruments have been employed.
For simultaneous measurements of SO_2 and CO_2, a *laser spectrometer* from Aerodyne was acquired which is based on two independent quantum cascade lasers working around 1300 cm^{-1} and 2200 cm^{-1}, respectively. The light is transmitted into a multireflection cell (Herriot cell), and then it is reflected numerous times so the total path length corresponds to 400 m. The wavelength of the light is scanned over a certain wavelength region, around 1300 cm^{-1}, in which the key species absorb the infrared light. The measured species include truly simultaneous measurements of SO_2, CO_2, CH4, CO and H_2O. For SO_2 a few tens of ppt can be reached in sensitivity at 1 s response time while 60 ppb for CO_2 at ambient levels. The instrument has a linear concentration response, is very stable and only shows small drift and in regular calibration are therefore not required.

2.1.3 Mini sniffer system

Mini sniffer systems can be carried on drones and helicopters. They are lightweight, reasonably fast yet slower than the standard and high-sensitive systems and sensitive enough to detect pollutants at about 50 m from the exhaust gas stacks. CO_2 is generally measured by small NDIR sensors, similar to the ones used by the standard system, but with shorter optical cavity and hence less sensitivity. This sensor has a nonlinear concentration response curve.
Other ship exhaust pollutants as SO_2, NO and NO_2 are measured by electro-chemical sensors (EC). In these sensors, the gas diffuses through a selective membrane into a cell with two electrodes and an electrolyte reservoir. These sensors show a linear response, and the response time of the sensors is inversely proportional to the concentration gradient. The EC sensors are known to be sensitive to relative humidity (RH) in the air sample and show strong cross-interference to other species. For instance, the SO_2 sensor for NO_2 used by Chalmers will read –120 ppb for 100 ppb of NO_2. At the same time the NO_2 sensors show 5 % cross sensitivity to NO, corresponding to 5 ppb reading for 100 ppb NO. Hence, each measurement corresponds to a chain of several measurements and for each of these, the quality needs to be controlled.

2.2 Remote measurements of FSC and fuel specific NOx emissions using sniffer systems

Emissions from ships under way can be measured from fixed stations, aircrafts or measurement vessels using standard sniffer systems or mini sniffer systems on drones or helicopters.
The sniffer system measures the sum of the background concentration and the concentration elevations caused by the emissions from the ship. The FSC is given as the percentage of the mass of sulphur to the mass of fuel and expressed as % S m/m. To determine the FSC of the ship, the background concentration must be subtracted from the measured concentration to determine the true ratio between the SO_2 and CO_2 in the exhaust gas. Therefore, the FSC can be determined according to Eq.1 with the units used as given in the subscript (Balzani Lööv et al., 2014, Beecken et al., 2014, Beecken et al., 2015, Kattner et al., 2015, and Mellqvist et al., 2017):

$$FSC_{\%S_{m/m}} = \frac{M(S)_{g/mol} \times \int [SO_2]_{ppb} - [SO_{2,bgd}]_{ppb} dt}{\frac{M(C)_{g/mol}}{0.87} \times \int [CO_2]_{ppm} - [CO_{2,bgd}]_{ppm} dt} = 0.232 \times \frac{\int [SO_2]_{ppb} - [SO_{2,bgd}]_{ppb} dt}{\int [CO_2]_{ppm} - [CO_{2,bgd}]_{ppm} dt} \quad \text{Eq. 1}$$

$M(X)_{g/mol}$ denotes the molecular weight of species X in g/mol and $-Y_{ppb}$ the volume mixing ratio of species Y in ppb. Taking the ratio between these substances cancels out dilution in the ambient atmosphere so that the FSC can effectively be determined remotely even hundreds of metres downwind of the vessel, as long as the contribution of the vessel can be separated from the background variability.

Similarly, the emission factor of NOx with respect to the consumed fuel can be calculated as:

$$EF_{\left(\frac{g}{kgfuel}\right)} = \frac{M(NO_2)\frac{g}{mol} \times \int [NO_x]_{ppb} - [NO_{x,bgd}]_{ppb} dt}{\frac{M(C)\frac{g}{mol}}{0.87} \times \int [CO_2]_{ppm} - [CO_{2,bgd}]_{ppm} dt} = 3.33 \times \frac{\int [NO_x]_{ppb} - [NO_{x,bgd}]_{ppb} dt}{\int [CO_2]_{ppm} - [CO_{2,bgd}]_{ppm} dt} \quad \text{Eq. 2}$$

The NO_x measurements can be converted to brake specific emissions in g/kWh using Eq. **3**. Here the brake specific fuel oil consumption (SFOC) has been calculated from Eq, 4 following recommendations in the IMO 4th GHG study. The baseline SFOC ($SFOC_{base}$) was obtained assuming that all ships used MDO fuel and the year the ship was manufactured (obtained from IHS Sea-web). The engine load can be calculated from the actual ship speed at the time of the monitoring, as reported by AIS, using the propeller law. Here we have used the STEAM model (Jalkanen 2009) which uses the propeller law and individual hull factors to calculate the power usage for the propulsion.

$$EF_{NOx}\left(\frac{g}{kWh}\right) = SFOC \cdot EF_{NOx}\left(\frac{g}{kgfuel}\right) \quad \text{Eq. 3}$$

$$SFOC = SFOC_{base} \cdot (0.455 \cdot Engine\_Load^2 - 0.710 \cdot Engine\_Load + 1.280) \quad \text{Eq. 4}$$

The IMO NOx technical code ((MEPC.177(58) and MEPC.251.(66)) requires that NO_x emissions of marine diesel engines above 130 kW output power are limited, see Figure 2. The regulation corresponds to varying engine emission limits depending on keel laying date and engine type. The engine type is defined according to the rated engine speed (crankshaft revolutions per minute). The weighted average emission level (E_{avg}) is given as brake specific emissions with unit gram NO_x (as NO_2) per axial power in kWh. The weighted emission average value is calculated according to Eq, 5 based on the g/h emission factors (q_i) at 4 engine operation modes (i) with different engine loads (P_i) and weighting factors (wf_i).

In general, the ships to be monitored follows the propeller law for their main propulsion and the weighting factors then correspond to 0.15, 0.15, 0.5 and 0.2 for the engine load points 25 %, 50 %, 75 % and 100 %, respectively,

$$E_{avg}\left(\frac{g}{kWh}\right) = \frac{\sum_{i=1}^{n} q_i\left(\frac{g}{h}\right) \cdot wf_i}{\sum_{i=1}^{n} P_i(kW) \cdot wf_i} \quad \text{Eq. 5}$$

Instead of using g/h emission values and actual power, the calculation can also be done by using the g/kWh values and relative power in percent ($P_{rel,i}$) in the following way (PROMINENT D5.8, 2017):

$$E_{avg}\left(\frac{g}{kWh}\right) = \frac{\sum_{i=1}^{n} E_i\left(\frac{g}{kWh}\right) \cdot P_{rel,i}(kW) \cdot wf_i}{\sum_{i=1}^{n} P_{rel,i}(kW) \cdot wf_i} \quad \text{Eq. 6}$$

Different levels (Tiers) of control apply based on the ship's keel laying date and the engine's rated speed given as crankshaft revolutions per minute. Tier I applies for ships built (keel layed) in 2000-2010, Tier II applies to ships keel layed after 2011 and Tier III applies to ships operating in special emission control areas (Northern Europe and North America). In more detail Tier III applies to engines that are installed on a ship (keel laying date) after December 31 2015 for ships in North American ECA and the United States Caribbean Sea ECA and after December 31 2020 for ships operating in the Baltic Sea ECA or the North Sea ECA. As seen in Figure 2 the limits for slow speed ships, rated engine speed below 130 rev/min, corresponds to 17, 14.4. and 3.4 g/kWh for Tier 1, Tier II and Tier III, respectively.

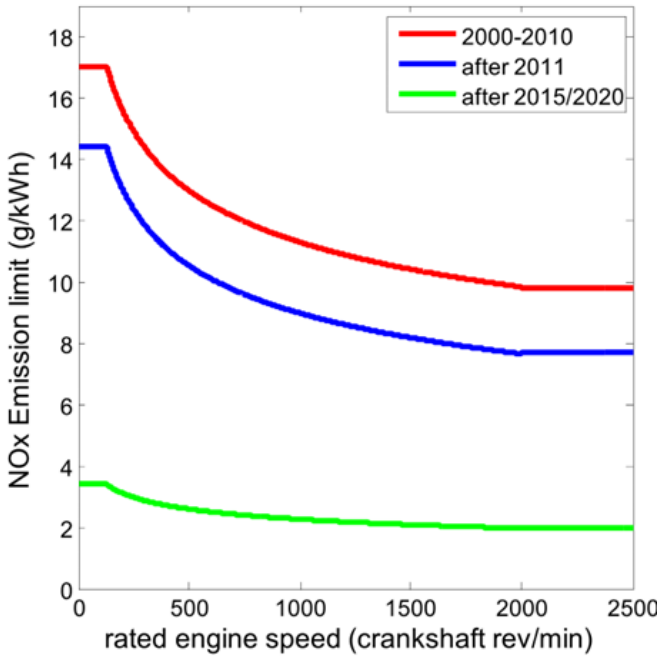

Figure 2. The IMO MARPOL Annex VI NO_x emission limits for ship engines. The emission levels depend on ship build year (keel laying date), divided into three tiers, and engine type. The engine type is defined according to the corresponding rated engine speed. The emission corresponds to g NO_x per axial power in kWh, corresponding to a weighted average of several engine loads. Tier III applies in ECAs in north America (ships built from 2016) and Northern Europe (ships built from 2021).

2.3 Implementation of standard sensors for fixed ship emission monitoring

Below the specific implementation and the quality assurance procedure of each partner are described. The information and text for each system has been provided by the individual SCIPPER partner.

2.3.1 BSH standard sniffer system

2.3.1.1 Calibration

The German ship emission monitoring sites are operated by BSH. The calibration procedure for the sniffer instruments consists of two main procedures, i.e. daily automatic internal function control and half-yearly calibration with certified gas cylinders. Every 25 hours an automatic internal function control is performed for the SO_2 and NO_x modules. A zero-concentration control is carried out followed by a span gas control. In the first step zero air is provided to the measurement modules for 720 seconds by pulling ambient air through a scrubber removing all SO_2 and NO_x. After purging the system for 600 s, the zero-concentration response check takes place. The scubber is filled with black charcoal to scrub SO_2 and with Purafil® to scrub NO_x in the ambient air. When the zero deviation exceeds 1.5 ppb, the modules were adjusted with the internally generated zero concentration air. In a second step SPAN gas is provided to the modules. The SO_2 and NOx modules are equipped with internal SPAN modules. Permeation tubes with selected permeation rates are installed into permeation ovens to provide concentrated SPAN gas of SO_2 of around 300 ppb and NO_2 of around 370 ppb, respectively. The duration of active SPAN valve is set to 720 s, as well. Like for the zero-concentration response check, after 600 s of purging, the response check takes place. For measurement and adjustments, the operating temperature of the whole system is stabilized to 50 °C.
Monitoring the data graphically over a long period of time (1-3 months), plausibility checks and instrument drift behaviour can be identified and evaluated. The SPAN deviation is only used for this quality checks, but not for short term instrument response adjustment. When the deviation exceeds 50%, maintenance is needed. Due to technical limitations no internal function control could be performed for the CO_2 module.
External calibrations with certified gas cylinders are carried out half-yearly directly on site at the stations. Because certain comparable environmental conditions are needed, i.e. temperature between 10° C and 25°C as well as dry weather conditions, this calibration is conducted preferably in spring and autumn. The gas cylinders are ordered 8

to 10 weeks prior to the planned calibration procedure. For zero air synthetic air is used. It consists of 20.5 ± 0.5 mol% O_2 in N_2 with a storage stability of 60 months. The gas cylinders for the CO_2 SPAN calibrations have a mixing ratio of 300 ppm CO_2 and 900 ppm CO_2 in synthetic air with a standard measurement uncertainty specified with 2% by the coverage factor k = 2. These cylinders are certified to keep a stable concentration for 36 months of storage. For the SO_2 SPAN calibrations, 100 ppb of SO_2 in synthetic air is used. The standard measurement uncertainty for the used SO_2 SPAN gas cylinders is specified to be 5% by a coverage factor k = 2 and a storage stability of 12 months. For calibrating the NOx module, cylinders of NO in N_2 are used with concentration of 200 ppb ± 5% and 40 ppm ± 2 % each by the coverage factor k = 2 and storage stability of 12 months. The latter gas cylinder is only used in combination with a gas diluting mixing chamber to provide different concentrations of NO.

In the HORIBA system, the calibration gas is fed into the sample line from the gas cylinders with a gas surplus for each gas individually. A covered collecting tube of 1.5 m length with a T-piece for pressure compensation is used. For calibration of the individual modules a sample flow rate of 1 l/min each is needed. For applying calibration gas to the airpointer, a special calibration gas inlet at the maintenance door is used for connecting the covered tube to the system. The calibration gas flow rate is set to 4.5 /min. The T-piece for bypass for pressure compensation of the calibration gas is installed inside the airpointer itself. The calibration gas flows through the T-piece to the sampling filter and further on to all sensors in parallel. Depending on the module, a calibration takes several minutes (CO_2) up to 20 minutes (SO_2) for the sake of stabilization for both HORIBA and airpointer systems.

The calibration procedure for the LICOR CO_2 module consists of three levels: zero, 300 ppm and 900 ppm, while two calibration steps are performed for the SO_2 module (zero and 100 ppb) and for the NOx module (zero and 200 ppb NO). For evaluating the cross sensitivity of the SO_2 module to NO, the calibration procedure consists of a four-step chain (zero, 100 ppb NO, 200 ppb NO and 400 ppb NO). The target concentrations are generated with the gas mixing chamber airQrate (mlu recordum) by diluting the 40 ppm NO SPAN gas with internally generated zero air. With the monitored SO_2 concentration a cross sensitivity factor is calculated, which is used to correct the measured SO_2 concentration in ship plumes according to the measured NO concentration. The zero level of a module is adjusted when deviation > 1.5 ppm for CO_2 and >1.5 ppb for SO_2 and NOx is detected. The SPAN is only adjusted when the measured deviation of the module compared to the gas cylinder certificate is >15 %. For adjusting a instrument, the detected SPAN level has to remain stable for at least 10 min.

2.3.1.2 Cross interferences

Due to the UVF measurement principle, the SO_2 instruments have a known cross-sensitivity to NO. Since the concentration of NO in ship exhaust plumes is known to be very high, the SO_2 measurement must be corrected with the NO measured in parallel. The SO_2 - NO cross sensitivity is measured during each calibration on each instrument and is typically between 0.5 % and 1.5 % of the detected NO concentration.

In test measurements in the laboratory, an anticorrelation to the relative humidity was found for the SO_2 instrument. The measured SO_2 concentration deviated from the expected value by 5 - 20 % at RH 75 % and by 10-25 % at RH 85 % (negative deviation). It is assumed that the SO_2 is partially absorbed in the measuring system at high RH. However, since the observed effect in the laboratory test could not be quantified enough to create a correction function, the mean value of 15 % deviation is considered for the error consideration as described below.

2.3.1.3 Calculation of uncertainty

The FSC is calculated from the measured concentrations of SO_2 and CO_2 inside the plume according to Eq. **1**in section 2.2, which can be also written as:

$$FSC = \frac{[SO_2(ppb)]}{[CO_2(ppm)]} \cdot 0.232 \qquad \text{Eq. 7}$$

The measurement uncertainty with respect to FSC is calculated individually for each single measurement following the Gaussian error propagation by calculating the square root of the sum of the individual errors. Therefore Eq. 7 needs to be partially derivate for SO_2 (ppb) and CO_2 and multiplied with the individual uncertainties. The overall formula for the absolute uncertainty of the FSC, ΔFSC, is therefore:

$$\Delta FSC = \sqrt{\left(\frac{\partial FSC}{\partial(SO_2)} \cdot \Delta[SO_2]\right)^2 + \left(\frac{\partial FSC}{\partial(CO_2)} \cdot \Delta[CO_2]\right)^2 + \left(Uct_{compl\ conversion}\right)^2 + \cdots}$$

$$\overline{\cdots + \left(Uct_{unspecific}\right)^2 + \left(Uct_{calib\ SO_2}\right)^2 + \left(Uct_{calib\ CO_2}\right)^2 + (Uct_{rH})^2} \quad \text{Eq. 8}$$

With the assumptions of:

The uncertainty of the assumption of complete conversation of Sulphur to SO_2 and Carbon to CO_2 during the combustion process is expressed with $Uct_{\ compl\ conversion}$. Various studies showed that there can be an underestimation of the FSC between 1 % to 19 % from assuming complete conversion (Schlager et al., 2006; Agrawal et al., 2008; Moldanová et al., 2009; Balzani Lööv et al., 2014). To get a reasonable representation of this uncertainty, a value of 15 % of the FSC was chosen.

Another uncertainty accounts from unspecific errors resulting from non-perfect distribution of gases within the plume and the deviation from the assumed 87% of Carbon within the fuel. In Eq. **8** this is written as $Uct_{\ unspecific}$. With a sensitivity study a value of 10% of the final FSC was estimated for this uncertainty.

The uncertainty of calibration for SO_2 ($Uct_{calib\ SO_2}$) is 5 % of the resulting FSC (specification gas cylinder).
The uncertainty of calibration for CO_2 ($Uct_{calib\ CO_2}$) is 5 % of the resulting FSC (specification gas cylinder).
The Uncertainty due to the RH effects as described in the above cross interferences section (Uct_{rH}) is 15 % of the resulting FSC.

Eq. 8 can be written as:

$$\Delta FSC = \sqrt{\left(\frac{0.232}{[CO_2]} \cdot \Delta[SO_2]\right)^2 + \left(-\frac{[SO_2] \cdot 0.232}{[CO_2]^2} \cdot \Delta[CO_2]\right)^2 + (0.15 \cdot FSC)^2 + \cdots}$$

$$\overline{\cdots + (0.10 \cdot FSC)^2 + (0.05 \cdot FSC)^2 + (0.05 \cdot FSC)^2 + (0.15 \cdot FSC)^2} \quad \text{Eq. 9}$$

Where $\Delta[SO_2]$ is the total SO_2 instrument uncertainty due to noise, which is calculated as:

$$\Delta[SO_2] = \sqrt{\left(SD_{SO2_1}\right)^2 + \left(SD_{SO2_2}\right)^2 + \cdots + \left(SD_{SO2_n}\right)^2} \quad \text{Eq. 10}$$

$$\Delta[SO_2] = \sqrt{(SD_{SO2})^2 \cdot peak\ width} \quad \text{Eq. 11}$$

with SD_{SO2_1}, SD_{SO2_2},…, SD_{SO2_n} gives the standard deviation of the SO_2 signal at the peak node 1, 2, …, n and $peak\ width$ is the number of peak nodes used for the calculation of the peak area

$\Delta[CO_2]$ is the total CO_2 instrument uncertainty due to noise, which is calculated as:

$$\Delta[CO_2] = \sqrt{\left(SD_{CO2_1}\right)^2 + \left(SD_{CO2_2}\right)^2 + \cdots + \left(SD_{CO2_n}\right)^2} \quad \text{Eq. 12}$$

$$\Delta[CO_2] = \sqrt{(SD_{CO2})^2 \cdot peak\ width} \quad \text{Eq. 13}$$

with SD_{CO2_1}, SD_{CO2_2}, …, SD_{CO2_n} gives the standard deviation of the CO_2 signal at the peak node 1, 2, …, n and $\boldsymbol{peak\ width}$ is again the number of peak nodes used for the calculation of the peak area.

Expanding the first term under the square root by [SO_2], Eq. 9 can be rewritten to:

$$\Delta FSC = \sqrt{\left(FSC \cdot \frac{\Delta[SO_2]}{[SO_2]}\right)^2 + \left(FSC \cdot \frac{\Delta[CO_2]}{[CO_2]}\right)^2 + +(0.15^2 + 0.10^2 + 0.05^2 + 0.05^2 + 0.15^2) \cdot FSC^2 \cdots} \quad \text{Eq. 14}$$

As $\Delta[SO_2]/[SO_2]$ gives the relative uncertainty of the measured SO_2 plume concentration and $\Delta[CO_2]/[CO_2]$gives the relative uncertainty of the measured CO_2 plume concentration, both with respect to the signal standard deviation, Eq. **14** can be generalized to:

$$\Delta FSC = FSC \cdot \sqrt{(Uct_1)^2 + (Uct_2)^2 + \cdots + (Uct_n)^2} \quad \text{Eq. 15}$$

This means that the absolute uncertainty for a measurement can be calculated from the square root of the sum of the squares of the individual relative measurement uncertainties multiplied by the absolute measurement value.
However, as the signal to noise ratio differs for every detected plume, within the automated BSH plume analysis algorithm the measurement uncertainty is calculated individually for each measurement.

2.3.1.4 Data reporting

Only if the plume allocation is distinct and the plume analysis indicates a potential non-compliance, i.e. when the calculated FSC minus uncertainty exceeds 0.11 % S m/m, the software automatically sends an e-mail alert to the responsible local authorities to trigger further actions (e.g. on-board inspection). In case of an alert the FSC results were also automatically transferred to the THETIS-EU database.

When the calculated FSC minus uncertainty exceeds 0.11 % S m/m but no automatic clear allocation of a ship had been taken place, an email is sent to the operators for verification. This is also conducted if the Sulphur peak areas are below the QS criteria (35 ppb for airpointer and 70 ppb for HORIBA). This trigger level was inserted to avoid sending false-positives to inspection authorities.

The operator verifies the specific situation at the measurement site manually and decides ether to send an alert or not according to the procedure as outlined above.

2.3.2 Chalmers standard sniffer system

2.3.2.1 Calibration

The measurement system of the Chalmers standard sniffer is calibrated using certified reference gases for its 3 target species: CO_2, SO_2 and NOx with typical values of ~ 300 ppm, 330 ppb, and 300 ppb respectively. The CO_2 mixture is also used as zero gas calibration for the SO_2 and NOx instruments while the SO_2 mixture is used in the same manner for the CO_2 sensors. The calibration procedure for CO_2 and NOx takes approximately 1 minute, due to the fast response of the corresponding sensors. However, the Fluorescence SO_2 sensor requires a longer calibration interval which typically ranges between 120 to 180 seconds.
The previously described procedure under normal operation is performed every 10 – 20 days. Though it has been observed that the calibration coefficients of the instruments are relatively stable under continuous operation in the long term, still remains unclear the variation of the calibration coefficients on daily bases and their quantitative role

on the measurement uncertainties. During the period when campaign 3 was carried out in Wedel, calibrations were performed on a daily basis, providing a large data set that allowed to observe the daily calibration drifts. Each calibration was performed using a reference gas with concentrations of about 300 and 330 ppb for the SO_2 and NOx instruments respectively. The CO_2 calibration was performed using a reference gas of 430 ppm. The calibration variability was estimated by comparing the difference between two consecutive measured calibration spans.
These observations lead to a better estimation of the calibration uncertainty as shown on Figure 3, Figure 4 and, summarized on Table 2.

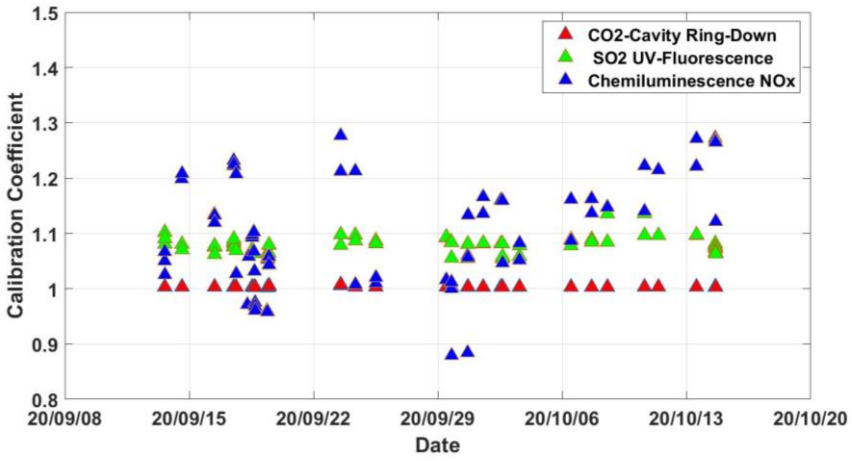

Figure 3. Estimated span calibrations coefficients of CO_2, SO_2 and NOx for the Chalmers standard system throughout the period corresponding to the Wedel field campaign.

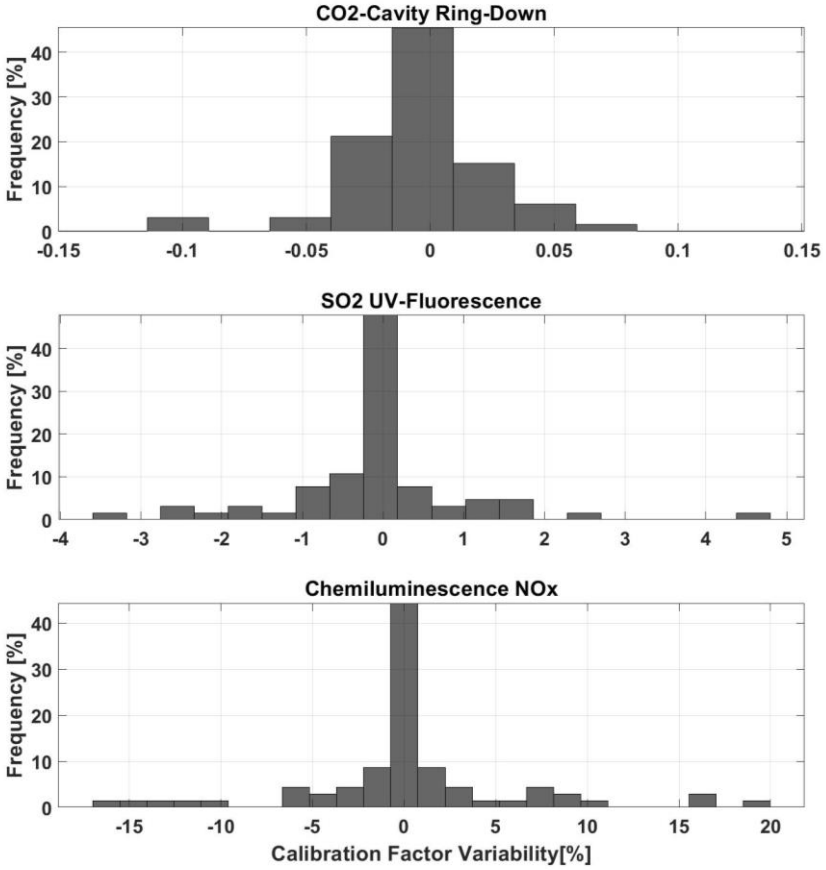

Figure 4. Frequency distribution of the relative variation in calibration factors of CO_2, SO_2 and NOx from Figure 3 for the Chalmers standard system.

Table 2. Standard deviation of relative variation in calibration factor with daily calibration frequency during the C3 Wedel measurement campaign using the Chalmers standard system.

Instrument and species	Relative $\Delta_{cal\text{-}factor}$ (1σ)
CO_2 (Cavity Ring-Down)	0.1 %
NO_x (Chemiluminescent)	6.5 %
SO_2 (Fluorescence)	1.3 %

2.3.2.2 Cross interference

The SO_2 sensors based on fluorescence are particularly sensitive to hydrocarbons and NO. While the hydrocarbon interference is physically removed by using a hydrocarbon kicker; the NO interference can only be estimated and corrected from simultaneous NO measurements or NO_x measurements assuming a certain NO/NO_x ratio of around 70%. Experimentally, it has been determined that this interference counts for 1.2±0.1% and 0.98±0.45% for NO and NO_x respectively. Under the assumption of a linear interference effect, the SO_2 measurements are corrected by just subtracting the estimated NO/NO_x interference.

In situations when a faster response time of fluorescence SO_2 measurements is required, like fast airborne measurements, the kicker is normally removed. However, it comes with the price of the possible cross interference in the SO_2 readings due to the presence of organic species in the plume sample. In general, this artifact has a very particular signature characterized by long tails in the plume measurements; hence, it can be identified and corrected. Ignoring this artifact can lead to an overestimation of up to 0.2 % S m/m in the calculated FSC value.

