



Glass plate sampling efficiency for trace gases in the sea surface microlayer

Lea Lange¹, Dennis Booge¹, Hendrik Feil^{1,2}, Josefine Karnatz¹, Ina Stoltenberg¹, Hermann W. Bange¹, Christa A. Marandino¹

5 ¹Marine Biogeochemistry, GEOMAR Helmholtz Centre for Ocean Research Kiel, 24148, Germany

²Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky Universität Oldenburg, Germany

Correspondence to: Lea Lange (llange@geomar.de)

Abstract. Many climate-active trace gases in the atmosphere are closely linked to production and consumption in the ocean, which are, in turn, influenced by the sea surface microlayer (SML). The SML is the upper most layer of the ocean with up to 1 mm thickness, often enriched in organics. Studies of trace gases in the SML aim to identify and quantify potential processes unique to the SML and to understand the SML's influence on the transfer between air and sea. Established sampling techniques of the SML (e.g., glass plate, mesh screen) are associated with high losses for the volatile trace gases. Despite the high losses, in this study we find that meaningful analysis of glass plate samples for trace gases is possible. We experimentally determined the sampling efficiency for the short-lived trace gases dimethyl sulphide (DMS), isoprene, and carbon disulphide (CS₂). Water temperature and trace gas concentration were the main drivers for sampling efficiency variability, while salinity and the number of dips of the glass plate were not significant. The effect of surfactants could not finally be untangled. Although our results are consistent, we do not quantify a sampling efficiency to correct individual measurements, as our experiments did not encompass the full suite of environmental parameters normally encountered in the field. . Instead, we suggest to use 0.13 ± 0.01 ($\pm$ standard error) for DMS and isoprene, and 0.12 ± 0.01 for CS₂ as thresholds to identify cases of net production in the SML. Future studies should extend to long-lived species (e.g., nitrous oxide, methane), include the effect of wind, and be repeated for the mesh screen. We hypothesize that a correction of individual measurements requires to determine sampling efficiency as a function of environmental parameters, for which the underlying physicochemical relationships need to be unraveled by increasing the parameter space studied here.

1 Introduction

Short- and long-lived trace gases impact Earth's climate via processes like hydroxyl radical chemistry, aerosol formation, cloud condensation nuclei formation, or the greenhouse effect. The oceans serve as important sinks and sources for many of these gases. Understanding trace gas cycling in the ocean mixed layer and exchange with the atmosphere across the air-sea interface is, therefore, crucial for climate and air quality predictions. (Liss et al., 2014) Trace gas data from the upper meter is scarce and the lack of information is even worse for the surface microlayer (SML). The SML is the uppermost layer of the water column with a thickness of less than 1 mm. It often shows organic enrichment, for instance with surface-active substances



(surfactants), especially during calm conditions (Wurl et al., 2011). Enrichment is often indicated as enrichment factor (*EF*), defined as the ratio between the quantity's concentration in the SML and in the ULW. Surfactants alter the physicochemical properties of the SML, like reducing the surface tension, and are addressed in most SML studies. For high concentration of surfactants so-called slicks form that are often visible as flat surfaces, as they dampen the waves. The SML has been studied
 35 for a wide range of physical, chemical and biological parameters, which resulted in the hypothesis of the SML being a biogeochemical reactor (Bibi et al., 2025). However, established sampling techniques are not well suited for volatile gases, causing a severe lack of information on trace gas concentrations directly at the air-sea interface. For ease and convenience, many air-sea gas exchange estimates are thus based on underlying water (ULW) concentrations sampled from the mixed layer (usually around 5 m depth). Studies have shown evidence that this neglect of processes occurring shallower than 5 m and in
 40 the SML may be the reason for discrepancies between computed fluxes and predicted or measured fluxes (Booge et al., 2016; Kock et al., 2012; Lennartz et al., 2017; Marandino et al., 2005, 2008). The SML is ubiquitous in the ocean, therefore inaccuracies in air-sea gas exchange estimates may accumulate and may show a significant effect on atmospheric chemistry and climate computations even on a global scale (Sabbaghzadeh et al., 2017; Wurl et al., 2011). To unravel the relevance of the SML on air-sea gas exchange and its potential as a source or sink for trace gases, it is necessary to sample the SML
 45 efficiently and accurately.