2.3.2.3 Quality Criteria

In the data evaluation the quality of the measurements is expressed through a quality flag that can alternate between the following levels: HIGH, MEDIUM, and POOR. The flagging criteria, which considers several measurement parameters, is summarized in Table 3 and, is empirically based on long term measurements carried out at the fixed station at the Great Belt.

The quality flag is a combination of measured parameters such as CO_2 peak signal and the surrounding environmental conditions when it has been acknowledged that the measurements are more certain. One important consideration here is the comparison of CO_2 in the ship plume against the variation of the ambient background CO_2, which comprises both variations of the background (upwind fixed source like a city) and the noise of the instrument. The quality level may also shrink if different hardware warning flags are raised while the instruments are operating. These flags are mostly associated with issues related to high/low temperature, low voltages, flow interruptions, etc.

In general, automatic data retrieval based on the previously described criteria, performed satisfactorily for high and medium quality measurements and filters out measurements with poor quality, which are considered uncertain.

Table 3. Example of the quality criteria for Chalmers applied for fixed measurements using standard system at the Great Belt site.

Criteria	Comment	High	Medium	Poor
Normal operation	Warning flags for the hardware not set, such as high/low temperature, low voltages etc	Required	Required	Depends
ΔCO_{2Max} in plume	Peak height	>3 ppm	>2 ppm	0.5 ppm
$\Delta CO_{2integrated}$	Integration of plume	>50 ppm.s	>25 ppm.s	3> ppm.s
ΔtCO_2 in plume	Time duration in plume	<100 s	<150 s	<240 s
Wind direction	Wind relative to ship movement	± 30º	± (30-60º)	± (30-60º)
Wind speed		3-8 m/s	2-3 m/s or 8-10 m/s	1-2 m/s or 10-12 m/s
No of ships with overlapping plumes		1	1	1
FSC	Filtering out low values	>-0.2	>-0.2	>-0.2

2.3.2.4 Calculation of uncertainty

This calculation is based on the signal properties of the concentration of each of the plumes, their cross interferences and calibration uncertainties. The total uncertainty of the FSC calculation, involving the ratio between SO_2 and CO_2 in the measured peak, Utot is described by Eq. 16.

$$U_{tot} = k\sqrt{\left(U_{SNR}(X_1)^2 + U_{BL}(X_1)^2 + U_{cal}(X_1)^2 + U_{gas}(X_1)^2 + U_{cross}(X_1)^2 + +U_{SNR}(X_2) \dots\right)} \quad \text{Eq. 16}$$

X Corresponds to the different species ($\mathbf{X_1}$=$\mathbf{SO_2}$, $\mathbf{X_2}$=$\mathbf{CO_2}$), etc.

U_{tot} is total relative uncertainty.

k is coverage factor, depends on N degrees of freedom typically ranging 1.96 - 3. It is the factor which is used to expand the standard uncertainty to CI 95 %, see section 5.3.1.

U_{SNR} corresponds to the inverse of the signal to noise ratio, calculated as the standard deviation of the baseline noise (ΔS) normalized with square root of the number of samples *N* (standard error of the mean) and the average peak height of volume mixing ratio (ppb | ppm) as shown on Figure 5 and described by Eq. 17.

$$U_{SNR}(X) = \frac{\left(\frac{STD(X)_{background}}{\sqrt{N}}\right)}{VMR_{avg}(X)} \quad \text{Eq. 17}$$

VMR_{avg} corresponds to average peak height of the measured volume mixing ratio as described below.

$$\boldsymbol{VMR_{avg}(X)} = \frac{\int_0^N \boldsymbol{VMR(X)}}{\boldsymbol{N}} \quad \text{Eq. 18}$$

U_{BL} corresponds to the baseline uncertainty calculated as the difference in baseline on both sides of the peak ($\Delta_{Baseline}$) normalized with the average peak height and $\sqrt{3}$. Here the latter corresponds to the 95 % CI coverage factor for a boxcar probability distribution (assuming the baseline value to equal probability for any value within the given range), as shown on Figure 5, and described by Eq. 19.

$$U_{BL} = \frac{\Delta_{Baseline}(X)}{2\sqrt{3}\cdot VMR_{avg}(X)} \quad \text{Eq. 19}$$

U_{cal} is the variability of calibration factor between calibrations as previously described on section 2.3.2.1.

$$U_{cal} = STD\left(\frac{\Delta_{cal\,factor}}{cal\,factor}\right) \quad \textit{Eq. 20}$$

U_{gas} *(X)*, corresponds to calibration gas uncertainty as specified by gas supplier, typically ranging 2-10 %

U_{cross} *(X)*, corresponds to the uncertainty in cross sensitivity of the instrument to other gases than the actual key as described in section 2.3.3.2.

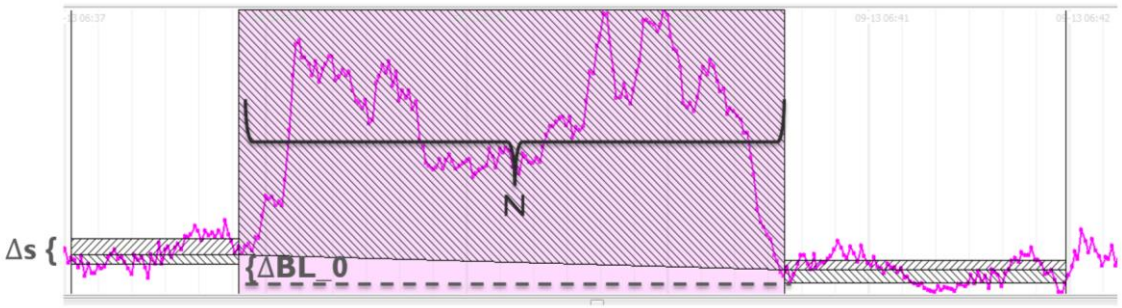

Figure 5. Measured plume and uncertainty parameters

2.3.2.5 Data reporting

The Chalmers Standard system continuously acquires data and automatically performs emission calculations including the FSC and other emission factors. After the FSC is calculated, the results are reported through the following protocol:

- The likely FSC emission vessel is identified by combining AIS signals and wind data.
- The system assigns data quality flags of the FSC measurement according to the criteria previously described.
- If the quality criteria classify the measurement as a valid one, the result and the AIS data is inserted in a database located at Chalmers University of Technology (Figure 6). The system also inserts in the database an animated figure that recreates the vessel passage and its corresponding measured signals (Figure 7).
- If the FSC is > 0.1% the system also generates alert emails to the authorized users.
- After an alert is generated, the animation is manually inspected to confirm the FSC alert or filter out possible false non-complaint vessels. Consequently, the measurement can either be removed from the database or reported to the corresponding authorities.
- A pdf with the animation with measurements results is generated and if the quality control is approved then this is submitted to the Thetis-EU database by the port state authority.

Date	type	Value	Quality tag	Quality control	Platform	Platform type	Ship name	Ship type	IMO	MMSI	Ship coordina
03/09/2016 - 10:56	SC	0.00	Poor	Automatic	Älvsborg	Stationary	HANS	Other Type	8996956	230034540	11.848055,57
03/09/2016 - 10:49	SC	0.06	High	Automatic	Älvsborg	Stationary	ICE STAR	Cargo	9142631	244264000	11.846465,57
03/09/2016 - 10:25	SC	0.02	Medium	Automatic	Älvsborg	Stationary	GEORGE	Other Type	8634077	230034510	11.846706,57
03/09/2016 - 09:44	SC	0.04	Medium	Automatic	Älvsborg	Stationary	STENA DANICA	Passenger	7907245	265177000	11.825803,57
03/09/2016 - 09:18	SC	0.02	Medium	Automatic	Älvsborg	Stationary	ASTINA	Tanker » Hazardous category A	9320063	266220000	11.850048,57
03/09/2016 - 08:56	SC	-0.15	Poor	Automatic	Älvsborg	Stationary	DELFIN	Tanker	0	265505100	11.853978,57
03/09/2016 - 08:44	SC	0.02	Poor	Automatic	Älvsborg	Stationary	STENA SCANDINAVICA	Passenger » No additional information	9235517	266343000	11.860925,57
03/09/2016 - 08:40	SC	-0.04	High	Automatic	Älvsborg	Stationary	STENA SCANDINAVICA	Passenger » No additional information	9235517	266343000	11.843681,57
03/09/2016 - 08:17	SC	0.07	High	Automatic	Älvsborg	Stationary	THUN GLOBE	Tanker	9229051	245573000	11.834666,57
03/09/2016 - 08:08	SC	-0.02	Medium	Automatic	Älvsborg	Stationary	HANS	Other Type	8996956	230034540	11.837226,57
03/09/2016 - 07:51	SC	0.01	Medium	Automatic	Älvsborg	Stationary	LEXUS	Tanker » No additional information	9310317	219157000	11.856116,57

Figure 6. Example of a database query corresponding to FSC measurements at the Älvsborg site in Göteborg.

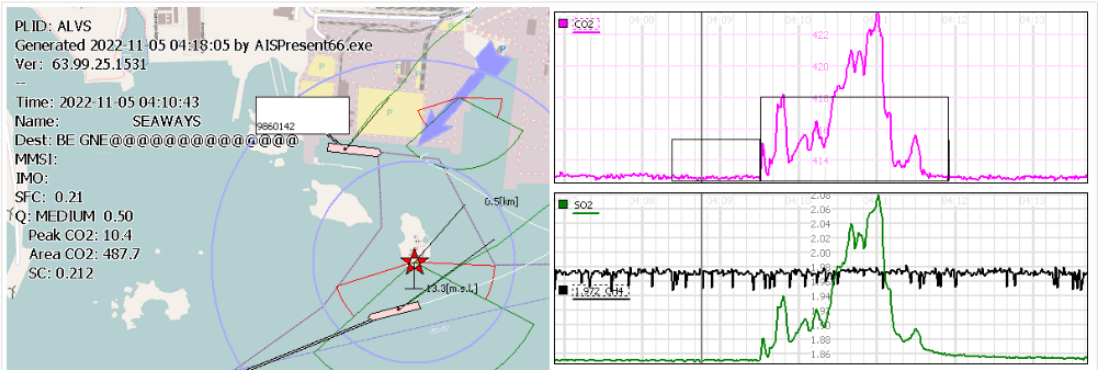

Figure 7. Example of an animation generated by the Chalmers Systems corresponding to a non-compliance vessel measured at the Älvsborg site in Göteborg.

2.3.3 TNO standard sniffer system

2.3.3.1 Calibration

Calibration of the monitors occurs once every month. All three monitors (SO_2, NO and CO_2) are calibrated using zero and a span in dry air. The span cylinders are calibration cylinder with given concentrations from the manufacturer. Typical concentrations used are around 100-200 ppb for SO_2, 400 ppb for NO and 450 ppm for CO_2. To make sure that there is an overpressure of the calibration cylinder, and no surrounding air is sampled after the pressure regulating valve the inlet is split into two, one to the monitor and the other to a rotameter (which measures the flow). The valve is adjusted in such a way that the rotameter measures a certain flow. After this is set, the inlet is flushed for at least 1 minute. During the calibration procedure a program displays the concentration of all three monitors simultaneously. The program shows the current, 300 s moving average (i.e. 5 minutes) and the 10-percentiel. This allows the operator to clearly see when the monitors register a stable concentration and set this as zero or span-concentration in the monitor. First, the monitor the zero is determined with synthetic gas. After that, the span of the calibration gas is offered to the monitors. The calibration constants are adjusted (when the values are more than 1 ppb or for CO_2 1 ppm off from the previous values) and logged. Alongside the calibration general maintenance is carried out. The 5 µm Teflon filter is replaced and flows are checked.

2.3.3.2 Cross interferences

A known cross interference is of the SO_2 monitor to NO. This cross sensitivity was investigated in the lab. It was found that a correction factor of 0.008 (0.8%) was appropriate. Thus, from the integrated SO_2 plume a factor of 0.008 of the integrated NO plume is subtracted.

2.3.3.3 Calculation of uncertainties

The signal to noise ratio is calculated by taking the standard deviation outside the peak (1 minute before and after the peak) and dividing this by the height of the peak. To calculate the relative error of the FSC the signal to noise ratio of the SO_2 and CO_2 signal are quadratic added, which gives the formula:

$$Error_{FSC} = \sqrt{\left(\frac{std(SO_2)}{peak(SO_2)}\right)^2 + \left(\frac{std(CO_2)}{peak(CO_2)}\right)^2} \qquad \text{Eq. 21}$$

where $Error_{FSC}$ is the relative error of the FSC, std(SO_2) and $std(CO_2)$ the standard deviations of SO_2 and CO_2 outside the peak respectively, and peak(SO_2) and peak(CO_2) the peak height of SO_2 and CO_2 respectively (maximum minus the background level). By multiplying $Error_{FSC}$ with the FSC the absolute error in FSC is obtained.

2.3.3.4 Data reporting

Data reporting occurs near real-time to the Human Environment and Transport Inspectorate (in Dutch Inspectie Leefomgeving en Transport, ILT) on a password protected website. Hourly an automated system calculates the FSC

of ship plumes and allocates it to ships that passed by the measurement location (with help of the AIS signal and the measured wind conditions). This allows ILT to take targeted fuel sample of ships.
For ILT to quickly see which ships are possibly non-compliant the FSC values are divided in four classes and coloured accordingly. These classes are:

- FSC≤0.10% → Highly likely compliant
- 0.10%<FSC≤0.12% → Possibly non-compliant
- 0.12%<FSC≤0.15% → Probably non-compliant
- FSC≥0.15% → Highly likely non-compliant

In the automated systems a quality scoring of the reliability of the FSC estimate is given in poor, fair and good. Only the FSC estimates with a quality scoring of good are reported to ILT. The following criteria are to be met to get the quality score good:

- The peak height in gas concentration should meet or surpass certain threshold, which are:
  - SO_2 → 2 ppb
  - CO_2 → 3 ppm
  - NO → 10 ppb
- The peaks must be at least 4 times above the standard deviation in the signals without a peak (i.e. the noise level).

2.4 Implementation of high-sensitive laser instrument for ship emission monitoring

2.4.1 Calibration

The laser measurement system of Chalmers is calibrated using certified reference gases for its 2 target species: CO_2 and SO_2 with typical values of ~ 300 ppm, and 300 ppb respectively. The CO_2 calibration mixture is also used as zero gas calibration for the SO_2 while the SO_2 mixture is used in the corresponding manner for the CO_2. Due to the fast response of the instrument to both CO_2 and SO_2, the calibration procedure typically takes approximately 1 minute. The experiment set up at Wedel, which the same procedures described in section 2.3.2.1, also allowed to characterize the calibration daily drifts as shown in Figure 8 and Figure 9.

Table 4. Standard deviation of relative variation in calibration factor with daily calibration frequency during the C3 Wedel measurement campaign using the high-sensitive laser system.

Species	Relative $\Delta_{cal\text{-}factor}$ (1σ)
CO_2	1.9 %
SO_2	0.6 %

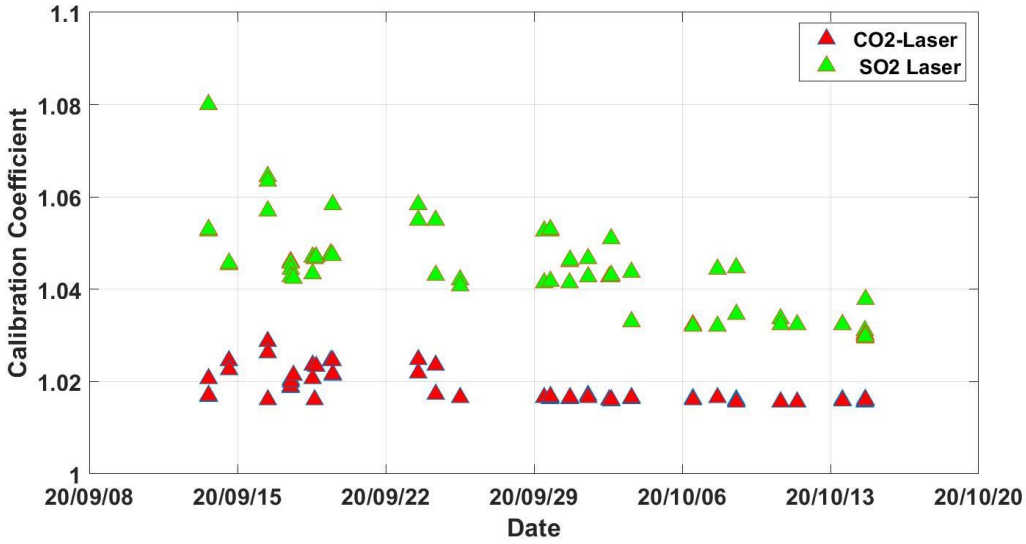

Figure 8. Estimated span calibrations coefficients of CO_2 and SO_2 for the Chalmers Laser system throughout the period corresponding to the Wedel field campaign.

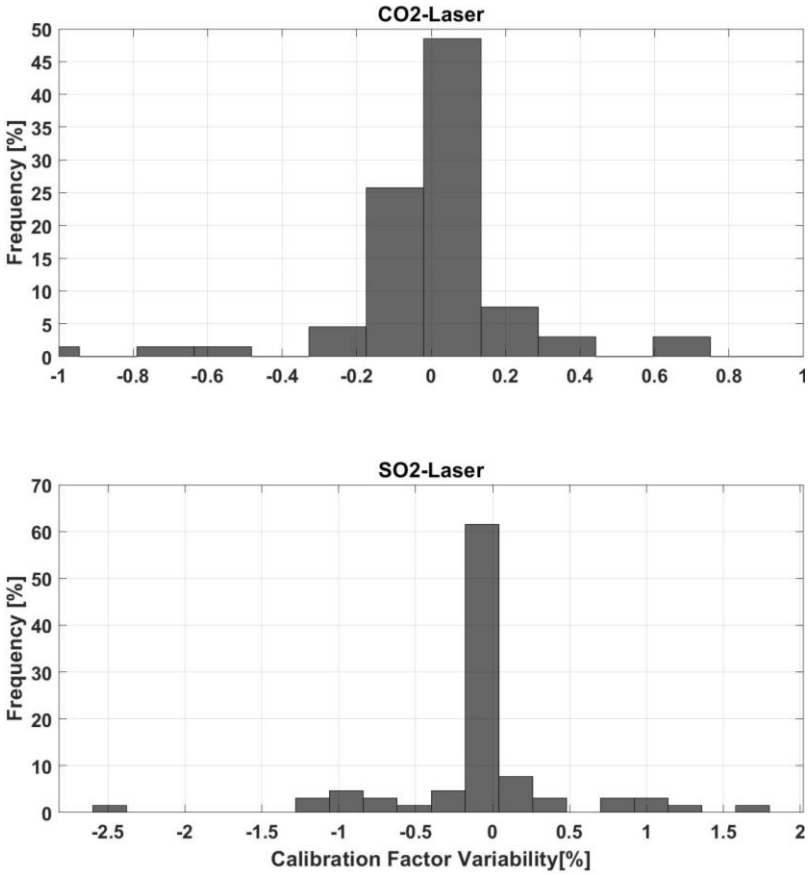

Figure 9. Frequency distribution of the relative variation in calibration factors of CO_2, and SO_2 from Figure 8 for the Chalmers Laser system.

2.4.2 Cross interference

The SO_2 and CO_2 measurements for the laser sensor does not have cross-interference for neither NO_2 nor CO_2.

2.4.3 Quality criteria

The quality criteria for the laser system are based on similar criteria procedures as the Chalmers standard system described in section 2.3.2.3. However, this system is new and has only been as a research system and therefore all calculations have been redone afterwards in a post analysis software.

2.4.4 Calculation of uncertainty

This calculation is based on the same principles and notation as the Chalmers standard system in section 2.3.2. Since the laser system does not have any cross interference, the amount of uncertainty parameters is slightly reduced.

$$U_{tot} = k\sqrt{\left(U_{SNR}(X_1)^2 + U_{BL}(X_1)^2 + U_{cal}(X_1)^2 + U_{gas}(X_1)^2 + U_{SNR}(X_2) \dots\right)} \quad \text{Eq. 22}$$

2.4.4.1 Data reporting

The Laser system is a variant of the Standard Chalmers system; and follows the same protocol described in section 2.3.2.5. However, the calculations of the FSC assume that there are no cross interferences in the SO_2 signal.

2.5 Implementation of mini sniffer sensors for ship emission monitoring from drones

Mini sniffer sensors, in the context of the applications presented in this report denote compact and light-weight sensors compared to the analysers used for sniffer measurements described in the previous sections. The applied sensors are electrochemical sensors (EC) and/or infrared sensors (NDIR) depending on the gas to be measured. Generally, these sensors are unsuitable for low concentrations observed in highly diluted plumes as they require higher gas concentrations (typically 100-1000 times higher) and longer response times to reach the target value. However, when applied on airborne platforms like drones or helicopters, these disadvantages can be overcome through plume navigation, i.e. the ability via piloting to monitor in the proximity of the stack outlet and the time spent in each plume.

Typical measuring ranges observed in the plumes from drones are as follows (Explicit, 2018):

- CO_2: 200 – 600 ppm over background
- SO_2: 0.5 – 1 ppm
- NO_2: 0.5 – 1 ppm
- NO: 2 – 10 ppm

As for the response time, in the case of EC sensors the gas needs to diffuse through a membrane which means it can take up to several tens of seconds to reach a steady signal.

Additionally, mini sniffers typically have cross-sensitivities to other gaseous constituents, temperature, pressure and/or relative humidity requiring the application of additional corrective measures to ensure the accuracy of the final results. This is typically handled post sampling and will depend on the exact design of the measuring system. Because EC sensors rely on membrane technology, their lifespan is also limited, typically up to 2 years, and dependent on exposure. In this case, it's the reactive chemicals in the membranes that are used requiring the sensors to eventually be replaced. Finally, EC sensors are observed to have larger quality variability from the manufacturing side. All of the above put high demands on the quality assurance procedures.

The companies Explicit ApS (Denmark) and Aeromon Oy (Finland) as well as the Chalmers Technology developed their own small-sized sniffer systems which are presented in this section.

2.5.1 Explicit

2.5.1.1 Flight approach

Explicit has developed a patented method[2] for how to approach and sample the vessel exhaust gases using drones. This method relies on a combination of weather input, real-time sensor feedback, and vessel AIS data to identify the optimal sampling spot in the plume, also referred to as the 'sweet spot'.

[2] EP 3 100 022 B1, A Method and an Unmanned Aerial Vehicle for Determining Emissions from a Vessel.

The sweet spot is most often located along the centre line of the plume in reasonable distance from the ship stack. The exact distance depends on the size and operation of the ship. The smaller the vessel and/or the less engine output, the closer the distance. The typical position of the sweet spot ranges from 25-100 meters, 40-60 meters above sea level. The relatively close proximity to the emission source also means the risk of contamination from other plumes is in practice non-existent.

The flight approach is illustrated below:

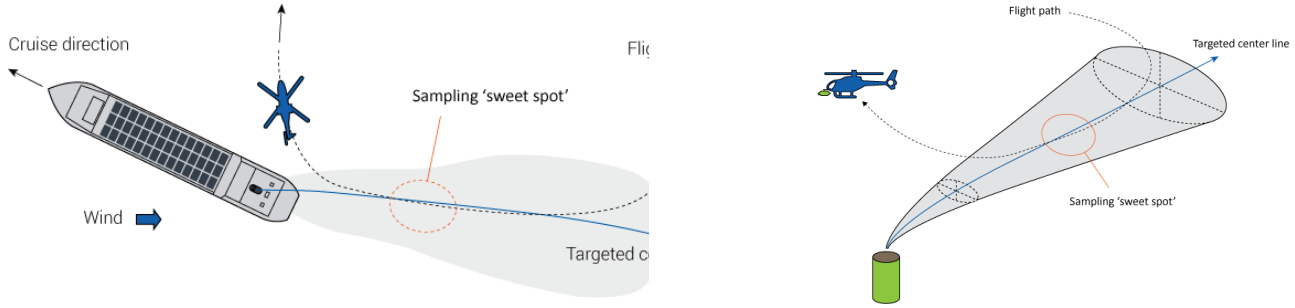

Figure 10. The flight approach used by Explicit.

To locate the likely position of the plume within the area of operation, the pilot uses a series of proprietary software tools, developed by Explicit, starting with a plume simulation model based on the vessel course AIS data and online data on local wind conditions. The model forecasts the rough position of the plume vis-à-vis the ship position and course.
Once in the vicinity of the plume, the pilot switches to Smart Flight mode, to locate the sweet spot inside the plume. Smart Flight allows the pilot to navigate intelligently using the live feedback from the sensors to locate the position with the optimal gas concentrations inside the plume. In order words, the pilot 'sniffs' his/her way to the optimal position, reducing the risk of failed or weak measuring attempts dramatically.
The feedback from the sensors is also what is subsequently used to evaluate the quality of each measurement, i.e. the pilot's ability to optimize the sampling position in the plume. See later sections on quality scoring protocols.

2.5.1.2 Calibration

To ensure the functionality of the sensors, Explicit has established the following test and calibration protocols, which were also followed during the SCIPPER campaigns:
Each Explicit Mini Sniffer Unit (EMSU) is performance-tested and calibrated before deployment. Test and calibration are done by FORCE Technology in accordance with ISO-standard EN ISO 6145-1. Each calibrated unit is subsequently issued with a certificate of calibration. Units that do not meet the performance standard inherent in the system (i.e. perform comfortably within the established uncertainties) are rejected. Because of EC sensor signal decay over time, each EMSU must be replaced at appropriate intervals, or as a minimum once a year, to ensure functionality. Through drift and performance tests before and after campaign operations, the durability of a unit - i.e. the time it can sustain its calibration without material drift - has been conservatively assessed to be minimum approx. 100 hours of powered operations. Explicit has built-in system logs as part of its QA procedures to keep track of usage and replacement frequencies.

2.5.1.3 Characterization of uncertainties and cross-sensitivities

Explicit uses a pre-deployment characterization of the uncertainties of their sensors according to ISO 6145-1 and ISO/IEC Guide 98-3:2008 at FORCE Technology as an external testing facility.

To establish the uncertainty budget and cross-sensitivities, the assembled EMSUs have been tested under representative environmental conditions which resemble those present under deployment in the field. As such not only the abundances of the measured gases are accounted for but also variations in temperature, pressure, and humidity.
For the error calculation, each relevant component has been considered and combined according to Gaussian error propagation. The entire sensor system is exposed repeatedly to short pulses of known mixtures of the gases of interest which simulate the exposure during sampling of a ship plume. The chosen range of concentrations of SO_2, CO_2, NO, NO_2, and H2O are reflecting values measured during typical sampling conditions in the centre of a plume as distances appliable to the method.

For the characterization of the uncertainty regarding sulphur emissions, the mixing ratio of SO_2 in the gas blend was varied so that it resembles various fuel sulphur contents. The random and systematic components of the measurement uncertainty of the analysed FSC values have then been assessed from repeated measurements. Also, the cross-sensitivity of sensors to other gases and the related uncertainty is assessed in this manner, namely that of the applied electrochemical sensor to NO_2. For each measurement, the relative standard uncertainty is reported depending only on the level of the sulphur content found in the plume and no other situational variables. The validity of the found uncertainties is assured during field operation by defined quality criteria matching the conditions under testing, i.e. only at certain minimum concentration levels and exposure times equal to a minimum quality score, do the uncertainties apply. The fixed uncertainty budget approach also means, the expressed error represents a conservative estimate of the actual uncertainties.
For the assessment of the uncertainty of fuel-mass specific emission factors of NO and NO_2, the sensor system was tested in a similar way, through repeated exposure to simulated exhaust gas mixtures at typical ratios as can be expected from marine engines. From this, an overall relative uncertainty was derived.
Where payload capacity permits, the measurement system is expanded to two identical EMSUs measuring simultaneously and independently. This approach, known as Multiple Parallel Sampling, reduces the uncertainty by the magnitude of $\sqrt{2}$. It also ensures redundancy on the system.

2.5.1.4 Data reporting

All data collected during missions is processed through the Sensor Operator Ground Control Station (processing unit in the field), where data from sensors, GPS and AIS tracks are merged using UTC time stamps to link a specific measurement to a specific ship in time and space. Data is subsequently relayed to the Explicit E-lab (cloud application environment) for the final emissions analysis before the collected and processed data is combined to form a final emissions report on a ship observation. Before issuing a final emissions report on a vessel, the results are subject to a manual quality control by the E-Lab, including verifying sensor performance and emission outputs, investigating low quality scores and rejecting measurements that do not meet the minimum quality score criteria.
Final emissions reports on all observed vessels are made available to clients via te Explicit Airborne Emissions Monitoring System (AEMS), a web-based reporting module that allows users access to the reported data from any device. The AEMS also includes map functions and an array of different statistical and search tools to enable overall analysis of compliance levels, operations, and historic records across larger datasets. Various APIs have also been developed to enable data exchange with other external services.

2.5.2 Aeromon

2.5.2.1 Aeromon BH-12 mini sniffer

The Aeromon BH-12 emission measuring device (Figure 11) is a patented, modular and portable piece of equipment intended for simultaneous detecting, measuring, and mapping of multiple airborne gaseous compounds, particulate matter, and noise. The device communicates its location and measured values in real-time to Aeromon Cloud Service (ACS) which is used to store, analyse, and visualize the results. BH-12 supports various different datalink options and is normally used either with IP radio or direct mobile LTE connection.

Figure 11. Aeromon BH-12 mini sniffer

2.5.2.2 Calibration procedure

Aeromon performs factory calibration with certified gases when sensor modules are assembled. After this the sensor modules are stored in controlled conditions. When sensor modules are used in field ops the calibration level is checked daily (calibration before starting and level check after measurement) with certified calibration gases and known cross interferences (including humidity) to maintain result traceability. This field calibration or calibration level check only takes a few minutes to complete, and the procedure is guided step-by-step for users by BH-12 software. When sensor response declines under pre-determined threshold levels the sensor modules are changed. Also, as sensor response to target gases and their cross interferences changes with T, RH and sensor ageing, this daily quality check procedure provides an essential quality control for small sensor, sample line and full system performance. The calibration gas concentration tolerances are included as a part of the uncertainty calculation.

Figure 12. Daily calibration level check is always performed through the sample line in use.

2.5.2.3 Error sources accounted for reporting

Aeromon has a minimum time window definition based on SO_2 sensor response time (slower than CO_2, NO, NO_2). When minimum time window requirement is exceeded, we calculate the FSC STD point-by-point within the FSC integration time window. This FSC STD is a part of the uncertainty calculation. This analysis step takes into consideration both the S/N ratio of FSC sensors and pilot performance during sampling time window.

$$\boldsymbol{U_{FSC}}(\%) = \sqrt{(\boldsymbol{U_{cal}})^2 + (\boldsymbol{U_{int}})^2 + (\boldsymbol{U_{FSCrel}})^2} \qquad \text{Eq. 23}$$

U_{cal} = calibration gas tolerance
U_{int} = intrinsic errors incl. sensor and device uncertainties from component manufacturer declarations and laboratory testing
U_{FSCrel} = relative error of FSC during the plume visit. The sampling or drone pilot performance and sensor S/N ratio are the main sources for this error. In cases where sampling concentrations vary notably during plume visit this error increases as different sensors react to altering concentrations at different characteristic speed, thus, giving rise to deviation in point-by-point calculated FSC value. In plumes with very low SO_2 concentration in relation to CO_2 the S/N ratio of SO_2 signal may become the dominant source of significant increase of this error.

The ambient baseline steps are not considered as a concern, because in general the ambient level changes of CO_2, NO_x and SO_2 are very small in comparison to our typical sample concentrations (CO_2 > 150 ppm above atmospheric background, SO_2 and NO_2 in range of > 0.1 ppm and NO > 0.5 ppm above ambient levels. However, in the FSC analysis tool in ACS the user can adjust the background level (level is selected for both before and after the plume sampling) for each gas parameter individually if needed.
In many critical applications, such as vessel compliancy monitoring, we utilize dual sensor modules such, that each gas parameters (CO_2, SO_2, NO, NO_2) are measured with two independent sensor module unit. This improves redundancy and decreases the uncertainty through S/N ratio.
Finnish Meteorological Institute reference laboratory (accredited according to ISO 17025) has performed an analysis of the expanded uncertainty of Aeromon system according to JCGM 100:2008 and the device performance characteristics have been defined following the draft technical document of CEN TC 264/WG 42: Air Quality – Performance evaluation of sensors for the determination of concentrations of gaseous pollutants and particulate matter in ambient air.

The results can be paired with additional quality indicator when needed, and, if this is needed, we follow a 4-step result quality classification:

1. **High** – Both CO_2 concentration is well above the higher threshold level and sampling is stable and met the system t90 limit.
2. **Moderate** – CO_2 concentration is above the lower threshold level and sampling is stable and met the system t90 limit.
3. **Low** – CO_2 concentration is above the lower threshold level but sampling is unstable or system T90 limit was not met.
4. **Failed** – Plume was entered but either CO_2 concentration threshold wasn't exceeded or sampling was extremely unstable or system T90 limit was not met.

The exact details of the concentration thresholds, sampling stability and system T90 limit depend on exact sensor and sample line component performances.

2.5.2.4 Data reporting

All the measurement data is visualized in real-time, and the analysed results are reported in a 2-phase process shown in Figure 13. First, automated analysis results are reported immediately after the drone has managed to stay in plume for minimum time to collect data for a reliable result. In the second phase, the automated analysis results are quality-controlled within minutes by an expert remote team and the quality-controlled results are delivered thereafter. Supported reporting formats are e.g., .pdf, .json and .html.

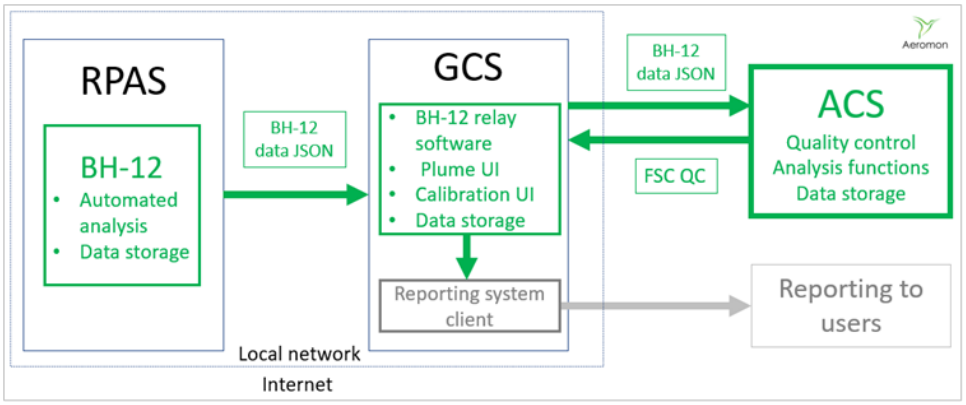

Figure 13. Data flow schematics example for reporting.

Typical individual result reporting includes the result and its uncertainty paired with measurement time (UTC), location (lat., lon., alt.), device and sensor module identification and measurement target identification information, such as IMO number, MMSI number, vessel name, location, and other standard AIS data fields.

2.5.2.5 Plume approach

The ship plume is directed towards the vector sum of true wind and created wind, as shown in Figure 14. Plume altitude depends heavily on the stack height and geometry in relation to (apparent) wind speed and direction. After the plume direction and altitude have been identified the indicators for plume entrance are optical observation of onboard cameras (e.g., visual, thermal), drone experiences turbulence and Aeromon Plume UI indicates visually to pilot crew that plume entrance has happened. The Plume UI uses the threshold concentration levels of plume compounds and indicates also the pilot crew on real-time concentration level when minimum time needed inside the plume has passed. The aim of sampling is not to fly through the plume, but to find a position with high enough concentration (indicated by Plume UI) and stay there for sufficient time, which depends on the slowest sensor T90 response time equipped onboard BH-12.

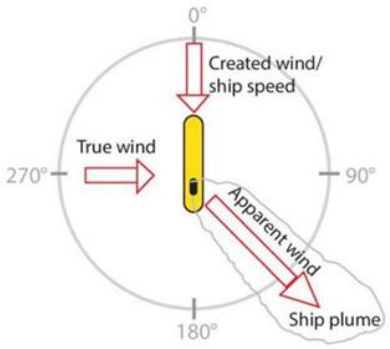

Figure 14. Ship plume direction (Berg et. al. 2012)

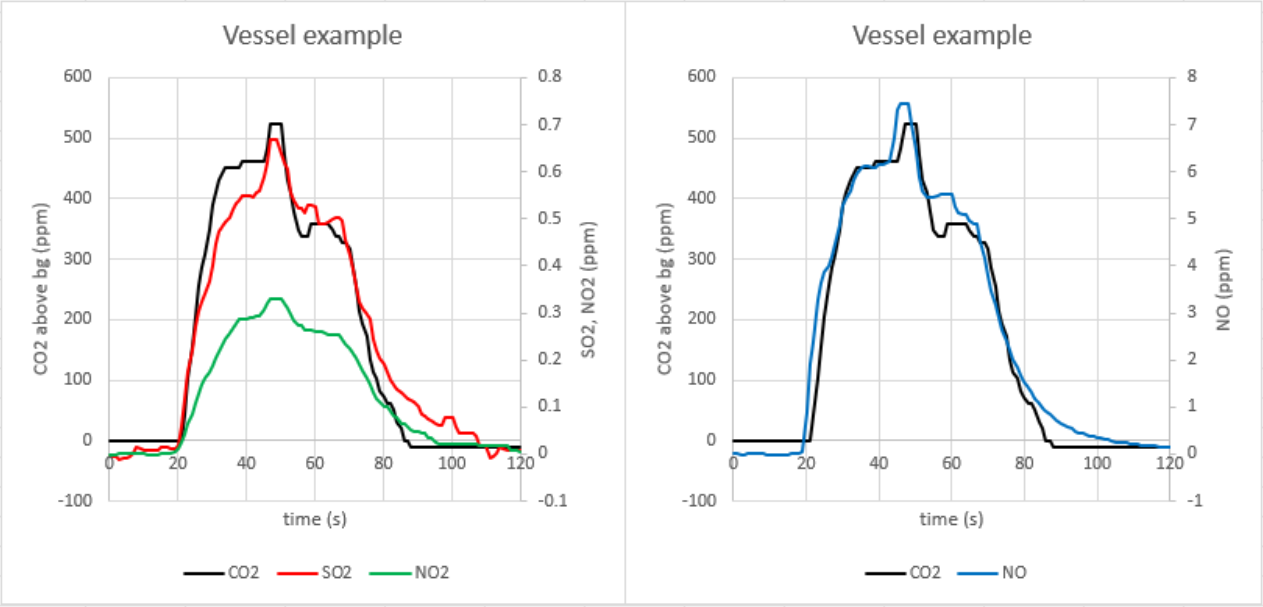

Figure 15. Vessel measurement example showing CO_2 (above atmospheric background), SO_2, NO_2 and NO signals.

Figure 15 shows an actual vessel measurement with Aeromon BH-12. The corresponding results for this vessel were FSC% = 0.30 % (uncertainty 19 % of reading) and $EFNO_x$ = 49.15 (uncertainty 24 % of reading).

2.5.3 Chalmers

2.5.3.1 Chalmers mini sniffer

A series of airborne mini sniffer prototypes have been developed by Chalmers University, and the latest version carried out ship plumes measurements during some of the SCIPPER campaigns. The sniffer was mounted over a hexa-copter in Y-shape configuration, developed by Sky-Eye Innovations in Sweden (Figure 16).

Figure 16. Mini Sniffer developed by Chalmers University

The sniffer box consists mainly of a network of electrochemical sensors (SO_2, NO_2, NO, O_3), an infrared CO_2 sensor and a photo ionization detector (PID) for volatile organic compounds (Figure 17). The gas sensor readings are complemented with other environmental parameters such as temperature, relative humidity, pressure and wind speed/direction. The combined measurement data frame plus GPS positions are transmitted via a 2.4 GHZ link, which is independent of the local mobile data network.

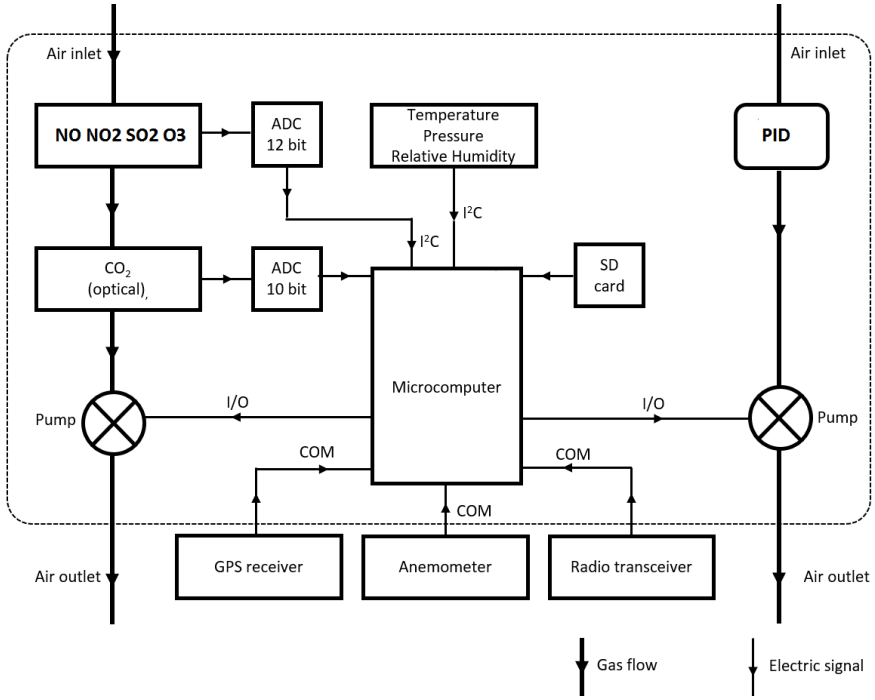

Figure 17. Schematics of the Chalmers drone sniffer box.

2.5.3.2 Calibration procedure

The mini sniffer calibration procedure can be divided into two main procedures: laboratory characterization and in-situ calibration.

The laboratory characterization is a cross test between the electrochemical sensors and more reliable reference sensors at different and controlled gas mixtures (Figure 18). Besides the calibration of the electrochemical sensors, this procedure allows a better understanding and quantification of the cross-interferences, which are not clearly specified in the default factory setting.

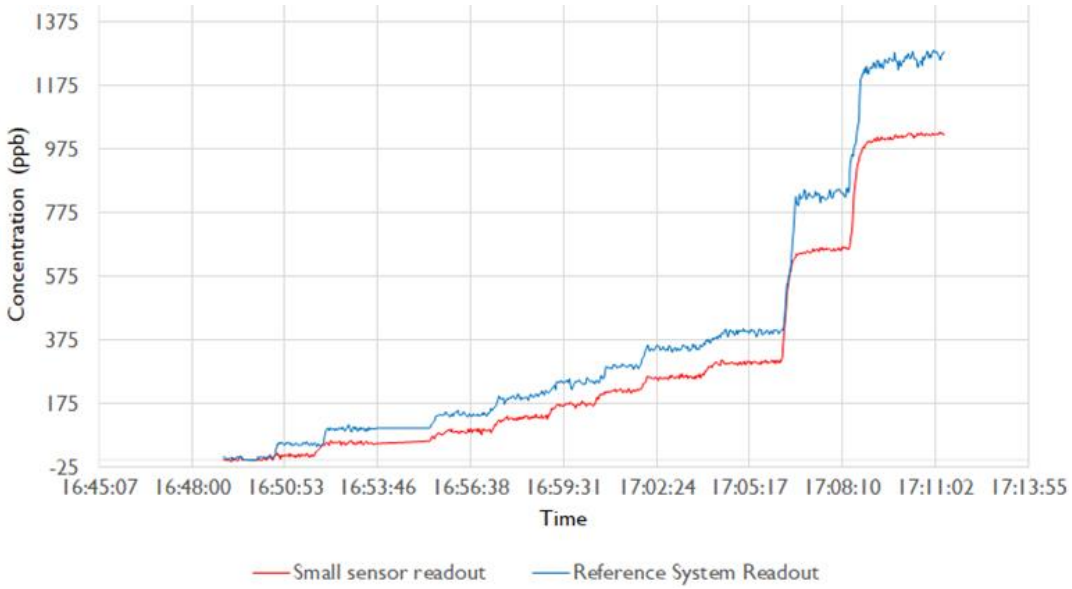

Figure 18. Difference between SO_2 readings from electrochemical sensor as compared to the standard sniffer system.

One of the cross-interferences that plays a major role is the negative interference of the SO_2 readings in the presence of NO_2, which is roughly 100 % negative according to the technical specifications of the sensors. A very coarse correction can be achieved by sampling adding the NO_2 signal as shown in Eq, 24.

$$SO_{2corrected} = [SO_2] + [NO_2] \quad \text{Eq. 24}$$

However, more detailed laboratory characterizations as the example shown on Figure 19 indicate that the NO_2 cross interference can be 25 % - 50 % larger than what is stated on the technical specifications.
In situ calibration with certified gas mixtures allows to either re-calibrate, or control any calibration drifts from the laboratory calibration.

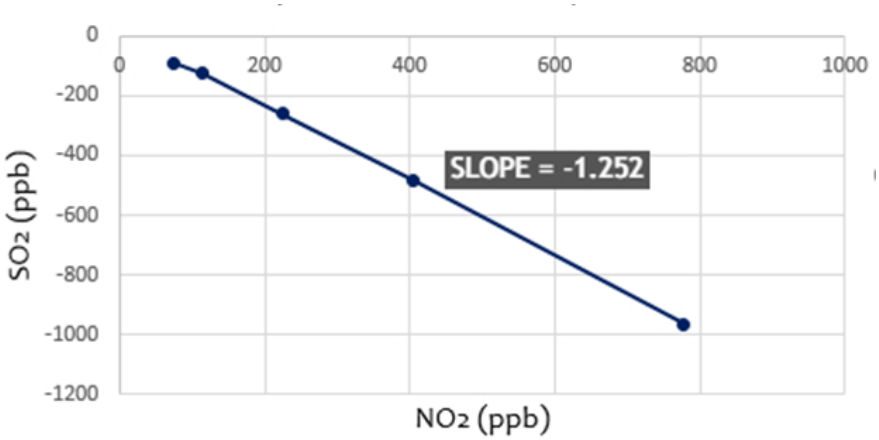

Figure 19. SO_2 cross interference of the NO_2

2.5.3.3 Quality criteria

This system is new and experimental and at the validation experiments in C3 there were no quality criteria yet developed; however, many of the principles are like those described in section 2.3.2.3. Considering the proximity of the mini sniffer to the source, the weather conditions are not playing a major role regarding the quality criteria; in contrast, the signal noise ratio of the measurements is probably the determinant factor.

2.5.3.4 Error sources accounted for reporting

The uncertainty estimation follows a similar approach as described in section 2.3.2.4, where the integrated concentration counts for the differences in the response time, and the uncertainties sources are dominated by the baseline variations and baseline average differences. The SO_2 electrochemical sensor has a NO_2 negative interference; thus, the SO_2 signal is corrected as described in Eq. 24. The expanded FSC uncertainty for the Chalmers mini sniffer is shown in Eq 25 , following the same approach as in Eq. 16 for the standard sniffer, including calibration, baseline difference and signal-to-noise of the measurements.

$$U_{tot} = k\sqrt{\left(U_{SNR}(X_1)^2 + U_{BL}(X_1)^2 + U_{cal}(X_1)^2 + U_{gas}(X_1)^2 + U_{cross}(X_1)^2 + Ut_{90}(X_1)^2 + U_{SNR}(X_2) \dots\right)} \quad \text{Eq. 25}$$

Here the notation is the same as in section 2.3.2.4 with only difference of an added Ut_{90}. This occurs for gas diffusion system (EC sensor) if the period inside the plume is shorter than the t_{90} sensor response time. In this study, this error is neglectable since all measurements intervals inside the plumes are longer than the response time. In addition, the EC sensors also are connected to a flow-through system which also affects the response time, in addition to the gas diffusion.

2.5.3.5 Data reporting

The mini sniffer data reporting follows the same approach from the Chalmers standard and laser system, where a datalogger and a presenter are running in parallel. The system software (section 2.3.2.5) has been upgraded to include the signals of the sensors from the mini sniffer, AIS and local wind parameters as shown on Figure 20. Since the Chalmers mini sniffer is still under an experimental stage, no data reporting protocols have been yet stablished.

Figure 20. Example of the IGPS software display. A) Position display of the mini sniffer and targeted vessel. B) Real time plume measurements.

2.5.3.6 Plume Approach

Though the apparent plume speed and the wind speed/direction indicates the location of the plume relative to the target ship; sometimes the plume altitude cannot be easily located due to very focalized wind fluctuations nearby the vessel. Once the plume has been identified, the flight strategy consists in approaching the ship, typically 10-20 meters away from the emissions source, and flight in a series of zig-zag trajectories in and out the plume. This approach creates a series of patterns where the measurements alternate between a sample in the plume and a clear air baseline. Figure 20b shows a sequence of plumes and their respective baseline as a result of the zig-zag flight strategy.

2.6 Other methods applied within project

A zenith sky DOAS (Platt and Stutz, 2008) system has been tested in the project. It is based on two spectrometers connected to zenith viewing telescopes via optical fibers. One of the spectrometers measures columns of SO_2 in the UV region between 300 and 320 nm and the other columns of NO_2 in the visible spectral region near 450 nm, see Figure 21. It is investigated in the project whether the ratio between SO_2 and NO_2 columns can be used as a proxy for FSC and a good example of this is shown in Figure 22. This system has furthermore been developed into a Multi-Axis DOAS system by combining it with an optical scanner that allows measurements at various viewing angles. From such observations it will be possible to derive vertical profiles of NO_2 and this system was applied in C5 for comparison to the satellite-borne TropOMI instrument.

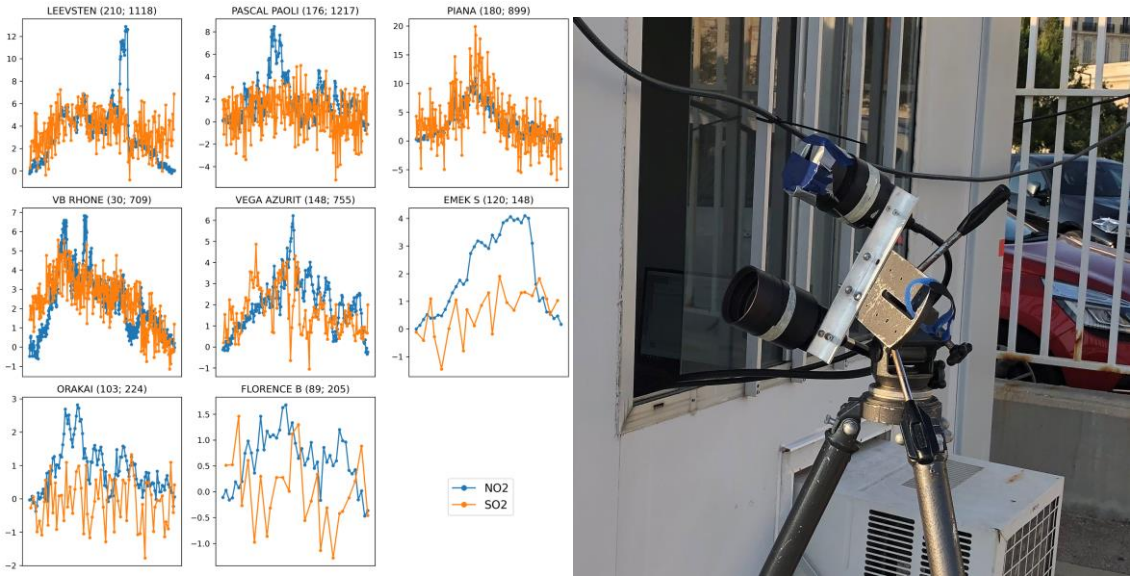

Figure 21. Example of column measurements of SO_2 and NO_2 in the plume of different ships during C4 using zenith sky DOAS by Chalmers. The measurements were carried out from a measurement vessel. The y-axis in the left graph corresponds to the column in mg m^{-2}.

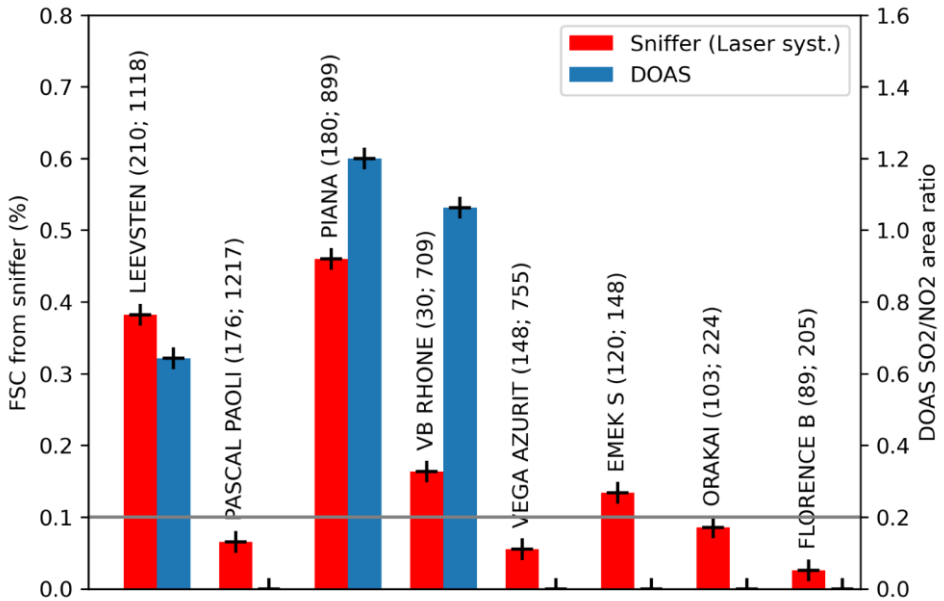

Figure 22. SO_2 over NO_2 column ratios from single ship plumes and comparison to FSC from the high-sensitive laser sniffer system. The numbers in parenthesis correspond to the length of the ship and the distance to the ship in meters.

3 Intercomparisons and validation measurements

Several field campaigns were carried out in SCIPPER in which part of the objective was to validate and compare measurement techniques. The campaigns are described in this section including key results for measurements of NOx emissions and FSC.

3.1 Campaign activities

3.1.1 Marseille campaign, September 2019 (C1)

Campaign 1 (C1) was carried out during September 17 to 28, 2019, in the waters outside port of Marseille water Area and at the nearby industrial port of Fos-de-Mer, see Figure 23. The campaign was coordinated by project partners AMU and Atmosud and measurements were carried out by Chalmers, Aeromon, and AMU. For 5 days, ship emission measurements were carried out from a measurement vessel, a leisure speed boat, which was stationed in the northern part of Marseille. The vessel was equipped with a suite of sniffer instruments for gas and particles (SO_2, CO_2, NO_x, BC, PN, PM2.5) and an optical sensor (zenith sky DOAS) measuring columns of SO_2 and NO_2. The instruments were all operated by Chalmers. The measurement vessel tracked ships that were in-, or out-bound to the two ports of Marseille and Fos-de-Mer. In addition, sniffer measurements were carried out by Chalmers from a fixed site in the port of Marseille from September 23 to 28, indicated in Figure 23. AMU did complementary measurements in a northern location at the same time. The measurement vessel carried a drone-based sniffer (BH-12) from project partner Aeromon as well. Measurements of 15 vessels were carried out simultaneously with the Chalmers sniffer system for comparison purposes.

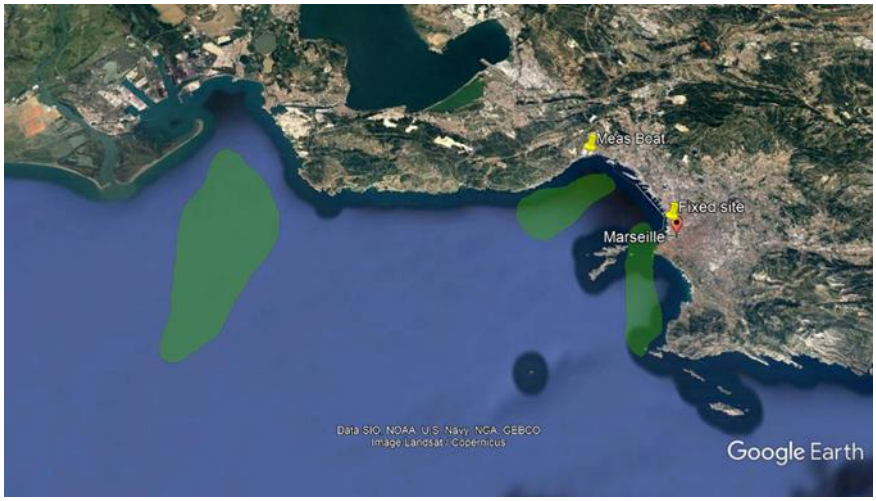

Figure 23. The measurement area in campaign 1 and 4. The underlying satellite picture is taken from google earth.