In order to measure SML concentrations and gradients at sea, samples must be taken by hand with specialized equipment from a small rubber boat, which is less convenient and more difficult than the common underway pumping system or discrete water samples from a CTD rosette used on board during many oceanographic campaigns. A widely used sampling device is the glass plate, which was first described by Harvey and Burzell (1972) for collection of microorganisms and organic content, but other
 50 techniques exist as well, such as the mesh screen (Garrett, 1965), rotating drum (Harvey, 1966), membrane sampler (Crow et al., 1975), cryogenic sampler (Turner and Liss, 1985) and more recently a gas-permeable floating tube specially designed for trace gases (Saint-Macary et al., 2023). Several studies discuss the limitations and difficulties to compare results from the different methods for non-volatile compounds (e.g., Cunliffe and Wurl, 2014) and additionally the need for a new technique for volatile compounds like trace gases is highlighted (e.g., Engel et al., 2017). There are only a few studies that sampled the
 55 SML for trace gases and even fewer that used the glass plate, although it is the most common device for other parameters measured in the SML. To the best of our knowledge, there are no studies with SML samples for the short-lived trace gas isoprene, and only one on carbon disulphide (CS₂; Turner and Liss, 1985), but dimethyl sulphide (DMS) measured from SML samples has been used more widely to determine *EF* (see Saint-Macary et al., 2023; Turner and Liss, 1985; Walker et al., 2016; Yang, 1999; Yang et al., 2001). In those and similar studies, glass plate and mesh screen sampling are usually described
 60 as being prone to sampling losses due to volatilization, but to the best of our knowledge the magnitude of the losses has not been tested and no corrections attempted. One study accompanied their field measurements of DMS *EF* with laboratory work to investigate the dependence of the *EF* on factors like temperature, salinity, and DMS concentration (Yang et al., 2001). They address potential losses, but their experimental design does not allow to infer reliable estimates of sampling efficiency. Neglecting losses might lead to wrong interpretation of calculated enrichment factors, especially when losses are high.



65 Trace gas losses in SML sampling are consistently addressed in relevant publications using the glass plate or mesh screen, because the sampled water is spread in a very thin film over a large surface exposed to air before it is transferred to a sample bottle. Diffusive losses driven by the concentration gradient between air and sample water are larger with stronger gradients. Additionally, the exposure to light might drive photochemical processes on the glass plate or mesh screen. Furthermore, the sample bottle is open for the duration of sampling, as usually one dip of the glass plate or mesh screen does not provide enough
 70 sample volume for analysis. Finally, the sample is wiped from the glass plate and drips into the sample bottle for both the glass plate and mesh screen, introducing additional turbulence that increases the losses further. Apart from losing gas to the surrounding air, there is also a potential loss source from adsorption to the glass plate and mesh screen surface, the wiper, and the funnel.

Studies using a gas-permeable floating tube (e.g., Saint-Macary et al., 2023) attempt to overcome those issues, but their results
 75 are difficult to integrate with quantities obtained from the glass plate or the mesh screen. Results from different SML sampling techniques are in general complex to compare for three reasons: (1) sampling thickness differs significantly between techniques (2) sampling losses vary (volatilization, adsorption, undersize, ...) (3) introduction of error (e.g., stressing phytoplankton by sample handling). Even for non-volatile compounds, like surfactants, sampling losses might not necessarily be negligible (see controversy on mesh screen sampling efficiency in Garrett (1974)), though in general they are considered to be. For these
 80 reasons, trace gas concentrations and resulting *EF* vary significantly between SML techniques (see Turner and Liss, 1985; Walker et al., 2016; Yang et al., 2001), which commonly is attributed to differences in sampling thickness.