3.1.2 Wedel campaign, September 2020 (C3)

Sulphur emissions from sea-going and inland vessels to the ambient atmosphere were measured by several different state-of-the-art and new monitoring systems during a six-week long measurement campaign between September 7 and October 15, 2020 at the river Elbe at Wedel near Hamburg, Germany, see Figure 24.

Figure 24. Setup of the side-by-side inter comparison systems at the BSH measurement site in Wedel. Left to right: Chalmers (white container and blue VW transporter car), BSH mobile measurement station (white trailer), TNO (artistic trailer), Explicit (hovering drone), BSH (elevated platform at building).

The campaign was organized by BSH and other participating groups were Chalmers, TNO and Explicit. The respective instrumentation is presented in Table 1, see Sec. 2.1. Various systems for land-based and UAV-borne application were operated by the four participating partners in a side-by-side measurement setup to observe emissions under similar conditions These were 5 standard sniffers (3 BSH, 1 Chalmers 1 TNO), 2 mini sniffers (Explicit and Chalmers), 1 laser spectrometer (Chalmers), 2 DOAS systems (BSH/Uni Bremen and Chalmers), 4 particle size classifiers (2 BSH, 1 Chalmers, 1 TNO), 2 aethalometers (Chalmers and TNO) and one particle counter (TNO), as well as 5 meteorological stations and 4 AIS receivers. In total 966 plumes of 436 different vessels were analysed. Additionally, 60 fuel samples were taken from onboard of 34 selected vessels by the German waterway police. The fuel samples were analysed in an ISO 17025-certified laboratory operated by BSH, to determine the FSC as reference to the respective results obtained with the remote methods. Furthermore, the bunker delivery notes for some selected ships were requested from the shipping companies.
The unique benefits of Campaign 3 are first the high number of simultaneously measured plumes using different monitoring systems and second that the measured SO_2 emissions could be compared to numerous fuel samples or bunker delivery notes. Hence, the derived sulphur content in the fuel measurements could not only be compared between the systems but also to the actual laboratory confirmed fuel sulphur contents which represent the true FSC.

3.1.3 Stena Germanica and onshore Kiel Campaign, September 2021 (C2)

In September 2021, the SCIPPER C2 campaign was carried out in the Baltic Sea on board the passenger ferry Stena Germanica operating between Gothenburg (Sweden) and Kiel (Germany). Details of the on-board measurements are given in Sec. 1.4.1 of the second SCIPPER mid-term report Part B and the SCIPPER Deliverable D1.3. To support the onboard measurements and to investigate the development of the ship plume during transport at different plume ages, additionally remote measurements were carried in the bay of Kiel about 9 km north of the STENA terminal from August 27 until September 13, 2021 by Explicit, BSH and Chalmers (Figure 25). Explicit was operating a UAV (DJI Matrice 300) inside the orange marked area were arriving and departing ships to/from Kiel and the Kiel-Canal pass through. The UAV was equipped with a mini sniffer and was launched from a small meadow on the eastern shore of the Kiel Fjord. BSH operated their standard sniffer and particle size classifiers about 1 km north of the UAV launching site inside a fixed shelter. Chalmers installed their sniffer system composed of the demonstrator laser spectrometer system for SO_2 and CO_2 measurements and a NO_x analyser into a multi van which was parked at a campsite near the UAV launching site.

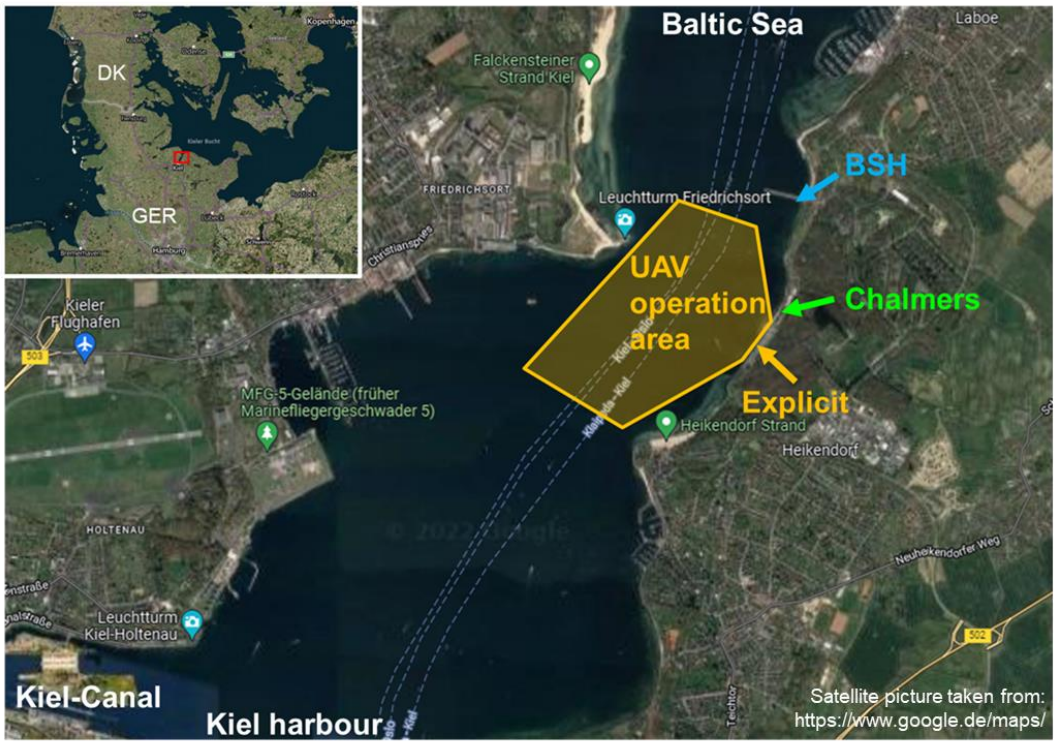

Figure 25. Location of remote measurement carried out by the project partners Explicit, BSH and Chalmers during SCIPPER C2 campaign in the Kiel Fjord (Germany). The underlying satellite picture is taken from google maps (https://google.de/maps/).

The ferry travelled overnight and passed the measurement area in Kiel every second day once in the morning (arriving) and once in the evening (departing) on the same day. Thus, in the measurement period the plume of Stena Germanica could be measured 18 times by the Explicit UAV. Since the wind conditions were sometimes unfavourable during the measurement campaign and due to occasional technical issues of one of the land-based sniffer systems, the ship plume from Stena Germanica could be measured with the fixed stations only 8 and 3 times, respectively. In addition, side by side measurements with the two fixed sites and the UAV were carried out on a large number of ships, namely 211 by BSH, 149 by Explicit and 87 by Chalmers, respectively.

3.1.4 Campaign 4: Marseille, July 2021 (C4)

The second Marseille campaign, C4, took place in July 2021. It came more than a year after the implementation of the global sulphur cap, providing the opportunity to check compliance rates of ships under the new regulations. Within this campaign, remote techniques, both, sniffer vessel and drones, were again deployed and intercompared, similarly to the campaign in 2019, with the aim to characterize emissions and compare techniques.
During the 2021 campaign sniffer and optical measurements were carried from a sniffer vessel in conjunction with measurement within SCIPPER Work package 3. In addition, there were fixed measurements carried out during a week from the southern part of the port at La Major. This included sniffer measurements for SO_2, CO_2, NO_x, particle size distribution, black carbon (Aethelometer), VOC measurements by FTIR and column measurement by zenith sky DOAS. Drone measurements were carried out again by Aeromon using a BH-12 mini sniffer equipped with SO_2, CO_2, NO, NO_2, PM1, PM2.5 and PM10 sensors. These measurements were performed from the northern part of the port of Marseille during the same days as the sniffer vessel was operated.

3.2 Key results for FSC and NO_x

3.2.1 FSC measurements

In C1 in Marseille in September 2019, the ships were still allowed to sail with a FSC limit of 3.5 % S m/m outside the SECA area. Remote measurements of ships were carried out using two systems that were operated in parallel from a fast-moving leisure vessel, i.e. an airborne version of the standard sniffer, and a drone-based mini sniffer. Note that the airborne version of the standard sniffer is different from the ones used on fixed stations since, in order to be able to perform fast measurements, it has had the "kicker" removed in the SO_2 UV-fluorescence instrument. This makes it cross sensitive also to organic compounds adding additional uncertainty, see section 2.1. In Figure 26, a comparison is shown of remote measurements by the airborne version of the standard sniffer (Chalmers) and a drone equipped with mini-sniffer equipment (Aeromon) for 13 different ships operating on the waters of Marseille and Fos-sur-Mer. There is overall good agreement between the measurements in terms of magnitude and correlation (slope 0.99. offset -0.02 % S m/m, R^2 0.95) and the 1σ–variability between individual measurements is 0.20 % S m/m. If the average of the two systems is assumed as the "true" emission, then the random variability for each system would be around 0.1 % S m/m and the expanded uncertainty (CI 95 %) approximately twice as large. In absolute FSC numbers this apparent uncertainty is larger than observed in other studies (C2 and C3) but not when considering relative uncertainties, given that the FSC values ranged up to much larger numbers than in the other studies. In addition, it should be considered that the used Chalmers sniffers is made for airborne measurements and therefore has had the "kicker" removed, which make the measurements susceptible to volatile organic compounds present in the exhaust gas and which can add 0.1 % S m/m units. Thirdly, the drone measurements were only allowed to be conducted at 200 m distance from the ship and this is longer than in C3, in which measurement were carried at 50 to 75 m distance.

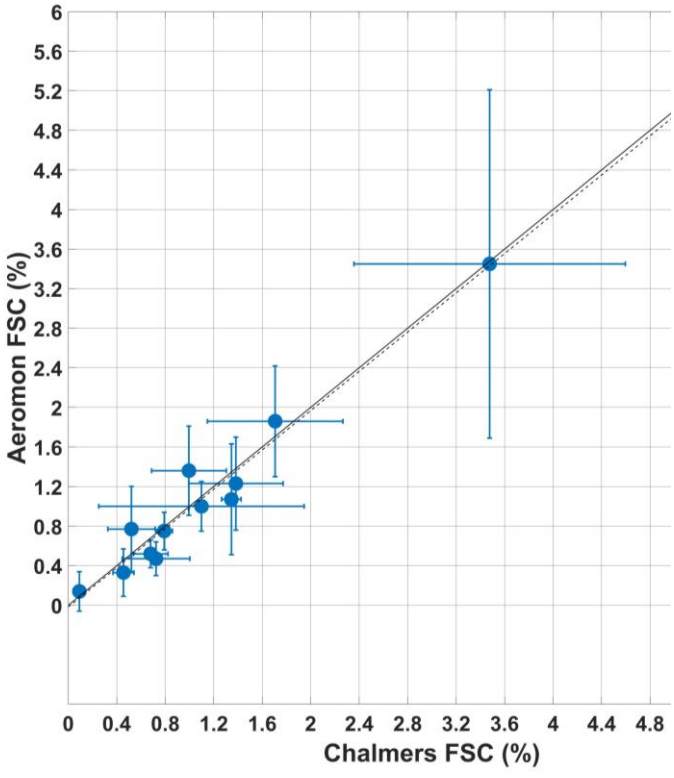

Figure 26. Comparison of sniffer (Chalmers) and drone (Aeromon) measurements of different ships operating on the waters of Marseille and Fos-sur-Mer during September 2019 (C1).

During the intercomparison campaign in Wedel in autumn 2020 (C3) one significant source of deviation and error was found in the quality of the applied calibration standards for SO_2, obtained from certified gas manufacturers. A later comparison of the calibration standards to a reference standard at TNO's laboratory confirmed that the

calibration standard supplied to one group deviated from the supplier's specifications by nearly 60 % and this has been accounted for in the results presented in this report.

The apparent measurement error is obtained from the absolute FSC difference between the remote sniffer measurements and laboratory analysed fuel samples. It is summarized in Figure 27 for the different measurements systems employed in the Wedel campaign. Note that this comparison was carried out as a blind exercise in which the data for the fuel samples were released first after all group had sent in their data.

It can be seen that most systems underestimate the actual FSC onboard the vessels, while this is not, or to a lesser extent, the case for the drone-based mini sniffer measurements as conducted by Explicit and Chalmers. A larger spread can be seen in the results from the Chalmers drone-based mini sniffer, cha.uas, when compared to the one operated by Explicit, exp.uas, but it should be noted that the former was an experimental system used for the first time, while the Explicit system was already at a mature state and had been used extensively in regular operation. In addition, the take-off position for the Chalmers drone-based mini sniffer was considerably further away from the ships than the Explicit one. This limited the operational range and thus restricted optimal flying conditions in several cases, and the measurements were generally carried out further from the ships (100-200 m) than for the mature drone. Note that in the analysis below we will generally consider the Explicit drone-based mini-sniffer.

From the standard deviation of the apparent measurements error, Figure 27, the expanded random measurements error (CI 95 %), which is not taking systematic biases into account, can be calculated for the different systems. This error corresponds to 0.04 % S m/m, for the high-sensitive system (cha.las), 0.05 % to 0.08 % S m/m for the different standard fixed systems and 0.08 % S m/m and 0.13 % S m/m for the mature (exp.uas) and experimental drone-based sniffer system (cha.uas), respectively. In addition, for the standard and high-sensitive fixed systems there is a systematic negative bias in the FSC data ranging from 0.02 % to 0.07 % S m/m, see discussion on this below. The total measurements error (CI 95 %), in the presence of a bias, can be obtained by combining the random and systematic errors by their sum of squares (Magnusson et al. 2008, RSSU table 3), see Eq. 26 in section 5.3. Using this approach, the total measurement error for the data shown in Figure 27 correspond to 0.05 % S m/m for the high-sensitive fixed system, 0.07 % to 0.09 % S m/m for the standard fixed systems, depending on instrument, and 0.08 % S m/m for the mature, and 0.14 % S m/m for the experimental drone-based systems, respectively.

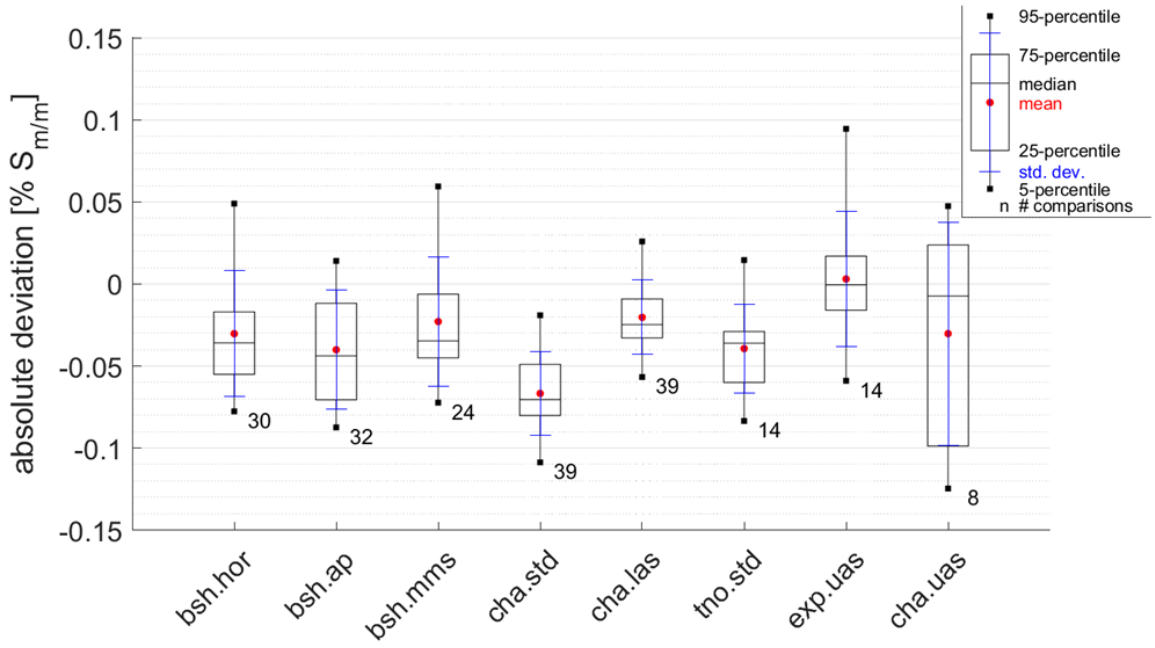

Figure 27. Summary of the measurement error, corresponding to absolute difference between remote FSC measurements and laboratory-analysed fuel samples (main engine), for 62 measured ship plumes. Here bsh.hor, bsh.ap and bsh.mms represent the three standard sniffers operated by BSH and cha.std, cha.las and cha.uas correspond to the results from the standard-, laser- and the mini sniffer systems operated by Chalmers. The TNO standard sniffer results and Explicit mini sniffer results are indicated with tno.std and exp.uas, respectively. The fuel samples used for this comparison were taken from the fuel system of the main engine.

In Figure 28 the apparent measurement error is shown versus the reported uncertainty, by each group and each individual measurement. From this plot it can be seen how well the apparent measurements error is explained by the reported uncertainty values. The estimation of the measurement uncertainties has been done in different

manners for the different institutes/companies, following their own practice as described in section 2. E.g. several of the groups (BSH, TNO) report their uncertainties as CI 68 % while others (Chalmers) use CI 95 %, all based on the individual measurement data for each plume. For the drone measurements by Explicit, the uncertainty is precalculated, based on the measured FSC and on that pre-defined quality thresholds have been met. See sections 2.3 to 2.5 for further details about the individual uncertainty calculations and section 5.3 for a general overview of our recommendations.

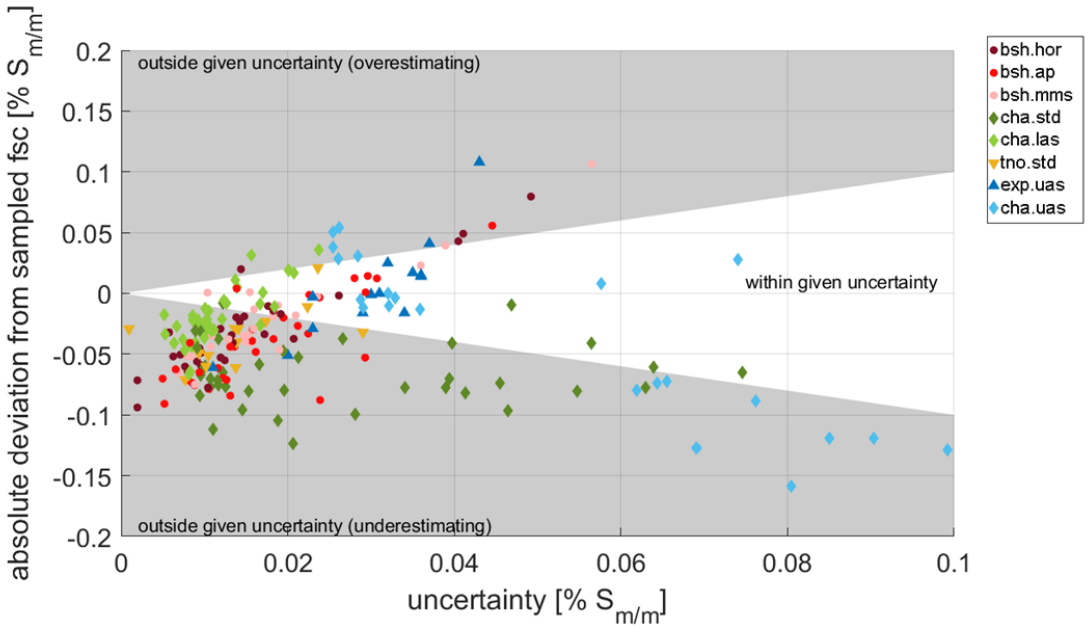

Figure 28. The measurement error versus the reported uncertainty which is reported by the respective systems for measurements obtained during C3 in Wedel. Different colours indicate the respective results of the different systems. The data points within the unshaded area indicate results which match the laboratory found FSC values within the reported uncertainty. Results in the grey-shaded area are either under- or overestimating the FSC when taking into account the individual measurement uncertainty.

The data shown in Figure 28 have been compiled in Table 5. For the fixed, land based, sniffer systems the reported uncertainties are too low in more than 70 % of the measurements, with one exception. The results measured with the drones are spread more evenly around the expected FSC and for the mature UAV-borne system (exp.uas) the reported uncertainties explain the measurement error in about 64 % of the cases.

Table 5. Summarizing table of absolute deviations of FSC from laboratory analysed fuel sample for all measured plumes per system during C3 in Wedel. The fuel samples used for this comparison were taken from the fuel system of the main engine. For clarifying the abbreviations for system see caption to Figure 27.

system	total number of samples	matching within reported uncertainties	overestimating, considering reported uncertainties	underestimating, considering reported uncertainties	median of reported uncertainties
bsh.hor	30	13 %	13 %	73 %	0.013 %$S_{m/m}$
bsh.ap	32	22 %	3 %	75 %	0.015 % $S_{m/m}$
bsh.mms	24	29 %	8 %	63 %	0.015 % $S_{m/m}$
cha.std	39	13 %	0 %	87 %	0.021 % $S_{m/m}$
cha.las	39	23 %	5 %	72 %	0.008 % $S_{m/m}$
tno.std	14	14 %	0 %	86 %	0.014 % $S_{m/m}$
exp.uas	14	64 %	14 %	21 %	0.029 % $S_{m/m}$
cha.uas	22	36 %	23 %	41 %	0.054 % $S_{m/m}$

In Figure 29 the apparent measurements error for each instrument has been plotted versus relative humidity. As can be seen, there appears to be a dependence of RH for the fixed systems bsh.hor, cha.std and cha.las. In addition, there

may be a weak RH dependence also for some of the other systems (bsh.ap, bsh.mms and cha.uas). Note that the data for the TNO sniffer (tno.std) has few points below RH 80 % and the RH dependence can therefore not be interpreted. The drone-based mini sniffer system (exp.uas) does on the other hand does not appear to be influenced by relative humidity.

The apparent effect of relative humdiity on the FSC measurements is not well understood and requires further studies. Nevertheless, it could be caused by increased adsorption of SO_2 on internal surfaces in the sampling system, e.g. tubings and filters, and in the instrument itself. This could likely be abated by shorter tubing lengths, active heating of relevant surfaces or excluding measurements with high relative humidity. Note that the length of the tubings varied in the Wedel study; the Chalmers standard sniffer, cha.std, which had the largest negative bias also had the longest inlet tubing, approx. 10 m, while the drone-based systems had the smallest bias and shorter tubings, i.e. less than 0.5 m. The drone-based system also measured quite close to the ships, with two orders of magnitude higher concentrations of SO_2 than the fixed systems, which may make it less susceptible to adsorption effects.

Noteworthy, is that the negative bias and the 2σ random error for the high-sensitive fixed system (cha.las) is significantly reduced if only using data acquitted for RH below 80 %, i.e. (-0.0035±0.028) % S m/m instead of (-0.017±0.04) % S m/m for the data in Figure 29 which are based on fuel samples and additional 24 ships with available bunker delivery notes. The corresponding total measurement errors, according to Eq. 26, then reduced from 0.044 % to 0.028 % S m/m.

A second reason for the bias in the standard measurements could be the fact that dry calibration gas is used in the calibration. Since UV fluorescence may be affected by water vapour this could cause systematic errors in the plume samplings. A third reason for the bias in the fixed standard systems could be caused by overcompensation of the cross sensitivity of the SO_2 instruments to NO. For the TNO and BSH systems the applied cross-interference factors for NO were 0.85 % and 0.5 %, respectively while the factor for Chalmers system was 1.5 % (2-3 times higher); this might explain the larger bias of the Chalmers standard sniffer system and needs to be revisited.

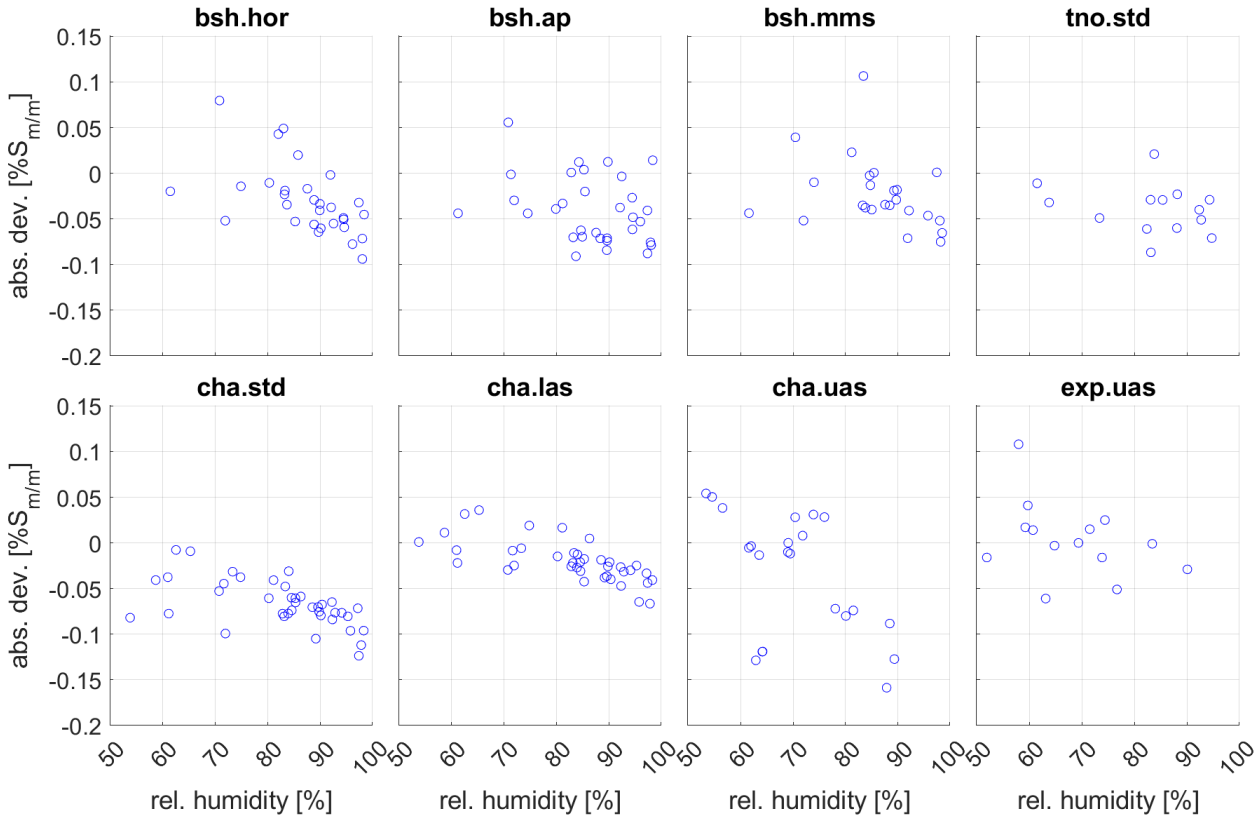

Figure 29. Absolute deviation of measured FSC in by-passing ships by the different remote systems to the laboratory analysed fuel samples.