This study has been conducted in the context of the joint, interdisciplinary campaigns within the project Biogeochemical processes and Air-sea exchange in the Sea-Surface microlayer (BASS), where the glass plate technique was used. To enable the best possible comparability, we aim, in this study, to identify sampling efficiency of the glass plate technique for volatile
 85 trace gases like DMS, isoprene and CS₂. Due to the volatile nature of these gases, sampling losses are expected to be high for glass plate sampling, yet, existing studies have shown that glass plate samples are still above the limit of detection. We therefore hypothesize that meaningful analysis of glass plate samples is possible. Physicochemical properties of the sample, like temperature, salinity, and trace gas concentration in ULW, as well as the individual steps performed in glass plate sampling are expected to affect the sampling loss. As those differ between samples, we expect to see variations of the sampling losses
 90 correlated with those parameters. Therefore, in this study we make the first attempt to derive a correction method for the classic glass plate technique. To do so we investigated a range of isotopically labelled atmospherically short-lived trace gases (DMS-d₃, isoprene-d₅, ¹³CS₂), addressing the following questions: (1) Are SML samples from the glass plate meaningful for trace gases? (2) If so, what drives the associated losses of those volatile compounds? (3) Can we derive a method to assess the SML enrichment accurately in experiments and in the field?



95 2 Methods

To quantify the sampling efficiency of the glass plate, one would have to compare in situ concentrations in the SML to the measured concentration from the glass plate sample. However, there is no established sampling technique or in situ sensor for trace gases in the SML. We, therefore, performed experiments in a tank where the water was mixed immediately before sampling. This ensured that the SML had the same concentration of trace gases as the ULW, which can be sampled without losses. The deviation of the concentration measured in the SML sample from the ULW samples reflects the sampling efficiency of the glass plate method. Additionally, the mixing ensures that the SML sample is not diluted by ULW (as $C_{SML} = C_{ULW}$) and we, therefore, capture the pure sampling error. Note, that when we use SML here we refer to the upper 1 mm of the water body, although there is no enrichment of organics present.

2.1 Experimental setup

105 In 2023, we conducted a series of lab experiments to assess the trace gas loss during glass plate sampling: (A) preliminary test (B) glass plate sampling without surfactants (C) glass plate sampling with surfactants. The preliminary test from experiment A are described in Appendix A Preliminary test, while we exclusively address experiment B and C further on. To exclude biological and chemical consumption or production of gases as potential sources or sinks affecting the measured concentrations, we used isotopically labelled gases in fresh water (FW) and artificial seawater (AS): deuterated DMS (DMS-d3), deuterated isoprene (isoprene-d5) and $^{13}\text{CS}_2$ as representatives for DMS, isoprene and CS_2 , respectively. Trace gas concentration, water temperature, salinity, and surfactants were varied in the ULW to assess their effect on sampling efficiency, whereas surfactants in the SML, sampling duration for SML, sample volume of SML samples, number of glass plate dips, pH in ULW, air temperature and person who sampled SML were recorded to accompany our measurements. To exclude sampling bias with glass plate sampling, only one person was sampling SML samples (with exceptions in experiment B). The experiments were performed with different media to address two research objectives:

- (1) Pure sampling efficiency was determined using ultrapure water (as fresh water) and artificial seawater with varying temperatures, salinities and trace gas concentration, as those parameters affect the solubility and diffusivity of gases.
- (2) Artificial surfactants were mixed into the artificial seawater in three different amounts to understand if their presence affects the losses.

120 Ultrapure water (18.2 MΩcm) was provided from lab water purification systems (Arium® Pro, Sartorius, Göttingen, Germany). Artificial seawater was either prepared with Tropic Marin® Pro-Reef Sea Salt (Tropic Marin AG, Hünenberg, Switzerland) or with ordinary aquarium salt in ultrapure water aiming for a practical salinity (S_P) above 30.0. Every day we spiked the water with isotopically labelled trace gases. The concentrations of isotopes varied between the experiments, as two different mixtures of standards were used.



125 Table 1 summarizes the experiments, which took place in two different locations in Germany: (1) under the roof of the SURF facility at ICBM, Wilhelmshaven and (2) in the laboratories at GEOMAR Helmholtz Centre for Ocean Research Kiel. Two different tanks were used: (1) aquarium made of glass (2) "AZ crate" (Polypropylene, Schoeller Allibert, Schwerin, Germany).

130 **Table 1 Overview of experiments performed in this study. Sample volumes are the targeted volumes. Analysis gas chromatography–mass spectrometry is abbreviated as GC–MS. Research objectives: (1) pure sampling efficiency (2) effect of surfactants on sampling efficiency.**

Experiment	Tank	Treatments	Medium	Analysis Sample volume	Additional measurements	Research objectives
A (GEOMAR) 13–15 March	incubation bath	fresh water	ultrapure water	GC–MS 10 mL		(1)
B (ICBM) 18–20 April	aquarium	fresh water	ultrapure water 63–68 L			
		artificial seawater	Tropic Marin® Pro-Reef Sea Salt 74 L, $S_p = 34.5$ 7.5 °dKH	GC–MS 10 mL	occasional T_{water} $T_{air} \approx 12.7$ °C	(1)
C (GEOMAR) 28 July – 31 August	"AZ crate"	fresh water	ultrapure water 78 L $pH = 6.3$	GC–MS 10 mL	T_{water}	(1), (2)
		artificial seawater	ultrapure water with ordinary aquarium salt		$T_{air} \approx 23.1$ °C	
		artificial seawater with surfactants (3 levels)	78–81 L S_p : 29.1 and 33.9 pH : 8.06–7.84	GC–MS, voltammetry 15 mL	($\Delta T_{air} = 4$ °C)	(2)

2.2 Sampling

Before each sampling, we mixed the water body carefully for one min and waited for two min for the turbulent motion to subside. We sampled the SML, ULW (10–13 cm) and (1–6.5 cm). All samples were collected in amber borosilicate glass bottles (20 mL). SML samples in experiment A (preliminary test) were taken with the mesh screen (Garrett, 1965). All other SML samples were taken with the glass plate technique (Harvey and Burzell, 1972). Following the recommendation by



Cunliffe and Wurl (2014) and examples from recent studies (e.g., Adenaya et al., 2021), we reduced the speed of extracting the glass plate to 5 cm/s. The glass plate (borosilicate glass, 30 cm × 25 cm × 0.5 cm in experiment B, 40 cm × 30 cm × 0.5 cm in experiment C) was immersed to a depth of about 20 or 25 cm. Both sides of the glass plate were wiped with a regular silicone or rubber wiper and the run-off was collected into a sample bottle with a plastic funnel. A single dip provided between 1.8–6.3 mL of medium. To collect sufficient sample volume for analysis two to five dips were required for analysis with gas chromatography–mass spectrometer (GC–MS) and voltammetry samples. The thickness of the sampled layer is determined by Eq. (1) (Cunliffe and Wurl, 2014).

$$h_{SML} \approx h_{sampled} = \frac{V_{sample}}{n_{dips}A} \quad (1)$$

where h_{SML} is the true thickness of the SML, $h_{sampled}$ is the thickness that was actually sampled in cm, V_{sample} is the total volume of the glass plate sample in mL, A is the wetted surface area of the glass plate in cm² and n_{dips} is the number of dips that were needed to collect V_{sample} .

Each SML sample is always associated with a reference ULW sample and sometimes with additional ULW or surface samples. All samples taken together after the same mixing event are referred to as a sample set. The reference ULW sample was taken immediately after the first dip of the SML sample was collected, interrupting the SML sampling for less than a minute (overall SML sampling took 90–235 s). All additional ULW or surface samples were taken after the SML sample had been finished (exception from this order is 19 April in experiment B). ULW samples were taken with a pipette (5 and 10 mL, Eppendorf SE, Hamburg, Germany) set to the required sample volume by inserting the tip of the pipette vertically and carefully into the water. For sample volumes of 15 mL, a pipette was immersed twice in the same location. ULW samples were taken from a water depth below surface of ~10 cm in experiment B and ~13 cm in experiment C. Surface samples were sampled the same way as ULW samples, with the only difference that we were aiming for a depth below surface of less than 1 cm in experiment B and ~4.5 cm in experiment C (with exception of 31 August: ~6.5 cm). Additional ULW and surface samples were collected once or twice daily to confirm that the tank was well mixed. Before each sample set was obtained, all of the equipment was rinsed with ultrapure water.