During C2, individual ships were measured by the BSH standard sniffer system, the Chalmers high-sensitive system and the Explicit drone-based mini sniffer during several weeks for ships traveling to or from Kiel and through the

Kiel channel. For more details about the campaign see section 3.1.3. The results for one of the ferries that passed regularly, i.e. Stena Germanica, were compared to the analysed fuel composition corresponding to FSC of 0.095 % S m/m, Figure 30. In a corresponding manner as for the Wedel campaign, the apparent measurements error was obtained from the absolute deviation between the measured FSC and the fuel composition of the oil. Note that the measurement statistics for the remote systems is appropriate for the drone-based measurement comparison but rather poor for the fixed sniffer systems since the wind seldom blew in the measurement direction. It can be seen in Figure 30 that there is a negative bias for the fixed systems, i.e. on average 0.03 % S m/m and 0.067 % S m/m for the high-sensitive sniffer and standard one, respectively. The drone-based mini sniffer on the other hand shows a minor positive bias of 0.015 % S m/m. The reported uncertainties of the drone-based mini sniffer are able to explain the measurements error in 90 % of the cases, while the reported uncertainties of the fixed systems are unable to match the apparent measurements error. From the data in Figure 30 the expanded measurement error can be calculated using Eq. 26, including also the bias. The expanded measurement error (CI 95 %) of the fixed sniffer systems is then 0.036 % S m/m for the high-sensitive one and 0.08 % S m/m for the standard one. The error is dominated by the bias here. For the drone-based mini sniffer the expanded measurements error is 0.03 % S m/m, dominated by the random measurement error. In Figure 30, the RH is also indicated, varying between 73 % to 98 %, and even in the dryer conditions, i.e. at relative a humidity of 73 %, there is still a bias, in contrast to C3.

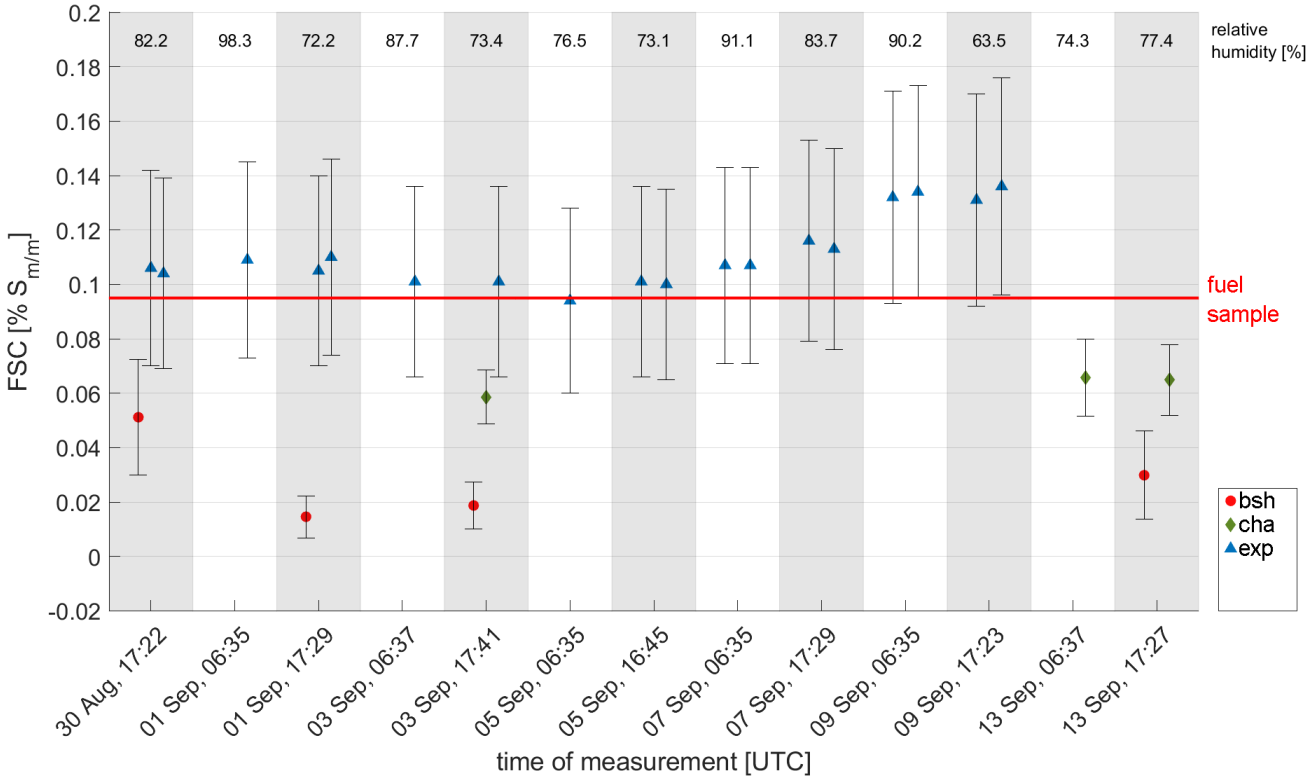

Figure 30. Comparison of the on-board and Kiel shore-based measurements of the FSC from Stena Germanica during SCIPPER campaign C2 in August and September 2021. Three systems were operated side-by-side, i.e. a high-sensitive fixed sniffer (cha), a standard fixed sniffer (bsh) and UAV-based mini sniffer system (exp).

At the Kiel site during the C2 campaign, 20 additional ships were measured, but with unknown FSC-values, see Figure 32. A similar pattern can be seen as in Figure 30, i.e. the FSC values of the drone-based mini sniffer (exp) are highest while the sniffer systems are lower. On the other hand, the high-sensitive system (cha) and the standard sniffer system (bsh) do not differ as strongly from each other. There is no apparent correlation between RH and the observed differences in FSC between the drone-based measurements and fixed systems.

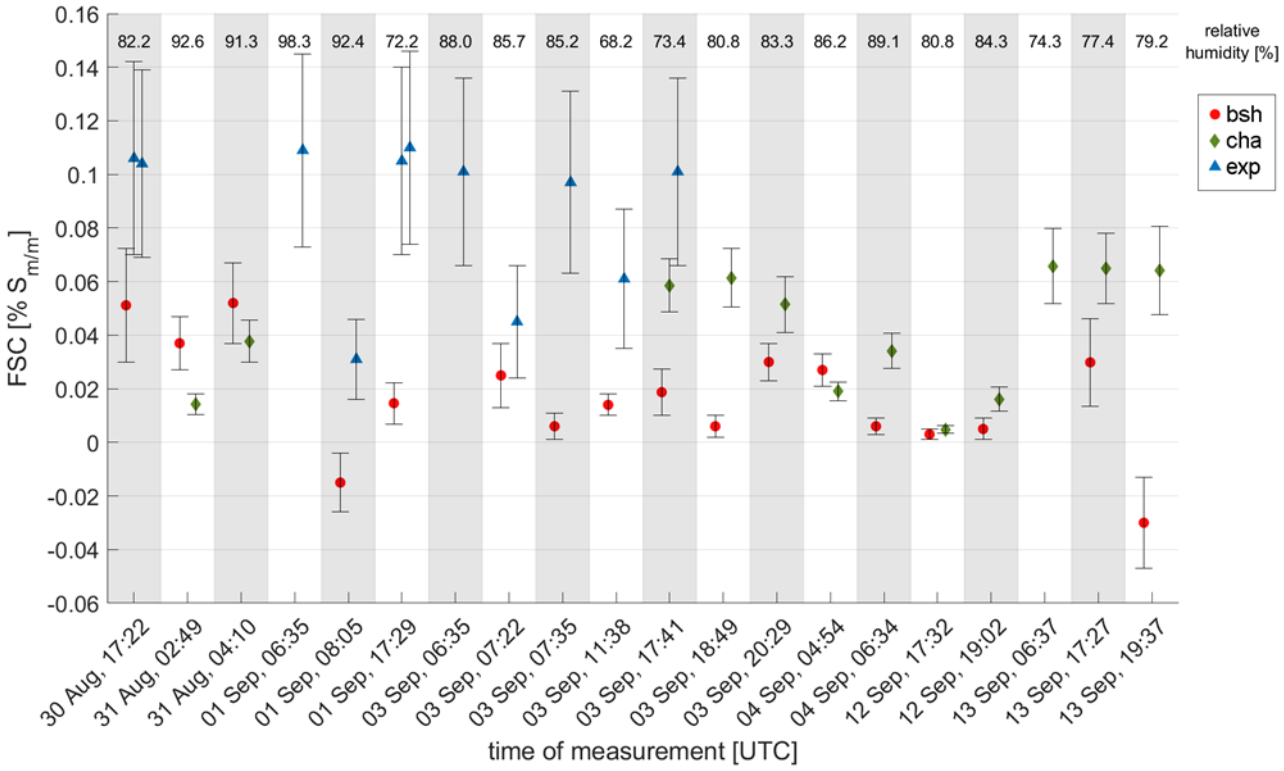

Figure 31. Side-by-side measurements in Kiel during campaign C2 in August and September 2021. Three systems were operated side-by-side, i.e. a high-sensitive fixed sniffer (cha), a standard fixed sniffer (bsh) and UAV-based mini sniffer system (exp).

In C4, measurements were carried out in the waters outside Marseille and Fos–sur-Mer from a measurement vessel (in the same manner as in C1). A comparison of the Chalmers standard sniffer system against the high-sensitive one is shown in Figure 32 together with the estimated expanded uncertainties. It can be seen that the standard system generally shows lower values (slope 0.85. offset -0.028 % S m/m). but that the two systems correlate well (R^2 0.94). The standard deviation of the difference between individual measurements is 0.05 % S m/m. One potential reason for the lower values by the standard sniffer system, as discussed in relation to C3, is that there could be an overcompensation of the NO interference.

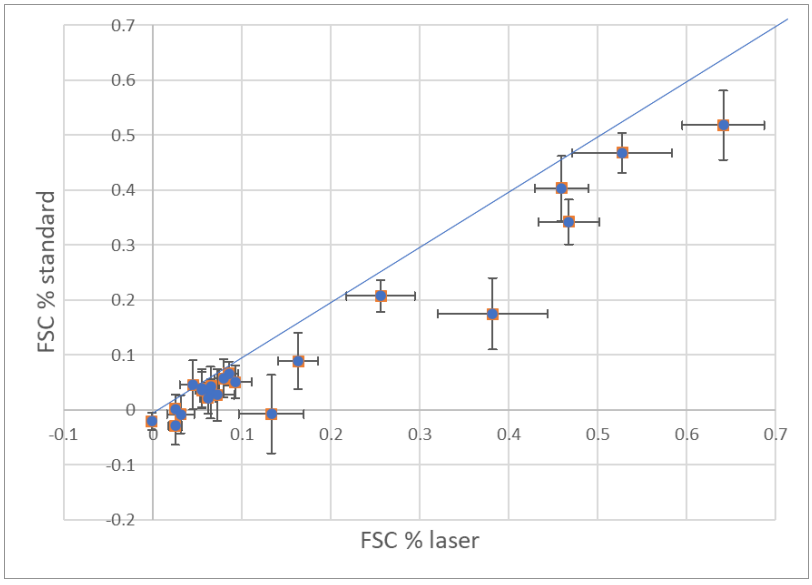

Figure 32. Measurements of 22 ships outside Marseille in 2021 during C4 using the Chalmers laser and standard sniffer instrument.

3.2.2 NO_x emissions

In several measurement campaigns, i.e. C1, C2 and C3, remote measurements of fuel-mass specific emissions of NO_x have been compared between different sensors. To assess the typical measurements error of the instruments, the data by the different groups have been compared against the ensemble average, corresponding to data from at least two or three coincident measurements. Note that the average is not necessarily the true value but here we assume so, in lack of other information, and in consistency with section 3.2.1 on FSC we assume the deviation from the ensemble mean to correspond to the apparent measurement error. Note, that in cases when the ensemble average corresponds to only two systems, both systems will have the same uncertainty.

Remote measurements of fuel-mass specific NO_x emissions (Emission Factors g/kg_{fuel}) were carried out by drone-based mini sniffer (Aeromon) and mobile standard sniffer (Chalmers) during C1 in Marseille, September 2019, Figure 33. The measurements were conducted simultaneously from a small and fast measurement vessel (Chalmers) outside the waters of Marseille and Fos-sur-Mer. The measurements by the two systems have a systematic difference of 5 g/kg_{fuel} (13 %) with 7 g/kg_{fuel} 1σ variability (17 %). If the average of the two systems is assumed as the "true" emission, then the total measurement error (CI 95 %) corresponds to 7.5 g/kg_{fuel} (17 %) following Eq. 26. The reported uncertainties by the two systems explain the differences in 90 % of the cases.

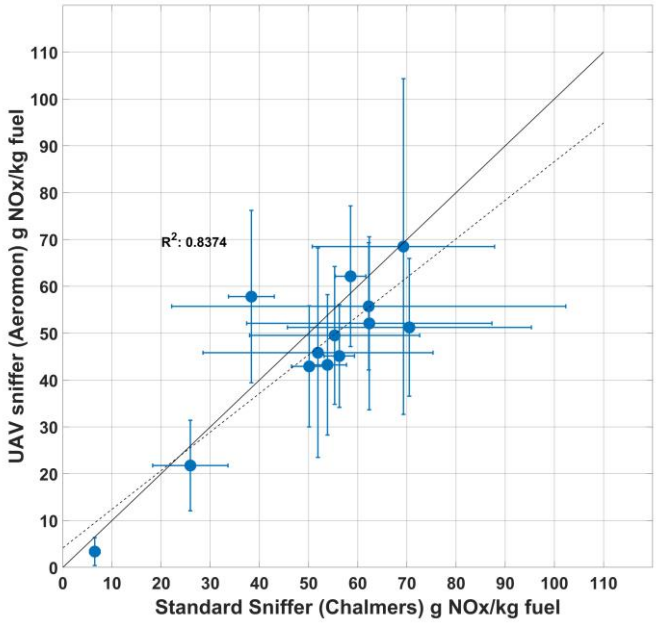

Figure 33. Comparison between Chalmers sniffer and Aeromon drone during C1 in Marseille in September 2019. Measurements were carried out from a small and fast vessel outside the water of Marseille and Fos-sur-Mer.

A similar comparison, as in C1, was carried out in the Kiel channel in September 2021 as part of C2, Figure 34. Remote measurements of NO_x Emission Factors (g/kg_{fuel}) were carried out on the eastern shore of the Kiel channel from two fixed sites, separated 1 km apart, by two standard sniffers operated by Chalmers (cha) and BSH (bsh), respectively. In addition, a drone-based mini sniffer (exp) was operated in vicinity of the fixed systems during most of the period. There is good correlations between the different measurements with an agreement better than 10 % between the data sets. The differences between all three systems can be explained by the estimated uncertainty in 85 % of the cases. It should be noted that the drone-based measurements on Stena Germanica (see days in Figure 30) where carried out only for the portside engines while the fixed sniffers systems measured the mixed emissions from all engines: the large difference on August 30 could be caused by this difference. The comparison shows a systematic and random (1σ) measurement error of -4.8±4.7 g/kg_{fuel} (-8.5±11) %, 4.0±4.9 g/kg_{fuel} (-6.9±11) % and -2±4.3 g/kg_{fuel} (-3.3±8.5) % for the NO_x EF measurements by Chalmers, BSH and Explicit, respectively. Here the measurements of Stena Germanica on August 30 were excluded for the reason explained above. When adding the systematic and random measurements error together according to Eq. 26 this corresponds to total measurement error (CI 95 %) of. 10.6 g/kg_{fuel} (23 %) for the fixed, shore based, systems and 8.8 g/kg_{fuel} (17 %) for the drone-based mini sniffer. Noteworthy is that the reported calculated uncertainties, obtained individually by the different groups, are around 20 % for all systems, which is almost double the apparent measurements error.

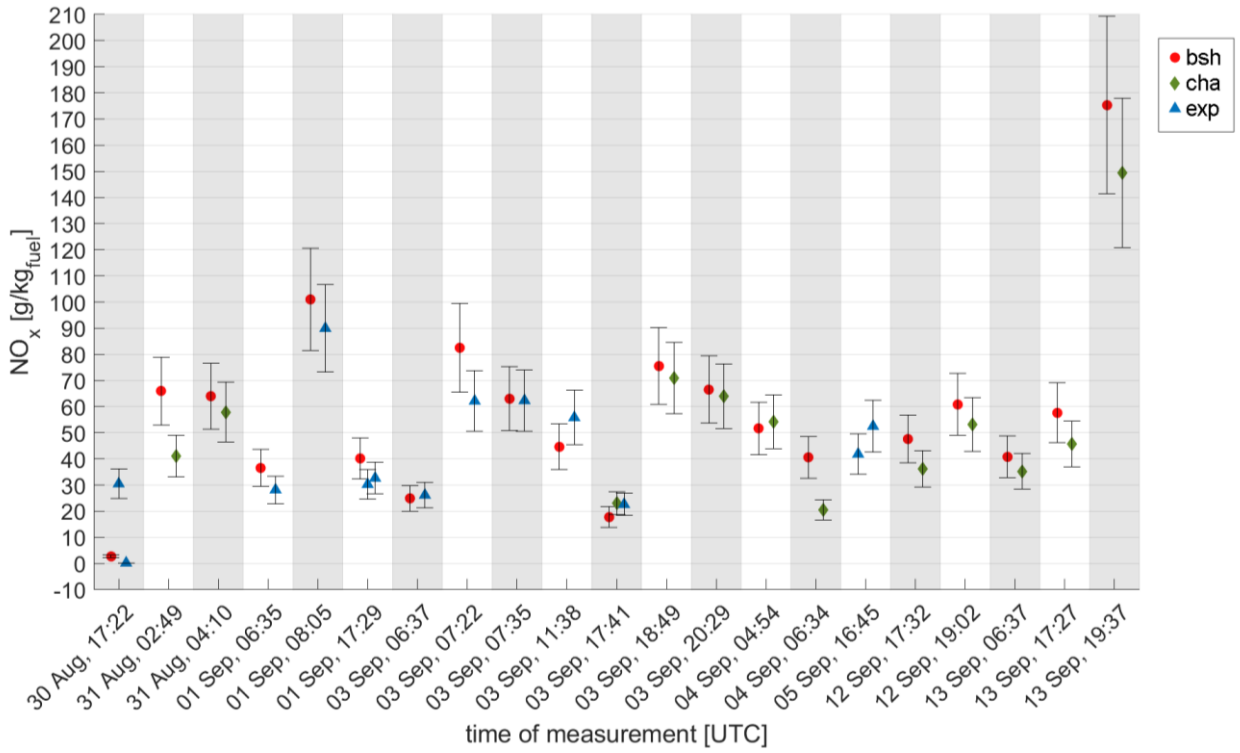

Figure 34. Comparison of NO_x emission factors (g/kg_{fuel}) for individual ships from fixed sniffers by Chalmers and BSH and drone-based mini sniffer by Explicit during C2 in Kiel. The fixed sites were positioned on the eastern shore of the ship channel with 1 km separation.

An extensive comparison exercise was carried out during campaign 3 in Wedel (Sec. 3.1.2) with more sensors and longer duration than in the other campaigns. Chalmers measured the NO_x concentration during the first half of the campaign and switched to NO measurements thereafter. The data was manually evaluated and, in the comparison, only the NO_x data were used. BSH measured NO and NO_x by switching every 10 s and TNO had a dual chamber instrument which measured NO and NO_x in parallel. The latter instruments had a failure in data storage leading to significant data loss. The NO_x data from TNO and BSH were evaluated automatically. In Figure 35 the NO_x data for the individual instruments have been plotted against the ensemble average and in Figure 36 the deviations from the ensemble average (absolute and relative) are shown for each instrument. Note that average deviation from the mean and the standard deviation is shown in the latter figure. The comparison shows systematic and random (1σ) measurement errors of 5.5±8.4 g/kg_{fuel} (12±19) %, -5.1±8.4 g/kg_{fuel} (-14±20) % and -1.2±8.8 g/kg_{fuel} (1.6±18) % for the NO_x EF measurements by Chalmers, BSH and TNO, respectively. This corresponds to a total measurement error (CI 95 %) of approx. 17 g/kg_{fuel} (40 %) for all groups, based on Eq. 26. This error is almost twice as large as obtained in campaign C1 and C2. Part of the differences in variability, to the other campaigns, could be caused by the fact that some of the data were evaluated automatically and that there is potentially a larger spread in ship types and engine load, than in the other campaigns, since the ships were sailing with, and against tide.

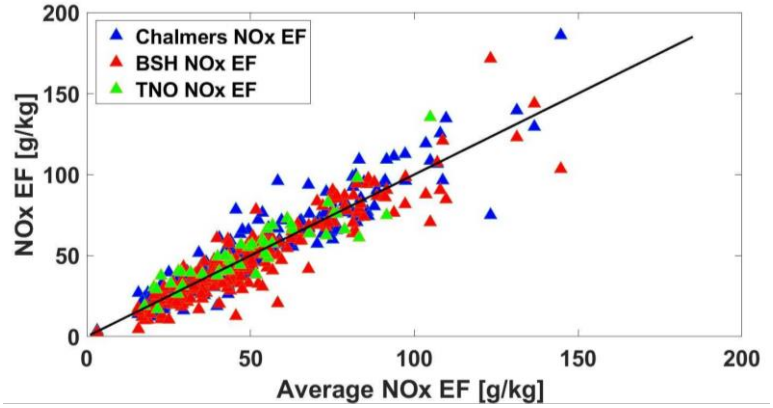

Figure 35. NO_x EF data (g/kg_{fuel}) for individual ship emission measurements by different groups versus the ensemble average. Data with two or three coincident measurements have been used.

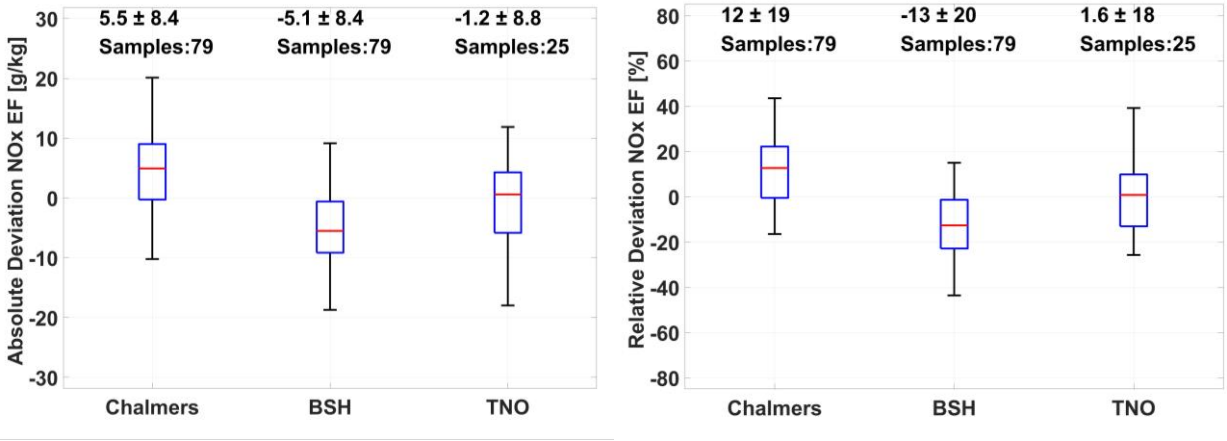

Figure 36. Statistics (median in red and 5th, 25th,75th, 95th percentiles in blue and black) of the absolute and relative deviation of NO_x emission factors measured individually by Chalmers, BSH and TNO for the same ship, against the corresponding ensemble average. Data were measured in Wedel 2020 during C3 and only data with three coincident measurements were used. The numbers given on top correspond to average value and the standard deviation.

In Figure 37, multiple remote NO_x measurements of the same ship, i.e. a suction dredger, by the different groups are shown. Note that the data were not always measured at the same ship passage nor same time period, due to meteorological and operational conditions as can be seen in the figure. This is hence not a strict intercomparison, but the differences and variability indicate how much information there is in a single "snapshot" measurement compared to it's average. The median value of the deviation varies between 10 % to 24 %, depending on technique and group, while the variability of the individual measurements is higher.

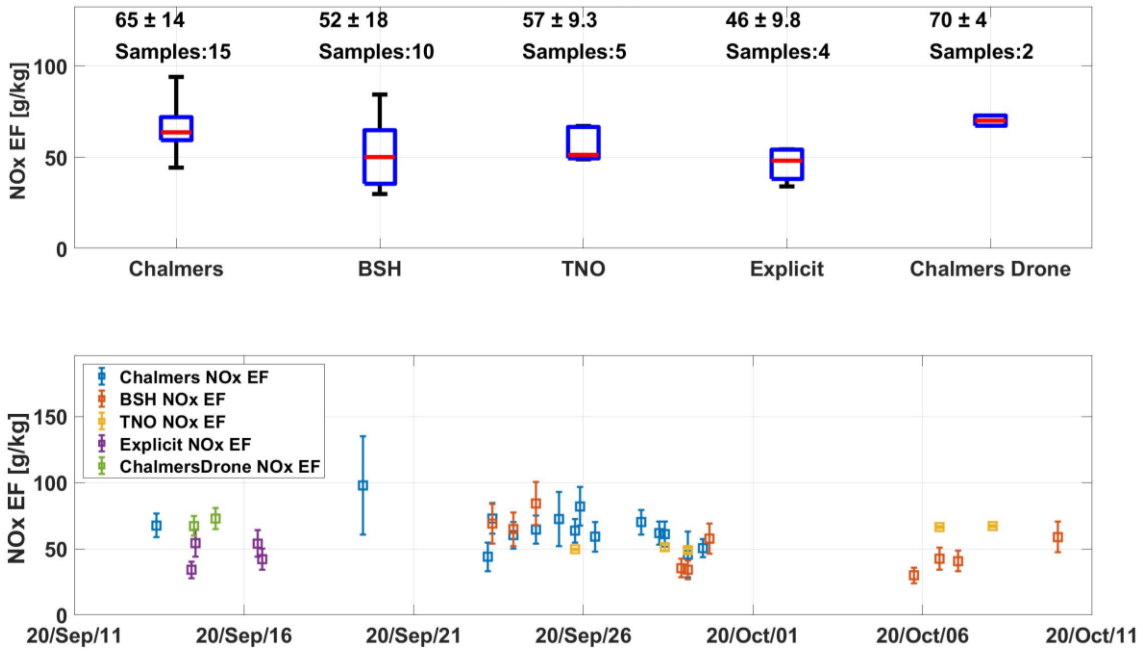

Figure 37. Multiple measurements during C3 by several instruments of the fuel-mass specific NO_x emission rates for a ship (Suction Dredger) that passed the Wedel site many times. The upper part of the figure shows statistics (median in red and 5th, 25th, 75th, 95th percentiles in blue and black) of the absolute and relative deviation of NO_x emission factors (g/kg_{fuel}) measured individually by Chalmers, BSH and TNO for the same ship, against the corresponding ensemble average. The numbers given on top correspond to average value and the standard deviation. The lower figure shows the measured emission factors of NO_x by the different systems versus time, and estimated measurements uncertainty.

In Figure 38 the reported measurement uncertainties of the standard sniffers in C3 are compared to the deviations from the ensemble average (here considered as a proxy for the measurement error), like it was done for FSC in Figure 28. The same data has also been compiled in Table 6. It is apparent that one of the groups assign too small uncertainties to its measurements, since they are realistic in only 20 % of the cases, while the uncertainties reported by the two other standard sniffer groups match the observations in 60 % and 70 % of the cases, respectively. It should be noted that this group has much fewer samples than the other groups, i.e. 25 instead of 79 ships.

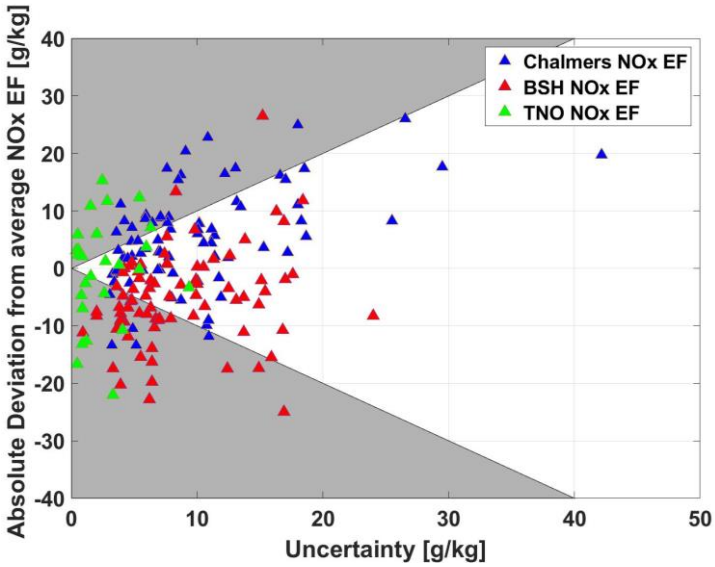

Figure 38. The absolute deviation of the NO_x EF measurements by the individual groups in C3 relative to the ensemble mean is plotted versus the estimated uncertainty.