Samples were taken for trace gas analysis and voltammetry (surfactants). Trace gases were analysed during all of the experiments, whereas voltammetry was only performed in selected cases (Table 1). The targeted sample volume for trace gas analysis was 10 mL. SML samples were filled to at least the targeted sample volume and later on weighed to calculate the sample volume based on the density of the medium. In the treatment with surfactants, a larger sample volume was required for voltammetry measurements to permit a subsequent 10 mL subsample. Given the limited SML volume in the tank and the presence of surfactants, collecting two true replicate SML samples was not feasible. Consequently, the target sample volume was increased to 15 mL for GC–MS analysis, and the purged sample for GC–MS analysis was subsequently reused for surfactant analysis. No effect on final concentrations from the GC–MS measurements was found between 10 mL and 15 mL samples, as gases are stripped completely out of the medium within the purge time. Sample bottles used for the voltammetry measurements were additionally acid-rinsed (10 % HCl) and pre-combusted at 500 °C before sampling.



2.3 Sample analysis trace gases

170 The system used to measure the trace gases DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ is a purge and trap GC–MS (Zavarsky et al., 2018). It consists of a self-constructed purge and trap system, a type 7890A GC, and a 5975C MS (Agilent, Waldbronn, Germany). The GC uses a capillary column (Supel-Q-PLOT, 30 m $\times$ 0.32 mm, Merck KGaA, Darmstadt, Germany). The MS is equipped with electrical ionisation, a quadrupole mass filter and a triple-axis electron multiplier detector. All three compounds are measured from the same sample. As samples cannot be preserved, no replicates could be taken for GC–MS

175 analysis. Samples were stored at room temperature in the dark and measured within less than one hour after sampling. Samples (both 10 and 15 mL) are purged with helium (99.999 %, Air Liquide, Düsseldorf, Germany) at a flow rate of 80 mL min $^{-1}$ for 10 min. The sample gas stream is dried with a Nafion $^{\text{®}}$ membrane dryer (Perma Pure, Lakewood, United States) and trapped with liquid nitrogen in a Sulfinert $^{\text{®}}$ stainless steel tube (Restek GmbH, Bad Homburg, Germany). Injection into the GC is semi-automated with a 6-port valve that channels the carrier gas flow through the trap while it is heated with hot water (70–

180 100 °C) to release the sample. The MS is operated in single-ion mode. The peak areas (PA) in the spectra were integrated manually with MSD Productivity ChemStation (version E.02.02.1431, Agilent, Waldbronn, Germany). The unit of PA is arbitrary and, therefore, is not used throughout the manuscript. Qualification and quantification of DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ was done using m/z ratios of 64 and 65, 72 and 73, and 77 and 79, respectively. The limit of detection (LOD, i.e., minimum PA detectable) was defined as 7 times the root mean square error of the baseline noise and was determined for each

185 compound individually from one representative chromatogram for each experiment. This yielded LODs for DMS-d3 of 209 and 90, for isoprene-d5 of 300 and 309, and for $^{13}\text{CS}_2$ of 258 and 181 respectively for experiment B and C. To extract only isotope signals, all PA were corrected for spectral fragmentation, which causes overlap in the mass spectra of natural and isotopically labelled compounds. A calibration was deemed not necessary, as we are only interested in ratios of SML trace gas content over ULW trace gas content, each sampled directly one after another. The calibration usually used with this setup

190 translates PA linearly into concentrations, which means that the ratio of PAs equals the ratio of concentrations, thus the ratio of PAs is sufficient. It takes less than 1 h to measure all samples from one set, therefore also the drift of the GC–MS is negligible for one sample set. By skipping the calibration, the number of measured SML–ULW pairs for each day was increased. Each PA was normalized by the sample volume to allow for comparability. Associated uncertainty of GC–MS measurements is 10 %. An additional step in quality control (QC) was added to identify measurements with poor quality. Peaks of mass 91 at

195 retention time $\sim$ 8.10 min were integrated for all SML, ULW and surface samples and normalized by sample volume (PA_{91}). Mass 91 is associated with toluene, as there is a stable toluene contamination in the purge system. Therefore, the variability in PA_{91} is indicative of poor-quality measurements (Appendix B Additional quality control using mass 91).