Table 6. The number of NO_x measurements for the deviation of the NO_x EF measurements of the individual groups relative to the ensemble mean is categorized according to the estimated uncertainty.

Institution	Total Number of Samples	matching within reported uncertainties	overestimating, considered reported uncertainties	underestimating, considered reported uncertainties	median of reported uncertainties
Chalmers	79	68%	6%	25%	8.0 [g/kg]
BSH	79	56%	42%	3%	6.7 [g/kg]
TNO	25	24%	36%	40%	1.5 [g/kg]

The fuel-mass specific emissions can be converted to brake specific emissions (g/kWh) by multiplying with the load dependent SFOC, obtained from Eq. 3 and Eq. 4. For C3, the engine load of the ships was calculated by the STEAM model (Jalkanen, 2009), using the actual ship speed relative to the tide at the time of the monitoring . Here the ship speed was obtained by AIS, while the tide information was provided by BSH.

In Figure 39 the NO_x emission rates measured for various vessels in Wedel (C3), given as fuel-mass specific and brake specific emission, respectively, are shown versus computed load. It can be seen that the emissions are considerably higher and more variable at lower loads. There are relatively few measurements at the higher loads, due to the location of the measurements site, but other studies by project partners, i.e. Chalmers and Explicit, confirms similar trends in other locations, i.e. Chalmers for measurements at Great Belt bridge (Mellqvist et al., 2017) and Explicit on water outside Denmark (Knudsen, 2022). The larger observed variability at lower loads is to be expected since the 25 % and even the 50 % load point of the E2 and E3 cycles (IMO NOx technical code MEPC.177(58)) are not very important for the weighted average NO_x emission (see Eq. 6, section 2.2[3]). Hence, the engines NO_x optimisation is focussed towards medium and high loads, which leads to convergence at higher loads. Additionally, there are no applicable not-to-exceed (NTE) limits in the case of Tier I and Tier II in the NO_x Technical Code –. Some of the variability is also likely explained by differences in ship and engine types.

Nevertheless, at high loads above 50 % the (g/kWh) brake specific emissions (g/kWh) converge, even though the data corresponds to many types of ships. This observation is consistent with measurements the partners have done elsewhere, and similarly consistent with the NO_x regulation where higher load points (>50%) carry more weight and thus trend lower to allow the ship to comply with the overall weighted NO_x limits. For this reason, using the remote measurements to monitor compliance with the NO_x limits above 50 % load should also be relatively straightforward for all tiers since any exceedance above the weighted average limit applicable to the individual ship has a high probability of signalling non-compliance.

For loads below 50 %, the picture becomes more complex as can be seen from the data. Here, exceedances above the average NO_x limits are not necessarily indicators of non-compliance since the NO_x Technical Code leaves a lot of room for ships to fuel optimize at low loads. This is illustrated in the data by the spread in factors below 50 %. This means even (very) high NO_x emission factors observed at low loads – well above the applicable weighted average for a ship – cannot automatically be assumed to be a sign of non-compliance. At least this is true for Tiers I and II.

For Tier III the implementation of a NTE limit in the certification test cycles of no more than 50 % above the applicable NO_x emission limit (max 3.4 g/kWh) at <u>any</u> load point means even low loads (<50 %) have a regulatory 'cap' which may be monitored via remote measurements (MEPC.177(58), section 3.1.4). In the case of Tier III, the NTE limit effectively means ships are significantly more restricted in how much they can optimize at low loads without exceeding the NTE limit, and this opens a possibility for effective use of the remote measurements across the full load spectrum, or at least as low as 25 % load which is the lowest load point covered by the NO_x regulation. However, while there is a conceivable path for using the remote measurements to control NO_x emissions for all ships above 50 % load, and for Tier III ships even from 25 % load, such an approach is in all cases highly dependent on the availability of correct SFOC data. Knowing the SFOC value for each ship is crucial for the factor calculation, and even if the IMO 4th GHG Study provides usable guidance on applicable SFOCs, an accurate emissions analysis will depend on individual SFOC data confirmed by the shipping companies. This data point is currently not readily available but could be envisaged reported to the authorities going forward for the purpose of NO_x compliance monitoring.

[3] Weight factors for 75% and 100% load are respectively 50% and 20%. Actual contribution of these points to end result is in practise actually 85%-90% because the g/h NOx mass flows are weighted and these are high for high load points.

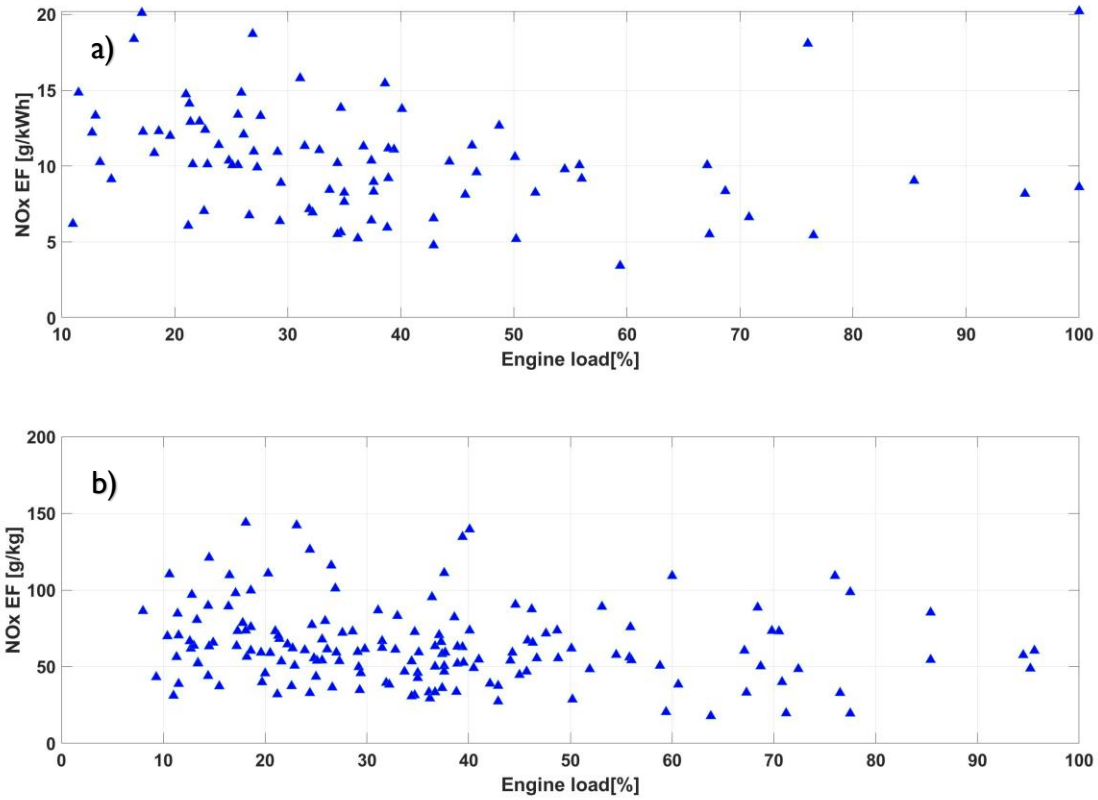

Figure 39. The ship's emissions of NO_x versus load, measured in the Elbe river approx. 10 km from the entrance to Hamburg harbour. a) Break specific mass emission (g/kWh) b) fuel-mass specific emission(g/kg). The data correspond to ships longer than 75 m for the Chalmers NO_x emission data during C3.

4 Current status of particle measurement techniques for ship emission monitoring

Ship emission is a very important anthropogenic source of atmospheric aerosols (e.g. Ausmeel et al., 2019). Nevertheless, there is still no legal regulation in place, limiting particulate emissions from seagoing vessels. However, since particulate emissions have a significant impact on public health and climate effects (Sofiev et al., 2018), it is expected that there will be legislation in the future. In order to better understand the particle emission behaviour of seagoing vessels under real word conditions and to be able to make recommendations for future regulations and monitoring mechanisms, the SCIPPER project has started to measure different properties of particles emitted by seagoing vessels. Section 4.1 provides a general summary of the measurement techniques available for measuring particles in ship plumes. Section 4.2 presents the measurement systems used within SCIPPER. Section 4.3 presents some selected preliminary results and first statistical analyses, and Section 4.4 compares some of the measurement systems used.

4.1 General description

From a monitoring and scientific point of view, the most important particle properties in ship exhaust plumes are the particle size, concentration and chemical composition. However, remote measurement of particulate matter in ship plumes must be relatively fast, since the residence time of the plume at the measurement location can be less than one minute. This limitation significantly reduces the number of suitable measurement systems.
To measure the integral particle number concentration so called **Condensation Particle Counters (CPC)** would be the best choice, because these instruments can measure with a temporal resolution of up to 1 Hz or even faster and therefore can temporally resolve plumes of individual passing ships. To capture the whole particle number concentration the CPC lower detection limit should be below 10 nm (cf. Sec. 3.2.). CPCs are available from different manufactures.
There are several methods for the rapid measurement of particle size. **Mobility spectrometers** are suitable for very small particles starting from a size of 6 nm. With these instruments, the particles are sorted to size classes by means of deflection in an electric field and their number is subsequently counted. **Optical methods** can be used to size and count particles with size starting at 90 nm up to several microns with 1 Hz resolution. Here the particles to be measured pass a light path where the scattered light from the particles is detected. Mobility spectrometers, as well as optical detectors are offered by different manufactures, which can significantly differ in detection limits and temporal resolution. It is noted that scanning mobility spectrometers (SMPS) may be too slow for full detection of ship plumes at fixed measurement sites, because a scan over the whole size range from below 10 up to 500 nm may take several minutes while the ship plume can pass by within less than one minute.
The best method to measure the particle chemical composition in individual plumes of bypassing ships is the use of an online **aerosol mass spectrometer (AMS)**. Here particles are sized and thereafter evaporated and ionized by different methods, depending on the type of AMS, to detect their chemical components by time-of-flight mass spectrometry. AMS systems are available from several suppliers. However, due to the complex analysis method they are very costly. If the user is mainly interested in the soot fraction of individual ship plumes a **soot photometer (Aethalometer)** can be used. This instrument collects aerosol particles on a filter tape and analyses the transmission of light through the filter tape, for both a loaded and an unloaded piece of tape, using different optical wavelengths. From the rate of change of the attenuation of light the concentration of the optically absorbing aerosol is determined.

4.2 Instrumentation

Within the SCIPPER project the consortium partner BSH operates two combined particle size spectrometers at their measurement sites in Wedel (approach to Hamburg harbour) and the Kiel fjord (approach to Kiel and Kiel Canal). The measurements began in Wedel in September 2020 and in Kiel in June 2021 and are ongoing at both sites. Both combined particle size spectrometers are composed of an mobility size spectrometer and an optical size spectrometer and are designed identical. For the mobility size spectrometers Fast Mobility Particle Sizers (FMPS, TSI Model 3091) were chosen to measure the particle size distribution with 32 logarithmic equidistant size channels in the particle diameter size range 5.6 nm < dp < 560 nm. For the optical spectrometers two Optical Particle Sizer (OPS, TSI Model 3330) were selected to measure the particle size distribution with 16 logarithmic equidistant size channels in the size range 0.3 µm < dp < 10 µm. Both systems can measure the full-size range with 1 Hz temporal

resolution. This means every second from each combined system a particle size distribution from 5.6 nm to 10 µm with a 48-size channel resolution is available. The setup is shown in Figure 40. The instruments are installed in air-conditioned shelters so that they can be operated at constant ambient conditions, and independent of the weather. To ensure comparable measurements at low relative humidity, the sample air is dried by using Nafion counterflow dryers. Particles larger than 10 µm are excluded by the PM-10 inlet head.

Figure 40. Combined particle size spectrometer (TSI Fast Mobility Particle Sizer + TSI Optical Particle Sizer) installed into a climate-controlled shelter. Sample air is dried with a Nafion counterflow dryer. The PM-10 inlet prevents particles larger than 10 µm to enter the system.

During the SCIPPER C3 measurement campaign, carried out from September to October 2020 in Wedel (cf. Sec. 4.1), both combined particle spectrometers were operated side-by-side to measure the same ship plumes under the same conditions. During this campaign also the project partner Chalmers operated a very similar system as BSH, consisting of an Engine Exhaust Particle Sizer (EEPS, TSI model 3090) and an OPS (TSI Model 3330), but without a dryer at the Chalmers system. Figure 41 shows that the agreement between the two BSH systems with dryer (left figure), but also the agreement between the BSH and the Chalmers system (right figure) was found to be reasonably good. The integral particle mass found by the Chalmers system in the individual ship plumes was little higher than the particle mass concentration reported from the BSH system for the same plume (slope 0.97 instead of 1.0), but this can be easily explained by the remaining humidity onto the particles measured with the Chalmers system, whereas the BSH system measured dried particles.

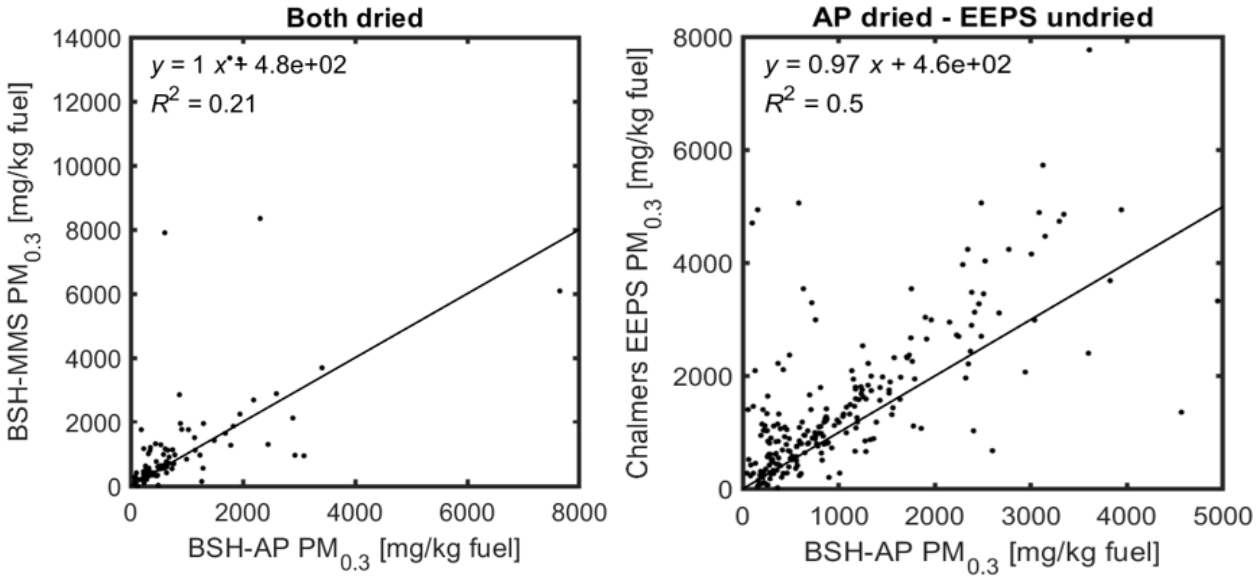

Figure 41. Comparison of particle mass concentrations measured in ship plumes with different combined particle size spectrometers operated by BSH (TSI Fast Mobility Particle Sizer + TSI Optical Particle Counter) and Chalmers University (TSI Engine Exhaust Particle Sizer + TSI Optical Particle Counter). The left graph shows the comparison of the two identical systems operated by BSH during SCIPPER C5. The right graph compares one of the BSH systems with the very similar one (EEPS instead of FMPS) operated by Chalmers also during C3 in Wedel, but without Nafion counterflow dryer.

Besides the particle size distribution, Chalmers additionally measured soot with an Aethalometer (AE33 from Magee Scientific). Also, soot was measured without a dryer between inlet and the AE33 and no size-selective inlet was employed. During the C3 campaign in Wedel the project partner TNO measured the particle size distribution with a LAS-X II, which measures particle size distribution between 90 and 10 000 nm in 99 channels. Besides the particle size distribution, the number concentration of UFP particles was measured by TNO with an EPC (Model 3783 of TSI). It is similar to an CPC, but uses water instead of butanol as condensation liquid. The EPC measure the number concentration of particles with a diameter of 7 nm or larger. A Nafion dryer was deployed to dry the aerosols before the LAS-X II and the EPC measured the particles. The same Aethalometer (AE33), as Chalmers University, was employed by TNO. However, TNO employed a PM2.5 selective inlet and a Sample Stream Dryer (PN M5610 also from Magee Scientific) during the C3 campaign.

4.3 Preliminary results on particle emissions

Figure 42 depicts, as an example, the particle number concentration (a), particle size distribution (b) and the soot (black carbon) concentrations (c) on September 29, 2020 during the C3 campaign in Wedel. In the displayed 45-minute time window seven vessels passed the measurement site in a distance of about 400 to 800 m. The plumes of those vessels are clearly distinguishable and could be allocated to the individual ships (cf. assignment in Figure 42 (b)). The integral particle number concentration in the plumes reached values up to 125 000 particles/cm³. Of course, the measured total concentration strongly depends not only on the total number of emitted particles, but also strongly on the dilution of the plume between the stack and the measurement site. By normalizing the measured particle concentration, with the in parallel measured CO_2 concentration, a particle emission factor in particle number per kg of burned fuel can be calculated and will be discussed more detailed with Figure 44. From Figure 42 (b) it is obvious that it is possible to identify also differences in the particle size for different ships. Those differences are probably caused by different types of engines, different fuels and different operational loads and will be investigated further in later studies. From Figure 42 (c) it can be seen that the measured plumes differ also in their amount of soot. At times the soot peak is less pronounced than that of the particle number concentration. Interestingly, there are also instances where the soot peak is more pronounced than in the integral particle number concentration (see for example the plume around 14:09 UTC). Also this interesting finding will be investigated further in future studies.

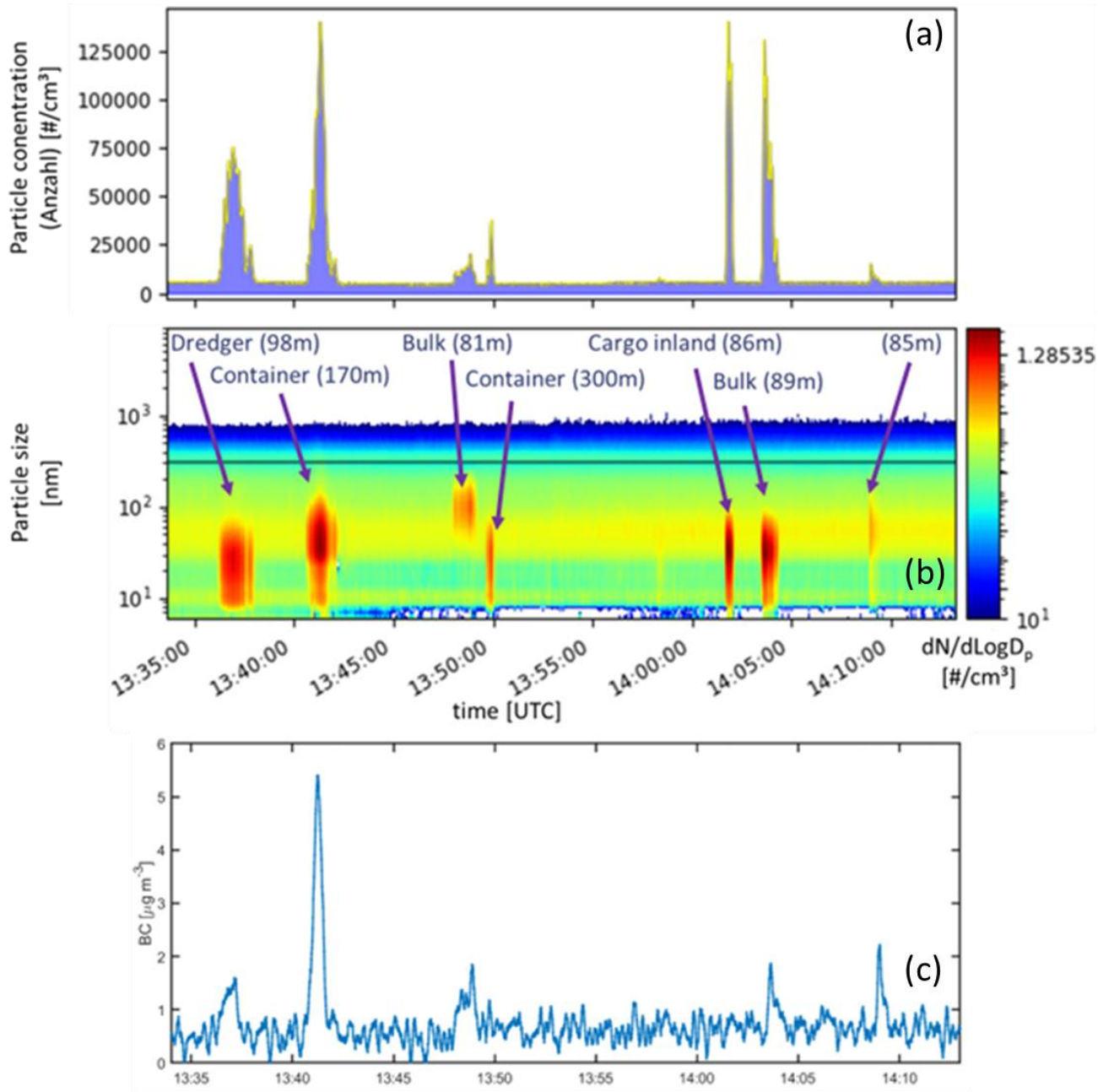

Figure 42. Time series of the number concentration of particle (a), particle size distribution (b) and soot concentration (c), also known as black carbon, BC) on 29th of September 2020 measured in Wedel, Germany during the C3 campaign.

From Figure 42 (b) it is visible that the majority of the emitted particles have a size below 100 nm. This finding was investigated further. Therefore, Figure 43 depicts the cumulative fraction of the total number- and mass concentration of particles for an average of all ship plumes measured during C3 in Wedel. The magenta dashed line indicates that 50% of the measured particles (number) had a size smaller than 40 nm in diameter and 90% were smaller than 80 nm (orange dashed line). Looking at the mass of the particles a different pattern occurs, since the bigger particles weigh more than the smaller particles. Hence, particles with a size up to 40 nm contribute only by about 5% to the total particle mass while 75% of the integral particle mass comes from particles smaller than 200 nm. The finding indicates which particle size range should be taken into account in order to determine the number- or mass concentration of particles in ship plumes. For particle number concentration measurements, a lower cut off of 10 nm would be sufficient, because less than 5% of the particle number are below this size. For integral particle mass measurements, a lower detection limit of below 40 nm would be sufficient and the chosen instrument should measure at least up to a size of 4 000 nm. Both the BSH and the Chalmers system fall within this range. The TNO LAS-X II system has a lower cut off of 90 nm and thus misses around 45% of the particle mass of the ship plumes.

For the integral number concentration of particles the TNO LAS-X II system is not suited, since it misses more than 90% of the particle number (which was the reason for operating an EPC in parallel to measure the total particle number concentration).

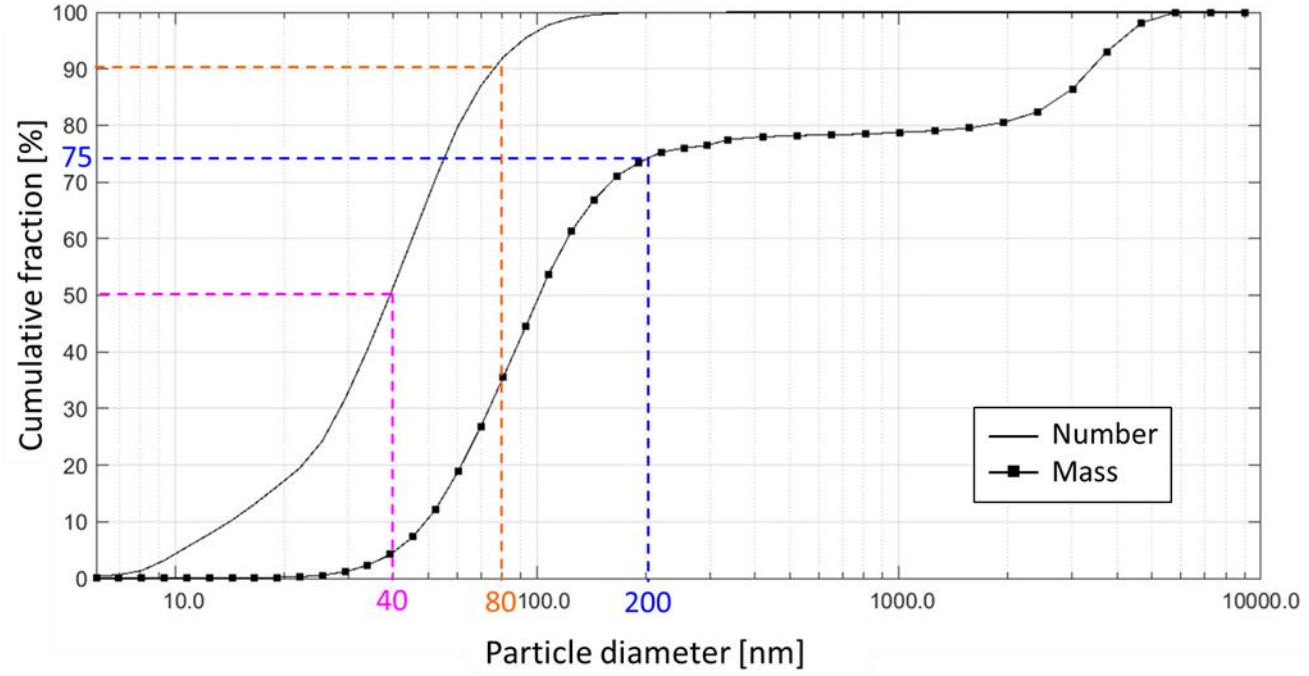

Figure 43. Cumulative fraction (in %) according to particle diameter for particle number and mass.

From Figure 43, it can be further deduced that efficient reduction of particulate emissions from ships is only effective if the very small particles are considered. For example, effective mitigation of particulate mass emissions would require regulations of particles larger than about 50 nm. For an effective reduction in the number of emitted particles, even particles as small than approx. 20 nm would have to be taken into account.

As stated in the beginning of section 4.3, it is possible to calculate plume individual particle emission factors when normalizing the measured particle concentration with the plume CO_2 concentration, measured as well. Figure 44 compares particle emission factors in the unit particles per kg of burned fuel for different ship classes and sizes. A total number of 6 110 analysed ship plumes, which were measured between September 2020 and April 2022 at the two BSH measurement sites in Wedel and Kiel were used for this plot. For seagoing vessels, total particle emission factors were found to be in the range from $0.8 \cdot 10^{16}$ to $1.5 \cdot 10^{16}$ #/kg_{fuel} on average. The highest average emission factors of $1.625 \cdot 10^{16}$ #/kg_{fuel} were found for dredgers, small cargo ships, and medium-sized tankers. It is interesting to see that there appears to be a dependence of the particle emission factor on ship size for cargo ships (decreasing trend for bars from light red to dark red). For cargo ships with a length of less than 100 meters, emission factors were found to be 60 % higher on average than for cargo ships with a length of more than 300 meters (mostly container ships). The emission factors found for inland vessels were on average four times lower than for the small seagoing cargo ships. Particle emission factors for seagoing passenger vessels were found to be on average at the same level as for the biggest cargo vessels and small tankers.
The found particle emission factors in Wedel and Kiel are of the same order of magnitude compared to what SCIPPER D4.1 has developed by exploiting in-funnel measurements accumulated from a significant number of literature sources. Specifically, an estimation of emission factors at 50% engine load provided values at the range of $2.3 \cdot 10^{16}$ #/kg_{fuel}. This same order of magnitude observed validates the findings of remote instrumentation. However, as the particle emission factors of individual ships do strongly depend on the quality of the used fuel, as well as on the type and operational mode of the ship's engine, further investigation is needed to explain the found differences between different ship types and sizes in detail.

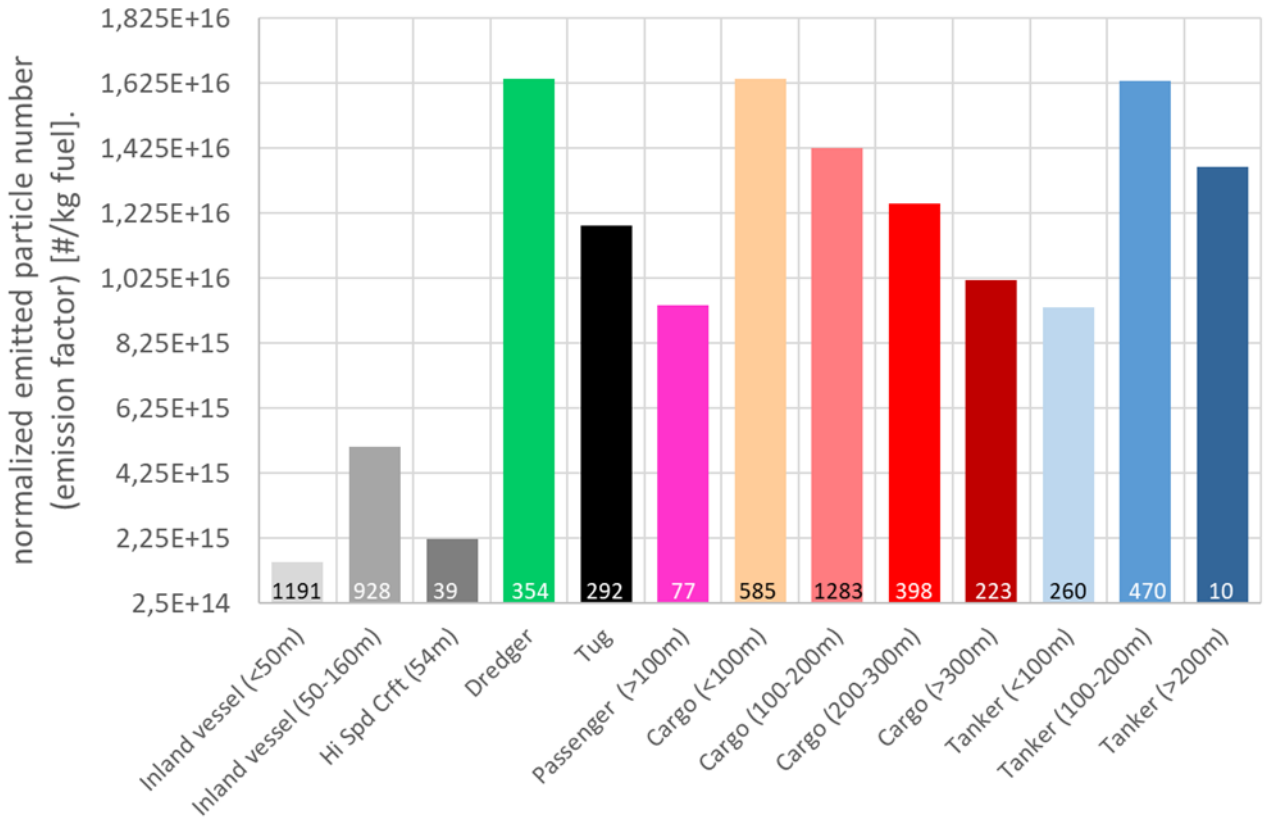

Figure 44. Integral particle emission factors for different ship types, measured between September 2020 and April 2022 at the two BSH measurement sites in Wedel (Hamburg approach) and Kiel (approach to Kiel and Kiel Canal).

4.4 Further Comparison of measured particle number- and mass emission factors from ship plumes measured during C3 in Wedel

In addition to Figure 41 for all the different particle size ranges the ships emission factors were calculated both for mass (mg/kg fuel) and number (#/kg fuel) for all sizing systems, introduced in section 4.2. Unlike the FSC, where fuel samples can be taken, actual particulate emission factors from selected ships are not known. For comparison the BSH AP system was chosen as the reference. Figure 45 depicts the emission of particle mass in the range from 0.09 to 0.3 µm (given in mg/kg fuel), outliers below 0 and above 2000 mg/kg fuel were removed. From this figure it is clear that the two BSH systems compare very well to each other with little scatter (R^2 of 0.83). The estimates of the emission of $PM_{0.09\text{-}0.3}$ of the TNO LAS-X II system also is comparable with values close to the one-to-one line leading to an R^2 of 0.65. There is more scatter between the Chalmers system and the BSH-AP, with an R^2 of 0.35. The Chalmers system also, in general, overestimates the particle mass emission factors of $PM_{0.09\text{-}0.3}$ compared to the other systems.

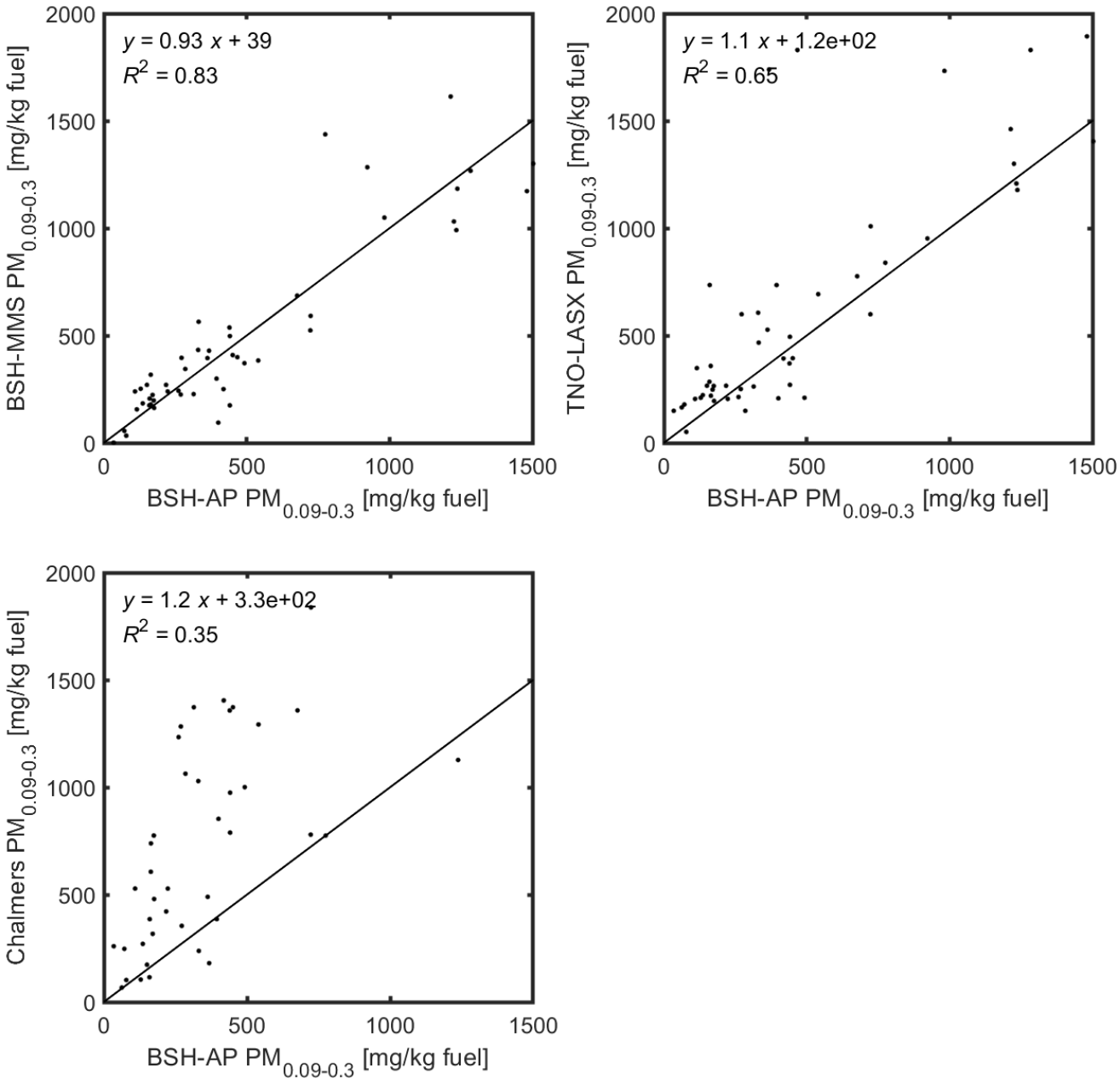

Figure 45. Scatterplot of the particle mass emissions measured during the C3 campaign in Wedel, Germany in a size range of 0.09 and 0.3 µm. The BSH-AP system is taken as a reference on the x-axis and on the y-axis the BSH-MMS (left top), TNO-LASX II system (right top) and Chalmers (left bottom) are plotted. The plots include the one-to-one line and the corresponding regression statistics.

For the same size range (0.09 – 0.3 µm) the emissions of the number of particles per kg fuel are given in Figure 46, values above $3 \cdot 10^{15}$ #/kg_{fuel} were considered outliers and removes. The agreement between the different measurement systems is better for the total number concentration of particles compared to the particle mass. The scatter between the two BSH-systems becomes even lower (with an R^2 of 0.88). Also the emissions of the TNO LASX-II system has less scatter with the BSH-AP system (an R^2 of 0.69). The scatter between the Chalmers EEPS system with the BSH-AP system is somewhat higher than that of the TNO LASX-II system (R^2 of 0.55 and 0.69, respectively). The measured particle number concentrations compare fairly well between the different systems. The values of the TNO LASX-II system is somewhat lower than the other systems.

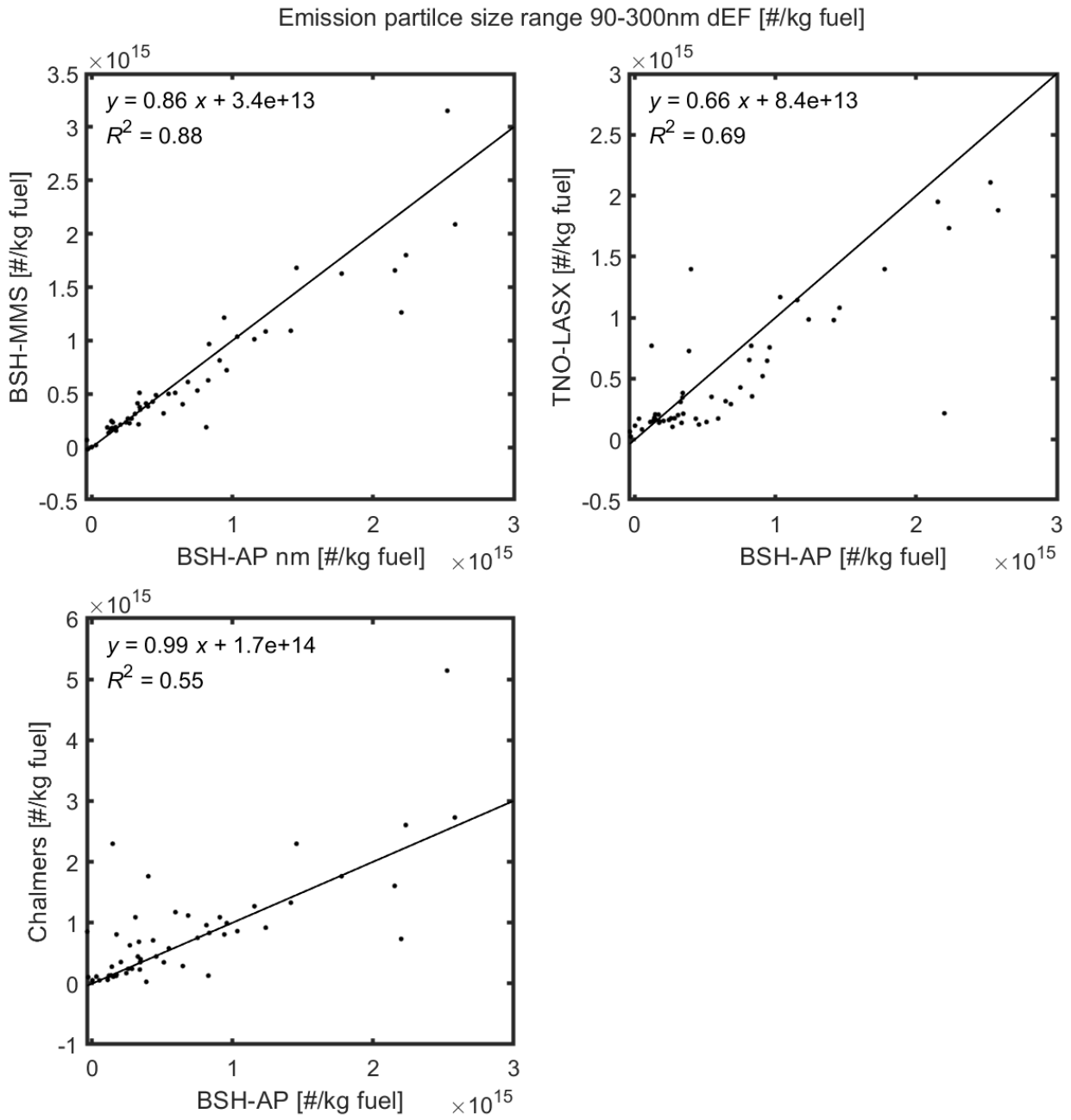

Figure 46. Scatterplot of the emission of the particle number measured during the C3 campaign in Wedel, Germany in a size range of 0.09 and 0.3 µm. The BSH-AP system is taken as a reference on the x-axis and on the y-axis the BSH-MMS (left top), TNO-LASX II system (right top), and the Chalmers-EEPS (left bottom) are plotted. The plots include the one-to-one line and the corresponding regression statistics.

The tendency is that the scatter in emission estimates of particles among systems become lager (i.e. lower R^2-values) when larger aerosol sizes are taken into account. Why this is the case can be explained by considering Figure 43. The number concentration of particles in ship plumes is clearly dominated by relatively small particle sizes (<100 nm). This means that the larger particles occur much less frequent. That means that statistically, there is less change to observe a larger particle. The systems in Wedel were set up besides each other (spaced up to 15 m apart) with separate inlets. This means that not exactly the same air parcel was sampled, but rather air parcels emitted from the same source. Given the lower number of particles that are large particles (>200 nm) it becomes statically less likely to sample the same number of particles. This explains why there is more scatter between the systems for larger size ranges.

5 Recommendations for harmonized ship emission monitoring

This chapter provides the key recommendations for conducting and reporting quality-assured remote measurements of gas emissions from ships. This includes in detail the topics of measurements, data validation procedures, uncertainty calculation and reporting. These recommendations are based on the common activities of the participating groups in the SCIPPER project. The focus is on sniffer instruments for quantitative remote measurements of FSC and fuel-mass specific NO_x emissions. Optical remote sensing systems are not considered since they are less mature and have more uncertain quantification. Fixed, shipborne and drone assisted measurements are covered. In addition, remarks are made on aircraft measurements using standard instruments since the partners have previous experience with such systems (Beecken et al., 2014 and 2015).

5.1 Measurements

5.1.1 Physical system

5.1.1.1 Parameters measured

When performing sniffer measurements, the relevant pollutants (SO_2, CO_2 and NO_x) should be measured in the plume with a sampling rate sufficiently (discussed in more detail in section 5.1.2.2) for capturing the variations of the plume and distinguish it from the background. It is advisable to make complementary measurements of cross-interfering species and ambient parameters such as air temperature and humidity. In order to identify which ships the measured plumes originate from, one should employ an AIS receiver. Furthermore, a windmeter has to be installed in an undisturbed location measuring both the wind speed and direction. For airborne measurement from drones and larger aircraft, wind information is not required since the identification of the ships can be done manually by the flight operator. For a fixed station the concentration data from the individual instruments needs to be collected in near real-time and saved together with time stamp and wind and the ships' identity and navigational information.

5.1.1.2 Sampling system

A sampling system must be used for the sniffer systems. Practically this means that the air is sampled through a rain-protected inlet and a PFA or PTFE tube (preferably short) towards the instruments by using a pump. A note of warning against using aluminium, since it affects the measurement of the SO_2 concentration.

For dirt protection, a filter is generally used to prevent larger particles (> 100 µm) from entering the optical cavities of the instruments. Since dirt and moisture accumulates in or on this filter, it needs to be replaced on a regular interval. The recommendation is 2 –4 weeks, depending on the particle pollution rate. The same is true if filters for optical SO_2 sensors are used for drones. For EC sensors on drones a filter is no necessity, due to the low pump flow and chemical sensing without optical cavity. It is advised to heat the inlet system and filter, as a precaution to remove possible condensation effects and adsorption of SO_2.

For manned airborne measurements, using the standard systems, the inlet should be perpendicular to the airplane flight direction, to minimize pressure variations and particles from being trapped in the sampling system. However, for particle measurements it is necessary to use an isokinetic inlet.

Fixed land-based sniffers should preferably be placed in less polluted areas with low background variability for the gases of interest to separate ship plumes from other pollution sources. For fixed sniffers it is also advisable to put the station on the predominant down-wind side of the shipping lane, in order to be able to measure the highest amount of ship plumes.

5.1.2 Performance requirements

The required performance for the instruments to be able to monitor remote ship emission measurements using different platforms, e.g. fixed sites, measurement vessels, aircraft and drones, are described below.

5.1.2.1 Detection limit

The detection limit for a single concentration measurement is usually defined as 3 times the standard deviation (3σ) of the instrument noise level and this is assessed from the measurement precision in clean background air or

when measuring on a reference zero gas. Note, that if several concentration measurements are averaged then the concentration detection limit is lowered by the inverse of the square root of the number of samples ($1/\sqrt{N}$).
By experience, for measurements from a fixed station, measurement vessel or manned aircraft for measurements of a few hundreds of meters away, the sensors need a detection limit of around 2 ppb for SO_2, NO_x and NO and 200 ppb for CO_2 at ambient levels. In this way the FSC can be measured with a 1σ precision of 0.04 % S m/m, as shown in section 3.2.1 for the standard instruments. However, this strongly depends on the individual signal to noise ratio during the measurement. The tested laser system has higher precision (0.02 FSC %) than the standard sniffers (cf. Sec. 3.2.1). It is therefore better in determining small exceedances at the FSC limit of 0.10 FSC % S m/m in situations with small plumes and peak heights. For measurements closer to the exhaust stack of the ships, using a drone or a helicopter, the required detection limits are higher, i.e. 10-20 ppb for SO_2, NO_x and NO and 2-5 ppm for CO_2 above background. This can be performed by the EC sensors described in section 2.1 and a small NDIR sensor for CO_2. In this way a 1σ precision of 0.025-0.04 % S m/m can be obtained, as demonstrated in section 3.2.1.

5.1.2.2 Response and sampling time

The response time needs to be fast enough to capture the variations of the plume and distinguish it from the background, i.e. typically 1s to 30 s response time, depending on application. One should distinguish between sampling rate, which is the data output frequency of the instruments (typically every 1 s) and the response time, which is the time it takes for the instruments to change from 10 % to 90 % of full value when the gas concentration changes quickly. The standard and laser instruments are all based on flow-through-cells in which the gas is measured using light at different wavelengths. The response time is typically the time it takes to exchange the air in the cavity. One exception here is the UV-fluorescence instrument, used in the standard systems on fixed stations, in which the diffusion tube, a hydrocarbon-kicker to avoid influence by VOCs, determines the response time.
In the fixed remote sniffer measurements application and manned fixed-wing-aircraft ones, the plumes are measured at 0.5 to 1 km from the ship's stacks. In the fixed application the plumes are present at the measurement site for 20 s to several minutes while in an airborne application it may only be present for a few seconds. Note, that even though the time in the plume is shorter than the response time, all the gas is actually measured in a flow-through system. To account for differences in response time, the detected plume peak areas are calculated for all species and then the ratio against the corresponding value for CO_2 is calculated according to Eq.1, Eq. 2 and Eq. 6.
For the EC sensors, used in the drone application, the principle is somewhat different. Here the gas species diffuse through a membrane at a varying speed which is proportional to the concentration, and this typically takes between 20 s and 30 s for measurements of the levels occurring in ship plumes. If the measurement is shorter than the response time or if the concentration is varying strongly, all the gas will not enter the detector and an uncertainty in the actual concentrations will arise. Such measurements should be removed by data quality control procedures as described in section 5.1.3.

5.1.2.3 Calibration and cross-interference to other species

In order to make the instrument response traceable, the instruments will be calibrated by comparing them to a certified, independent gas standard.. This should ideally be done in the field to make sure that the ambient conditions, i.e. temperature, background air mixture, humidity, plume concentration levels, plume mixture, are the same as for the actual measurements.
The required calibration procedure, including frequency of calibrations, depends on instrument characteristics and the conditions of the measurement site. It should be demonstrated by documented testing for each instrument system that the calibration procedure is appropriate. Ideally the effect of the procedure should be included in the measurement uncertainty calculation. For instance, if calibrations are done in the laboratory only, or done infrequently, there needs to be documented tests showing that this is appropriate.

It was found in this project that certified gas standards from gas suppliers are not always trustworthy, given that deviations of 60% were found. Therefore, it is recommended to check new gas standards against old ones and to perform round-robin tests of calibration gases, once-in-a-while, with other groups.
For the standard sniffer systems, and laser systems, the calibration can be done by replacing the inlet air by premixed gas of known concentrations. This can be obtained by acquiring a gas cylinder with calibration gas from

an accredited gas supplier with low levels of the key species, e.g. 300 ppb, mixed in dry air or nitrogen, or higher concentration gas, e.g. 80 ppm, diluted to appropriate levels using a gas mixer during the calibration procedure. This is the most common calibration method used by the groups in this study, but it has the disadvantage that the mixture is dry and hence does not correspond to the surrounding air.

The sensors can be affected by changes in humidity in two ways; either in the physical detection process or the fact that species could stick more easily to surfaces at higher relative humidity, leading to loss of substance or delayed response time. For instance, the highly sensitive laser system does not have any spectroscopic cross interference to water vapor, but it has many surfaces onto which adsorption could occur at high relative humidity. For the SO_2 UV fluorescence system, used in the standard sniffer system, it´s response is affected by water vapour since the fluorescence becomes lower when quenching from water vapour increases at high RH. This response change appears to be relatively small when switching between dry zero gas and clean air, but it may cause an impact when SO_2 is present in the gas. Since the standard sniffers are calibrated with dry gas this may be a reason why they appear to measure too low FSC values even at low relative humidity. This is in contrast with the highly sensitive system which is only affected by adsorption. This needs to be studied further. One problem with using premixed air at very low levels (typically 100-300 ppb), is the fact that they have limited lifetime and that sticky gas (SO_2) could get adsorbed in the measurement system. This was observed during C3, for low concentration levels of the calibration gas (around 100 ppb).

An optional way to calibrate the standard system is to inject a controlled amount of high concentration calibration gas into the inlet of the system, using mass flow controllers. In this way the calibration air has the same humidity and other properties as the background air. However, this results in a more complex system, and this puts high demands on maintaining the quality of the injection equipment over time.

The drone-based sensors, EC and NDIR, can be calibrated in the same manner as the standard systems, but these systems appear to be more sensitive to pressure variations that may arise when using calibration gases, due to the low flows and relatively weak pumps. The EC sensors are also rather sensitive to humidity changes and the use of dry calibration gases should be done with care. The EC sensor's humidity response depends on the component's relative humidity history, so component selection, correct storing conditions and possibly regular testing are important.

It is also important to assess how linear the measured response is versus actual concentration for all systems. The standard sniffers and laser technique generally have linear response against concentration, due to the physical detection principle. This is not the case for the NDIR technique, used to measure CO_2. NDIR has a rather nonlinear response but which the manufacturers are compensating for by software algorithms. It is important when using such a system to calibrate at least at two concentrations, like background values, e.g. 420 ppm and ship plume values, e.g. 450 ppm, respectively, to minimize potential non-linearity problems.

Most instruments are generally sensitive to other species than the key species they are designed for called cross-sensitivity. This is described in section 2.1, Table 1, for the different instruments used in SCIPPER, as well as in section 2.2 for the individual partners. It is recommended to correct the instrument readings for the known cross-sensitivity, Table 1**.** It should be noted that cross-interference is not necessarily linear in response and could depend on the relative humidity. For instruments with significant cross-interference, it is advised that the cross-interference is treated in the same manner as for normal calibration. Therefore, it is necessary to document, based on testing, how and how often such cross-interference calibrations are performed and included in the measurement uncertainty estimates.

5.1.3 Quality control

The objective is to validate normal performances of the used sensors and whether any drift has occurred. This can be done in multiple ways, as demonstrated by the SCIPPER groups:

a. Frequent calibrations, as discussed in section 5.1.2.3.
b. The usage of permeation tubes to provide specific span concentrations of SO_2 and NO_2 to the instruments. If the measured concentration values from the permeation tubes stay within a certain range, the quality of

the measurement is flagged as high. Note that the permeation devices are less suitable for calibration since they may drift, due to their strong dependence on temperature.

c. The usage of two independent systems. When the systems show significant differences, the quality is flagged as poor.
d. Puffing of high concentrated gas mixtures of SO_2 and CO_2 in front of the instrument inlet. For instance, a gas mixture of 100 ppm SO_2 and 23 % CO_2 in front of the instrument inlet correspond to the same ratio as a 0.10 % S m/m FSC ship. This can be repeated regularly to test the instrument response in different conditions and potentially correct for artifact die to the ambient conditions.
e. Traceability of sensor lifetime. The lifetime of an EC sensor is limited by a threshold in its total exposure (ppm-h) and secondly it decays in ambient air due to effects of oxygen and humidity. The typical behaviour of a certain type of detector can be investigated by long duration and stress tests. To investigate an individual detector, parallel measurements with multiple detectors can be used as described in bullet (c) above.

5.1.4 Measurement methodology and calculation of FSC and fuel specific emission factors

The calculation of FSC should generally be obtained from the elevated SO_2 concentration values in the plume relative to background, normalized against the corresponding values for CO_2, and integrated across the full plume, as indicated Eq. 1 in section 2. This is applicable for standard sniffers and the drone-based EC sensors. For the latter sensors, several of the SCIPPER groups are discarding both low and variable concentration values. An additional difficulty lies in the relatively strong relative humidity response of the sensors and the fact that this may change in the plume, and this should be thoroughly tested. The fuel specific emission factors of NO_x should be calculated according to Eq. 2 in section 2, respectively.

5.2 Data validation procedures

The data quality criteria correspond to the measurement conditions that have to be fulfilled for valid measurements. These criteria are empirically derived and generally relate to the uncertainty of the measurement, or they state whether the measurement is valid. An example of the latter is ship plume allocation. The criteria can impact the uncertainty of the results *directly*, such as the signal-to-noise level, or *indirectly*, such as the wind speed. The latter also has an impact on the signal-to-noise ratio since it will affect the measured concentration levels. There is generally redundancy in the quality criteria, meaning several parameters will have a similar impact. In Table 7 the data quality criteria used by the different SCIPPER groups are described. Note that this is not a comprehensive list since for some groups the criteria may already be included in the uncertainty calculation.
In the case where the measurement uncertainty is calculated explicitly for each measurement (see section 5.3.1), the *direct* quality criteria should instead be included in the uncertainty calculation, and they can therefore be omitted as a data quality criterion. This is not the case when the validity of the measurements more heavily relies on the quality criteria (see section 5.3.1). Note that there are many groups, also outside SCIPPER (cf. SCIPPER deliverable D2.1 by Beecken et al, 2019) that include data quality in their reporting, e.g. as a number (0-10), text (poor medium, high) or colour (green, orange, red). However, it is advised not to report data quality since it has ambiguous units which is hard to compare between different instruments. Instead, only measurements above an individually defined quality threshold should be reported together with the associated uncertainty.

Table 7. Quality criteria that are used by the individual groups on fixed stations (BSH, Chalmers, TNO) and drone-based mini sniffers (Aeromon, Chalmers, Explicit) in the data validation procedures.