2.4 Sample analysis surface activity

Surface activity (SA) of surfactants was measured only for samples when Triton X-100 (TX-100, with molecular weight

200 625 g mol $^{-1}$, Merck, Darmstadt, Germany) was added to the tank (experiment C, see Table 1). Of those samples, only a subset



was measured, which was a representative ULW triplicate at the beginning and end of the day, as well as each SML (singlets) sample. SA was quantified using phase-sensitive alternating current voltammetry with a hanging mercury drop electrode 797 VA (Computrace, Metrohm, Switzerland), following the method established by Čosović and Vojvodić (1982, 1998). This electrochemical technique exploits the adsorption behaviour of surfactants at the interface between the mercury electrode and electrolyte, altering the capacitive current (Scholz, 2015). Samples were stored at $-20\text{ }^{\circ}\text{C}$ until analysis. Prior to measurement, all samples were brought to room temperature and adjusted to a uniform ionic strength corresponding to a S_p of 35.0 ($0.55\text{ mol L}^{-1}\text{ NaCl}$) to ensure comparability across measurements. Depending on the TX-100 concentration added to the tank, measurements were performed with a deposition time ranging between 10–60 s and a voltage sweep between -0.6 and -1.0 V . Additionally, high TX-100 concentration samples required dilution before measuring (dilutions from 0.7–0.97). For evaluation, the initial current response at -0.6 V was used, and the difference in capacity current between sample and blank was calculated with $\Delta I = I_{\text{blank}} - I_{\text{sample}}$. Three scans were recorded per replicate and the mean of the three scans was used for final analysis. Calibration was performed using TX-100 across a concentration range of $0.01\text{--}1.3\text{ mg L}^{-1}$. Only the linear portion of the response curve was used for quantification. Analytical precision was assessed using daily standards and blanks, yielding an average precision of 6 %, with all values remaining below 10 %.

For each triplicate of ULW surfactants samples the standard error (SE) of mean was calculated with Eq. (2). Only a few representative ULW samples were taken per day and then were averaged over the day. Errors were propagated with Eq. (3) to show the precision of the daily mean accounting for replicate uncertainty.

$$SE = \frac{SD}{\sqrt{n}} \quad (2)$$

$$SE_{\text{day}} = \frac{1}{m} \sqrt{\sum_{i=1}^m SE_i^2} \quad (3)$$

where SE and SE_i are the standard error of the triplicate, SD is standard deviation of the triplicate, n the number of samples per triplicate, SE_{day} is standard error of the mean of triplicates per day and m is number of triplicates per day.

Spread of triplicate means is calculated as SD of the mean values with Eq. (4).

$$SD_{\text{day}} = \sqrt{\frac{1}{m-1} \sum_{i=1}^m (\bar{x}_i - \bar{x}_{\text{day}})^2} \quad (4)$$

where SD_{day} is standard deviation for the daily mean, m the number of triplicates, $\bar{x}_i$ the mean of the triplicate and $\bar{x}_{\text{day}}$ the daily mean.

2.5 Sampling efficiency

Sampling efficiency is calculated as ratio of trace gas content per volume in the SML over content per volume in ULW (Eq. (5)). Units should be a concentration (e.g., mol L^{-1}) or linearly proportional to the resulting concentration (e.g., PA as per millilitre of sample). Sampling efficiency is given as a dimensionless fraction. The complement of this fraction represents the



sampling losses (Eq. (6)). Note, that this resembles the EF commonly computed to determine enrichment in the SML in field samples, e.g., $EF = x_{SML}/x_{ULW}$, where x is a measured property.

$$E = \frac{PA_{SML}}{PA_{ULW}} \propto \frac{C_{SML}}{C_{ULW}} \quad (5)$$

$$L = 1 - E \quad (6)$$

$$C_{SML,true} = \frac{C_{SML}}{E} \quad (7)$$

where E is sampling efficiency (unitless), PA is the peak area per millilitre of sample, C is a concentration in, e.g., mol L^{-1} , taken from the SML or ULW, as indicated by the subscripts and L are the sampling losses.

Sampling efficiency corresponds to the integrated error introduced by sampling with the glass plate, used in Eq. (7) for correction. This is based on a multiplicative error model, where the error (i.e., sampling efficiency) scales linearly with the quantity needing correction (i.e., measured concentration in glass plate samples), without an offset.

2.6 Additional measurements and recorded parameters

In experiment B temperature of water and air was measured occasionally with common thermometers. In experiment C, additional parameters were measured for each sample set. Practical salinity and temperature of the water were measured with a Cond 330i with a TetraCon 325 (WTW, Weilheim, Germany) at a depth of about 2 cm below surface. The pH of water was measured at about 1.5 cm below surface with a digital pH meter PH-100 ATC (Votcraft, Hirschau, Germany), calibrated every morning with buffer solutions with pH of 7.0 and 4.0 (Scharlab, Barcelona, Spain). Air temperature was measured with a common digital thermometer (about 46 cm above water surface) and a liquid expansion thermometer (about 15 cm above water surface, low precision of $\Delta T = 0.5^\circ\text{C}$). The number of dips, sample volume in mL (or sample weight in gram) and the person who was sampling were recorded in both experiments. In experiment C, the duration of glass plate sampling in seconds was recorded as well.

2.7 Statistical analysis

Data processing, visualization and statistical analysis was performed in Python version 3.11.11 (The Python Language Reference, 2025). The libraries used for data processing and visualization were pandas version 2.3.2 (McKinney, 2010; The Pandas Development Team, 2025), NumPy version 2.3.3 (Harris et al., 2020), seaborn version 0.13.2 (Waskom, 2021) and matplotlib version 3.10.0 (Hunter, 2007; The Matplotlib Development Team, 2024). For statistical analysis, SciPy version 1.16.2 (Virtanen et al., 2020) and statsmodels version 0.14.5 (Seabold and Perktold, 2010) were used. The main function for each statistical calculation is given below. Parameters of the function call are only mentioned if they were changed from default. Observations with missing values were only excluded if the variable concerned was used in the respective statistical analysis. Outliers were included in descriptive statistics and statistical tests, as large variation was expected and removing outliers would potentially remove true values. In boxplots, however, outliers are shown with the common 1.5 times



interquartile range (IQR) in order to present all of the data and to highlight the skewedness. Pair-wise differences between means of treatments and experiments for each trace gas were assessed with Welch's t-test (SciPy, `ttest_ind(...,equal_var=False)`), as for a few cases the heteroscedasticity could not be safely assumed (minimum and maximum variances were off by a factor of more than 3). No multiple testing correction was applied. One-way ANOVA was used to determine if purging of surfactants samples had an effect on the measurements and to identify (categorical) factors (without order) driving well-mixed ratios in Fig. Figure 3 (SciPy, `f_oneway()`). The null hypothesis was rejected at $p < 0.05$, and statistical significance assumed accordingly. Where applicable, p -value levels of <0.05 , <0.01 and <0.001 are indicated by *, **, and *** respectively. Levene's test was used to assess similarity of variances for treatments and experiments, since sample sizes per group were small ($n \leq 11$) and, consequently, non-normality had to be assumed. Linear regression was calculated with ordinary least-squares (SciPy, `linregress()`) using Eq. (8). Coefficient of determination R^2 was calculated as squared Pearson correlation coefficient. Exponential fit was fitted using Eq. (9) (SciPy, `curve_fit(...,p0=(max(y),0.1,min(y)))`). Multiple linear regression (MLR, statsmodels, `smf.ols(...).fit(cov_type='HC3')`) was used to assess which independent variables drive sampling efficiency. Data was standardized using Z-scores before applying MLR with ordinary least-squares (Eq. (10)).

$$y = mx + y_0 \quad (8)$$

$$y = ae^{bx} + c \quad (9)$$

where y is the dependent quantity, x is the independent quantity, m is the slope and y_0 is the intercept at $x = 0$ of the linear function, a is the amplitude, b is the exponential rate constant, and c is the asymptote of the exponential function.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (10)$$

where Y is the response variable, β_0 is the intercept, β_i ($i \geq 1$) are the coefficients determined by the MLR, X_i are the independent variables and ε is the error term. Reported are R^2 and adjusted R^2 , which accounts for the number of predictors vs samples size, and p -value for F -statistics. Heteroscedasticity was not assumed.