Criteria	Criteria type	Platform	Group	Comment
Instrument Validation	Indirect	Fixed and drone	All	Validation that the instruments do not drift and work properly
Absolute Calibration	Direct	Fixed and drone	All	Calibration of absolute response and linearity for key species*

Cross interference calibration	Direct	Fixed and drone	BSH (only SO_2 to NO)	Calibration of absolute response and linearity for interfering species*
S/N CO_2	Direct	Fixed and drone	All	STD of background relative to peak height
Peak height CO_2	Direct	Fixed and drone	All	
Peak area CO_2	Direct	Fixed and drone	Aeromon Chalmers, Explicit, TNO	
Peak height of NO	Indirect	Fixed and drone	BSH, Explicit, TNO	Usually the strongest indicator for a plume and useful for plume allocation
Time duration of plume	Indirect	Drone and Fixed station	Aeromon Chalmers, Explicit, TNO	Sampling: sufficiently long compared to plume instrument response time, but not too long to ensure stable background*.
Wind direction	Indirect	Fixed	Chalmers, BSH, TNO	Affects the length of the detected plume and peak height
Wind speed	Indirect		Chalmers	Affects the peak height of gas concentrations
S/N SO_2	Direct	Fixed	BSH, TNO	Only in case of non-compliance: If above 3 STD of background noise level
Plume allocation	Indirect	Fixed	All	Includes - wind speed, wind direction, number of ships, plume vector, distance
Signal to noise ratio NO	Direct	Drone and fixed	BSH, Explicit Chalmers	Signal to noise ratio NO, if used for SO_2 correction and only in case of non-compliance
Plume shape (stability, ups and downs)	Indirect	Drone	Explicit	Aeromon includes this in result uncertainty analysis
integrated plume area NO	Direct	Drone	Explicit	

*The required period between calibrations needs to be underpinned with long term testing of system. The groups have different procedures and calibration frequencies within field or laboratory calibration.

5.3 Calculation of uncertainty

In this section some of the basics of uncertainty calculation of measurements is introduced. The implementation by the SCIPPER partners is then summarized in this context.

5.3.1 Basic introduction

The uncertainty of a measurement can be divided into a random and a systematic part.
For any measurement consisting of several observations, it is possible to calculate the sample mean and the standard deviation. The standard deviation is a measure of the variability of the observations about the measured mean value. The standard deviation encapsulates a number of sources of variability occurring during the measurement, including:

- variability in the components of the observation (SO_2, CO_2 and NO_x) with respect to their mean values used in the derivation of the observation (random measurement uncertainty);
- variability in the final quantity (parameter) being measured, i.e. FSC or fuel-mass specific NO_x emission (g/kg).

Random uncertainty can be decreased by increasing the number of measurements (repetition of the same measurement). Systematic uncertainties are difficult to measure and in general they are estimated by identifying inputs to the measurement that might produce systematic effects and quantifying these through sensitivity analysis. This could for instance include spectroscopic calibration factors, uncertainties in calibration gas concentration and influence of relative humidity.

It is also possible to verify the measurements and their uncertainties through the use of validation experiments as done in the validation campaigns in SCIPPER. In SCIPPER the measurements errors were obtained and divided into systematic and random parts.

The uncertainty analysis can be divided into two types following the recommendations in GUM (Guide to the Expression of Uncertainty in Measurement, ISO/IEC 2008). In the type A approach, the random uncertainty for each observation is assessed from the measurement itself and then the systematic uncertainties are added to calculate the full measurements uncertainty. In the type B approach, the uncertainty is based on available information and assumptions about the typical random uncertainty. For instance, data obtained from other laboratory and field tests, combined with making certain that the data fulfil quality criteria.

In the type A approach, the standard uncertainty of the mean for each measurement is calculated from the standard deviation by dividing it by the square root of the number of observations n. The standard uncertainty does include any variation in their impact on the specific observation set. However, it does not include the effect of systematic effects that are constant during the recording period of the n observations.

When adding a number of different measurements n, the uncertainty of the calculated averaged value is the square root of the sum of squares of the individual measurement standard uncertainties (Gaussian error propagation). It can also be shown that for a product or a ratio the resulting relative uncertainty corresponds to the square root of the sum of squares of the individual relative uncertainties (see 2.3.1.3).

The expanded uncertainty is the standard uncertainty, absolute or relative, multiplied by a coverage factor k. The expanded uncertainty is used to define a confidence interval for a mean. The value of the coverage factor depends upon the distribution of values being averaged and on the value of confidence desired. Here we have use CI 95 %. In this document it is assumed that values being averaged follow a Students t-distribution with a certain number of degrees of freedom. For a single measurement the number of degrees of freedom to be used is (n - 1). The k value, which can be found in a lookup table, corresponds to 12.7 for one measurement. It approaches the value of 1.96 (Normal distribution) when the number of values being averaged increases.).

Furthermore, to obtain the total uncertainty, the random and systematic uncertainties are added in quadrature and then the square root is calculated. This is also the case when the systematic uncertainty corresponds to a measurement bias. For instance, if the measurements are always 10 % low due to loss of species in tubing. The random and mean value of the bias plus the uncertainty of the bias is the added in quadrature (Magnusson et al. 2008, RSSU table 3). Neglecting the uncertainty of the bias yields Eq. 26 below. Here it is assumed that the average bias ($\bar{U}_{bias}$) has rectangular probability distribution rather than a gaussian and therefore it is divided with $\sqrt{3}$ (ISO/IEC, 2008). In this study we have used Eq. 26 by adding the systematic (bias) and random measurements errors that were determined in the validation campaigns.

$$U_{tot} = k\sqrt{\left(U_{Random}(x)^2 + \left(\frac{\bar{U}_{bias}}{\sqrt{3}}\right)^2\right)} \qquad \textbf{Eq. 26}$$

5.3.2 Implementation

The different groups in SCIPPER have agreed to recommend using the GUM type A approach for standard sniffers at fixed stations and either the type A or B approach for drones. The uncertainties should correspond to the 95 % confidence level.

The principles of the GUM type A or B approach are used already today but interpreted and implemented slightly differently, see section 2 and Table 8. It was not possible during SCIPPER to fully agree on a common approach and to give completely harmonized recommendations. An overall summary on how the individual groups have done their uncertainty calculation and which parameters have been considered is given in Table 8. This table additionally gives an overview of the differences between the groups and indicates what to do to further harmonize the uncertainty analysis.

The objective in the harmonization work was to assess the reported uncertainties in validation experiments carried out in the project. In campaign 3, the reported uncertainties were compared to the actual measurement error, obtained from the deviation between sniffer measurement and the specifications based on on-board fuel samples (analysed in a certified laboratories) or bunker delivery notes (kindly provided by the ship's operators). The results of this work in section 3.2.1 shows that the sniffer measurements at the fixed station have a negative bias, which is apparently not considered as part of the uncertainty analysis.
With the strong bias present, it is difficult to assess how well the uncertainty estimates of the different stationary measurements explain the observed random variability. This requires further work. For the drone measurements there was no apparent bias, and the individual reported uncertainty could therefore be better compared to actual measurement error. Explicit, using a type B approach, were able to explain their deviation in 64 % of their measurements and Chalmers (drone), using a type A approach, in 36 % of their measurements. Here it should be noted that the Chalmers system was an experimental system used for the first time in this application. Another similar comparison to the Stena Germanica ferry in C2 showed that the Explicit drone-based measurements overestimated the FSC but when considering the type B uncertainty, the agreement was good; this contrasted with the Chalmers high-sensitive sniffer and the BSH standard sniffer systems who showed biases which obscured the comparison of their uncertainty estimates.

For NO_x there is a lack of real comparison data and the comparison had to be made against the ensemble mean. Nevertheless, the comparison in section 3.2.2 between the Aeromon mini-sniffer drone and the Chalmers sniffer, operated from a measurement vessel in Marseille, shows reasonable agreement when the uncertainty estimates are included. Also, the comparison between the Chalmers high-sensitive system(cth.las), the BSH standard sniffer and the Explicit drone sniffer in C2 shows good agreement (Figure 34).

There is still a need for further work on a common method for the uncertainty calculation, even though good progress has been made in SCIPPER. Ideally this work should be done in a future standardization project.

Table 8. Summary of Implementation of the uncertainty analysis of each group.

Parameter	Type	Reported value	Group	Comment
Total Uncertainty	A	CI 95 %	Chalmers	
Total Uncertainty	A	CI 68 %	TNO, BSH, Aeromon	
Total Uncertainty	B	CI 68 %	Explicit	
S/N of SO_2, CO_2, NO_x	Random		Chalmers (standard, laser, drone)	Expanded standard uncertainty of the mean of background relative to average peak height
S/N of SO_2, CO_2, NO_x	Random		TNO, BSH, Aeromon	STD of background relative to peak height Aeromon uses the STD of plume measurement value relative to peak height which is a combination of S/N ratio and sampling stability during plume visit.
Baseline variability	Random		Chalmers	Change in background value (baseline) between both sides of the plume
Calibration variability	Random		Chalmers	Variability in response between calibrations
NO Interference	Systematic		Chalmers	
Calibration gas	Systematic		All	Uncertainty stated on gas certificate by supplier
Sulphur and Carbon conversion	Systematic		BSH	Possible loss of Sulphur and Carbon to other species than SO_2 and CO_2 during combustion (cf. Sec. 2.3.1.3)

Validity of FSC formula and gas distribution in plume	Random		BSH	Assumption of 87% Carbon in fuel and non-perfect distribution of gas within plume
Water interference	Systematic		BSH	Possible cross interference of standard SO_2 instrument for H_2O
FSC_RH_bias	Random		Aeromon	Negative bias in FSC estimations at high RH for standard sniffer. Aeromon performs a humidity shock response test during daily calibration level check and RH response is compensated.

5.4 Reporting

Table 9 lists the parameters that are required or optional when reporting the FSC values obtained from remote measurements to the inspection authority. It is also recommended that the measurement meta data is provided, which should describe the monitoring system and where data can be found, ideally in an open central facility, e.g. through an internet link. For NO_x, fuel-mass specific emissions should be reported with the same type of reporting as for FSC. However, it is also recommended to report break specific emissions (g/kWh) and then include the ship speed and the calculated or reported engine load and SFOC. The two latter can be obtained from the propeller law and an engine type specific SFOC curve as described in the IMO 4th greenhouse gas report or better obtained from shipping companies or authorities. It is suggested to report break specific emissions only for ships above 50 % load but, more discussion on this subject is needed, especially for Tier III ships.

Table 9. Suggested parameters to report related to remote FSC measurements.

Parameters	Comment
Measurement time	UTC
Measurement Location of sensor	Longitude, Latitude
FSC in % S m/m	All valid measurements should be reported (independent of value) and no usage of quality flags. Negative values and measured SO_2 values below the detection limit (3σ) should be reported as the minimum detectable FSC based on 3σ level.
Absolute Uncertainty in % S m/m	Report at CI 95 % according to GUM Reporting also at 68 % for triggering on board inspection.
Compliance flag	Boolean dependent on measurement result subtracted by absolute uncertainty compared to regional IMO limit
Ship Identity: IMO	Required
Ship Identity MMSI	Required
Port of call	Required
Site name (fixed sites)	Optional
Reporting Entity	Contact details of reporting entity
Ship speed over ground	Optional
Ship Course	Optional

6 Conclusions

6.1.1 Harmonization

The objective of this report is to align the measurement procedures and reporting of the different institutes/companies for remote sniffer measurements of Fuel Sulphur Content and future measurements of NO_x and PM. Three general types of sniffer techniques have been included, i.e. high-sensitive, standard and mini sniffers. Optical methods are not included. This work should lay the foundation for present and future standardization work. The main conclusions are:

- Data validation
  - It is necessary to assess whether a ship plume measurement is valid.
  - A quality scoring system can be empirically derived based on various parameters such as: sensor technique, weather conditions, signal-to-noise ratio, traceability (calibrations, linearity checks, cross interference calibration) and whether the ships are uniquely identified.
  - The quality scoring should not be part of the reporting since it is given in arbitrary units and hence not comparable between different operators and instruments. However, it can be an integral part of the uncertainty estimation (next main bullet).
  - Regular comparisons between the different measurement groups is recommendable, possibly as intercomparison measurement campaigns or with Round-Robin tests.

- Uncertainty calculation

  Two types of uncertainty assessment can be applied, based on GUM:
  - Type A: the uncertainty is calculated for each single measurement, based on the signal to noise, other measured parameters and systematic uncertainties.
  - Type B: a typical uncertainty (derived by performance testing) is assigned based on the assessment of measurement quality.
  - It should be investigated in field validation work whether the reported measured uncertainties by the individual institutes/companies are realistic.

- Reporting
  - All valid measurements should be reported with the expanded uncertainty given as CI 95 % . However, for triggering on-board inspections the standard uncertainty (CI 68 %) is recommended and it needs to be subtracted from measured results when assessing whether a ship is operating on compliant fuel or not.
  - Negative values should be reported as the minimum detectable FSC based on the SO_2 concentration detection limit (3σ).

6.1.2 Field validation FSC

The error of the FSC measurements was assessed by side-by-side measurements and by comparison to on-board taken fuel samples which were analysed in certified laboratories. It is assumed that the measurement error, which is the actual deviation from the true value, corresponds to the measurement uncertainty (with the same distribution), i.e. the predicted deviation from the measured value that should be reported by the instruments. The main conclusions are:

- The tested remote measurement systems are able to measure the FSC with a total uncertainty (CI 95 %) varying between 0.03 % and 0.14 % S m/m at 0.1 % FSC level, varying between system, as demonstrated in the field validation campaigns.

  - The high-sensitive sniffer system had random 2σ-measurement errors of 0.01 % and 0.04 % S m/m, respectively, for the measurements in C2 and C3. It also had corresponding negative biases of

0.03 % and 0.02 % S m/m. This resulted in total measurements errors of 0.036 % and 0.05 % S m/m for C2 and C3, respectively, including both bias and random errors (CI 95 %).

- The standard sniffer systems had random 2σ-measurement errors between 0.05 % and 0.08 % S m/m and negative biases varying between 0.02 % and 0.08 % S m/m, depending on instrument, in both C2 and C3, which resulted in total measurement errors of 0.07 % to 0.09 % S m/m.

- The Explicit (mature) drone-based mini sniffer system had a random 2σ-measurement error of 0.025 % S m/m and a bias of 0.015 % S m/m in C2, resulting in a total measurement error of 0.03 % S m/m including both bias and random errors (CI 95 %). The same system operated in C3 showed a random 2σ-measurement error of 0.08 % S m/m and no bias, resulting in a total uncertainty of 0.08 % S m/m.

- The Chalmers (experimental) drone-based mini sniffer system had a 2σ-variability of 0.13 %, a bias of -0.033 % S m/m and a total uncertainty of 0.14 % S m/m in C3.

- The Aeromon (mature) drone-based mini sniffer was compared against a variant of a standard sniffer (made for airborne measurements) at FSC ranges between 0.1 % to 3.5 % S m/m in C1. The comparison showed overall good agreement between the measurements in terms of magnitude and correlation (slope 0.99. offset -0.02 % S m/m, R^2 0.95) and the 1σ–variability between individual measurements was 0.20 % S m/m. If the average of the two systems is assumed as the "true" emission, then the expanded uncertainties (CI 95 %) for both instruments are 0.20 % S m/m.

- The reported uncertainties of the standard and high-sensitive sniffer systems varied between 0.01 % to 0.02 % S m/m. They matched the actual measurement error poorly in C3 (20-30 % of all cases) and not at all in C2, due to the negative measurements bias that was observed. For the Explicit drone-based mini sniffer the reported uncertainties were around 0.03 % S m/m and they matched the measurements error in 2/3 of the cases in C3 for 14 individual ships and in 9/10 cases in C2 when measuring on 11 occasions on the same passenger ferry.

- An apparent dependence between measurement error and relative humidity (RH) for the high-sensitive system and several of the standard sniffer systems were observed in C3. For the high-sensitive system, using only FSC data acquired below RH 80 %, the negative bias basically disappeared. Also, the total uncertainty (CI 95 %) decreased from 0.044 % to 0.028 % S m/m. However, this effect was not observed in C2 in which the bias was present in all data even though all sampling were below RH 80 %. The apparent relative humidity effect hence needs to be studied further.

- The influence of using dry calibration gas should be investigated in future studies, since this is a potential error in the SO_2 calibration of the standard system since the measurements occur in wet outside air. Also the NO cross-interference effect on the SO_2 reading of the standard systems should be investigated further.

- It was found that one of the calibration gas cylinders for SO_2 used in the project differed by 40 % from its certificate. To minimize such problems, it is recommended that newly delivered calibration gas cylinders are checked against old cylinders and occasionally checked in certified laboratories.

6.1.3 Field validation NO_x

The measurement error of the fuel-mass specific emission of NO_x was estimated by comparing individual measurements to the ensemble averages of the different instruments in three field campaigns. This was done since on board taken (true) reference values are not available as for FSC. It is assumed that the measurement error, which

is the actual deviation from the true value, is the same as the uncertainty, i.e. the predicted deviation from the measured value that should be reported by the instruments. The conclusions are:

- The tested remote measurement systems are able to measure the mass-specific NO_x emissions with about 17-40 % (8-17 g/kg$_{fuel}$) with 95 % confidence level. The reported uncertainties agree reasonably well with observed measurement errors with some exceptions.

  - In C1 (Marseille), the measurements of a standard sniffer and a mature drone-based mini sniffer, for ships operating on open waters, outside Marseille and Fos-sur-Mer, had a systematic difference of 5 g/kg$_{fuel}$ (13 %) with 7 g/kg$_{fuel}$ 1σ variability (17 %). If the average of the two systems is assumed as the "true" emission, then the total measurements error (CI 95 %) corresponds to 7.5 g/kg$_{fuel}$ (17 %). The reported uncertainties by the two systems explain the differences in 90 % of the cases.

  - In C2 (Kiel), 19 ships were measured using two, shore based, standard sniffers (Chalmers and BSH) and a mature drone-based mini sniffer (Explicit) and the agreement was better than 10 % between the data sets. The reported uncertainties by the systems explain the differences in 85 % of the cases. The comparison shows systematic differences and random 1σ variability of -4.8±4.7 g/kg$_{fuel}$ (-8.5±11) %, 4.0±4.9 g/kg$_{fuel}$ (-6.9±11) % and -2±4.3 g/kg$_{fuel}$ (-3.3 ±8.5) % for the NO_x EF measurements by Chalmers, BSH and Explicit, respectively. This corresponds to a total measurements error (CI 95 %) of 10.6 g/kg$_{fuel}$ (23 %) for the fixed shore-based systems and 8.8 g/kg$_{fuel}$ (17 %) for the drone-based mini sniffer. Noteworthy is that the reported calculated uncertainties, are larger for all systems (20 %) than the apparent measurements error.

  - In C3 (Wedel), several hundred ships of varying size were measured using several fixed sniffers and results were compared to the ensemble mean. The reported uncertainties matched the apparent measurement errors in 60 - 70 % of the cases, with exception for one system which only matched in 25 % of the cases. The comparison shows systematic differences and random 1σ variability of 5.5±8.4 g/kg$_{fuel}$ (12±19) %, -5.1 ± 8.4 g/kg$_{fuel}$ (-14±20) % and -1.2 ± 8.8 g/kg$_{fuel}$ (1.6±18) % for the NO_x EF measurements by Chalmers, BSH and TNO, respectively. This corresponds to a total measurement error (CI 95 %) of approx. 17 g/kg$_{fuel}$ (40 %) for all groups. The measurement error is almost twice as large as obtained in campaign C1 and C2 and this could partly be caused by the fact that some of the data were evaluated automatically and that there is potentially a larger spread in ship types and engine load, than in the other campaigns, since the ships were sailing with, and against tide.

- The brake specific NO_x emission for a certain engine load is obtained by multiplying with the specific fuel oil consumption (SFOC) obtained in this study using equations in the IMO 4th greenhouse study. The brake specific emissions for Tier I and Tier II ships do not appear to depend on engine load above 50 % and therefore the measurements will correspond quite well to the load weighted average in the NOx technical code[4]. For loads below 50 % more care has to be taken when assessing whether a ship complies with the IMO rules for its specific Tier. For Tier III the implementation of a NTE limit of no more than 50 % above the applicable NO_x emission limit (max 3.4 g/kWh) at <u>any</u> load point, means that even low loads (<50 %) have a regulatory 'cap' which may be monitored via remote measurements. However, knowing the SFOC value for each ship is crucial for the factor calculation and an accurate emissions analysis will depend on individual SFOC data confirmed by the shipping companies. Such data are currently not readily available but could be envisaged reported to the authorities going forward for the purpose of NO_x compliance monitoring. An even better solution would be to define an NTE limit value for NO_x in g/kg fuel specifically for monitoring and enforcement purposes (in addition to the laboratory test procedure in g/kWh). It is recommended to add these elements for monitoring and enforcement to the MARPOL NO_x technical code.

[4] The end result of the test cycle is also heavily weighted towards the 75% and 100% load points, namely 70%. Also since g/h NOx are weighted, the end results is for almost 90 % determined by the high load points .

6.1.4 Particles

Particulate emission measurements from sea-going vessels were intercompared by several techniques in the C3 campaign in Wedel. These techniques were also employed in C1, C2 and C4. The most important conclusions with respect to recommendations for future monitoring of particulate matter are:

- Particle measurements are more difficult to compare than gaseous species measurements since different instruments measure various properties of the particles. The instruments that were used to determine information on the particle size and concentration are based on size classification of single particles by measuring optical scattering or electromobility. In addition, light absorption was used to measure the mass of black carbon. During C3, only some of the instruments dried the particles before measuring. Nevertheless, a satisfactory agreement was found with differences in the derived particle number-, and mass concentration ranging from far below 10 % to maximum 35 %.

- In general, it was found that 85 % of the emitted particles (measuring particles with diameter above 5.6 nm) have a diameter between 10 and 80 nm. If the particle mass emission is considered instead of the particle number emission, it can be concluded that 70 % of the emitted particle mass comes from particles with a size between 40 and 200 nm. This leads to three recommendations:

    - For an effective reduction of particle emissions from sea-going vessels, with possible future regulations, the emission of very small particles should be considered. Specifically, the following particle size ranges should be taken into account: 10 nm to 80 nm for number concentration and 40 nm to 200 nm for mass concentration.

    - Potential future regulations should give a limit value related to the fuel mass, such as particle number or particle mass per kg of burned fuel. In that way possible violations can be monitored more easily than for NO_x. PM can also be reported as brake specific emissions in a corresponding manner as for NO_x as discussed above.

    - For the measurement of particle emissions from vessels, sufficiently sensitive instruments must be used, covering at least the above given size ranges. However, to cover most of the particles and their mass it is recommended to extend the size ranges to the following: 10 nm to 100 nm for number concentration and 40 nm to 2.5 µm for mass concentration. For all devices it is recommended to use a temporal resolution of 10 seconds or smaller, during which all channels are scanned through.

- Total particle emission factors were found to be in the range from 0.8 ·1016 to 1.625·1016 #/kg_{fuel} on average. However, different emission behaviors were found for different types of vessels. The particle number seems to be inversely proportional with the size of the vessel, ranging from below 100 to above 300 m. However, this needs to be investigated further.
-
    - Emission factors found for inland vessels were on average four times lower than for the small seagoing cargo vessels.

    - Emission factors for seagoing passenger vessels were on average at the lower end of the size range.

7 Acknowledgements

We thank Jukka-Pekka Jalkanen and Lasse Johansson for providing ship emission CO_2 data from the STEAM model. We also thank the German Wasser und Schifffahrtsamt (WSA) for hosting measurement campaign C3 at their facilities in Wedel, as well as the German Waterway Police (WSP) in Hamburg for collecting the fuel samples from selected vessels, measured during C3.

8 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.

Ausmeel, S., Eriksson, A., Ahlberg, E., and Kristensson, A.: Methods for identifying aged ship plumes and estimating contribution to aerosol exposure downwind of shipping lanes, Atmos. Meas. Tech., 12, 4479–4493, https://doi.org/10.5194/amt-12-4479-2019, 2019.

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://www.atmos-meas-tech.net/7/2597/2014/amt-7-2597-2014.pdf, 2014.

Balzani-Lööv, J. M., et al.: Field test of available methods to measure remotely SOx and NOx emissions from ships, Atmos. Meas. Tech., 7, 2597-2613, 2014, www.atmos-meas-tech.net/7/2597/2014/ doi:10.5194/amt-7-2597-2014, 2014.

Beecken et al, 2019, Review of available remote systems for ship emission measurements, report D 2.1, SCIPPER Horizon 2020.

Beecken J, Mellqvist J, Salo K, Ekholm J, Jalkanen JP, Johansson L, et al. Emission factors of SO2, NOx and particles from ships in Neva Bay from ground-based and helicopter-borne measurements and AIS-based modeling. Atmospheric Chemistry and Physics. 2015;15(9):5229–41.

Beecken, J., Mellqvist, J., Salo, K., Ekholm, J., and Jalkanen, J.-P.: Airborne emission measurements of SO2, NOx and particles from individual ships using sniffer technique, Atmos. Meas. Tech., 7, 1957–1968, 2014.

Beecken, J., Mellqvist, J., Salo, K., Ekholm, J., and Jalkanen, J.-P.: Airborne emission measurements of SO2, NOx and particles from individual ships using a sniffer technique, Atmos. Meas. Tech., 7, 1957–1968, https://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 SO2, NOx 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.

Berg, N., Mellqvist, J. et al.: Ship emissions of SO2 and NO2: DOAS measurements from airborne platforms, Atmos. Meas. Tech., 5, 1–14, 2012.

Berg, N., Mellqvist, J., Jalkanen, J.-P., and Balzani, J.: Ship emissions of SO2 and NO2: DOAS measurements from airborne platforms, Atmos. Meas. Tech., 5, 1085–1098, https://doi.org/10.5194/amt-5-1085-2012, 2012.

EU: Council Directive 1999/32/EC of 26 April 1999 relating to a reduction in the sulphur content of certain liquid fuels and amending Directive 93/12/EEC, OJ L 121, 11.5.1999, p. 13–18 (EN), https://eur-lex.europa.eu/eli/dir/1999/32/oj, 1999.

EU: Directive 2012/33/EU of the European Parliament and of the Council of 21 November 2012 amending Council Directive 1999/32/EC as regards the sulphur content of marine fuels, OJ L 327, 27.11.2012, p. 1–13 (EN), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32012L0033, 2012.

Explicit: Airborne Monitoring of Sulphur Emissions from Ships in Danish Waters, 2017 Campaign Results, Environmental Project no. 2001, Danish Environmental Protection Agency, https://www2.mst.dk/Udgiv/publications/2018/04/978-87-93710-00-9.pdf, April 2018.

ISO/IEC Guide 98-3:2008, Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995).

Kley, D. and McFarland, M.: Chemiluminescence detector for NO and NO2. Atmos. Technol., 12:63–69, 1980.

Knudsen B. et al., 2022, NOx Emissions from Ships in Danish Waters, Assessment of Current Emission Levels, and Potential Enforcement Models, ISBN: 978-87-7038-384-4, Report to Danish EPA.

Mellqvist, J., Beecken, Conde, V, and J. Ekholm, CompMon report: Certification of an aircraft and airborne surveillance of fuel Sulphur content in ships at the SECA border, Chalmers University of Technology, 2017,, https://research.chalmers.se/publication/500250/file/500250_Fulltext.pdf, https://compmon.eu/.

Mellqvist, J., Beecken, J., Conde, V., and Ekholm, J.: Surveillance of Sulphur Emissions from Ships in Danish Waters, https://doi.org/10.17196/DEPA.001, 2017.

O'Keefe, A., and Deacon, D. A. G.: Cavity ring-down optical spectrometer for absorption measurements using pulsed laser sources. Review of Scientific Instruments, 59(12):2544–2551, ISSN 00346748 (ISSN), 1988.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy – Principles and Applications. Springer, Berlin, Heidelberg, ISBN 978354021193897835407577 64. doi: 10.1007/978-3-540-75776-4, 2008.

PROMINENT D5.8, 2017 D5.7 Technical recommendations on options for specifications of monitoring equipment and database set-up. TNO, TÜV Nord, SGS, IMST. PROMINENT WP5 report, grand agreement 633929, Dec, 2017.

Repka et al., Envisum final report, 2019, Clean Shipping: Exploring the impact of emission regulation, https://projects.interreg-baltic.eu/fileadmin/user_upload/Library/Outputs/EnViSuM_FinalReport.pdf, https://blogit.utu.fi/envisum/.

Sofiev, M., Winebrake, J.J., Johansson, L. et al. Cleaner fuels for ships provide public health benefits with climate tradeoffs. Nat Commun 9, 406 (2018). https://doi.org/10.1038/s41467-017-02774-9.