3 Results

A total of 155 samples were taken (53 SML–ULW pairs), of which 149 samples (51 SML–ULW pairs) satisfied formal QC. An additional 13 samples (5 SML–ULW pairs) were removed, as our experimental assumptions were not satisfied, leaving a total of 46 SML–ULW pairs for sampling efficiency calculations.

The results presented here are based exclusively on experiment B and C, as experiment A were preliminary tests. Furthermore, our assumption that the tank was well-mixed is tested. PA of the trace gases for SML and ULW and their correlation are presented, as this is the basis for our choice of a correction procedure. Final sampling efficiency is shown for water without additives first, and then including surfactants.



280 3.1 Experimental conditions and parameters

3.1.1 Water temperature and salinity

The overall water temperature range across all experiments and treatments was 17.2–23.1 °C. Water temperature was similar in both treatments in experiment B. The temperature in the FW treatment was only measured once at the end of the day (16:30 UTC, 19 April), yielding 20.9 °C. Having been filled the previous day and stored under the SURF roof overnight, the water
 285 likely warmed over the course of the day, reaching 20.9 °C at the time of measurement. The AS treatment on the next day started off warmer and cooled down over the course of the day, averaging to 21.5 °C ($\Delta T = 4.6$ °C). In experiment C (located in lab), a larger range of water temperatures was targeted by using refrigerated water that would heat up in the course of the day to room temperature. This was only partially successful, as the targeted water temperature of <10 °C to start with was not achieved. The FW treatment averaged at 17.6 °C, AS at 19.5 °C and the treatment with artificial surfactants at 20.6 °C (with
 290 ΔT of 1.2, 3.0 and 1.5 °C respectively). Water temperature of FW treatment overlaps partially with the AS treatment, and the AS treatment partially with the treatment including surfactants, allowing for continuity across treatments.

Salinity was set to either zero or to oceanic levels. S_P was between 28.9 and 34.6 in all treatments with salt amendments. Salinity was kept constant for each treatment level (FW, AS, and each of the SA levels), with exception of the AS treatment without surfactants, which combines two lab days. On 8 August the targeted practical salinity of 34.5 was not reached, but
 295 29.0 ± 0.07 . This was corrected to 34.6 on 10 August 2023.

3.1.2 Surface activity and enrichment of surfactants

Artificial surfactants were added in experiment C to AS only. Since GC–MS samples were reused for voltammetry analysis, it was tested, if the delayed freezing and the purging had an effect on measured SA. Comparison was performed on a specially collected sample set ($n = 6$) and across all ULW samples with voltammetry measurements ($n = 36$). Neither purging (one-way
 300 ANOVA failed to reject null hypothesis, $p_{all} = 0.8864$ with $n = 36$, $p_{subset} = 0.5138$ with $n = 6$) nor the delayed freezing (one-way ANOVA failed to reject null hypothesis, $p_{all} = 0.3788$ with $n = 30$) had a significant effect. SA of purged and non-purged samples is thus treated the same (data points shown in Appendix C).

The goal was for the surfactants to reach $EF = 1$ after wet surfactant TX-100 was added to AS. We targeted a range of SA, so on each day a different amount was added to the tank, increasing over time (Table 2). Mean SA_{ULW} ranged from 0.160 to
 305 1.515 mg L^{-1} TX-100 equiv., while mean SA_{SML} was much higher between 2.144 to 8.244 mg L^{-1} TX-100 equiv., consequently the SML was enriched on average EF from 3.7–11.7 (with three exceptions where $EF < 1.0$, once for each level of TX-100 added). SA_{ULW} was usually pretty consistent for one added amount of TX-100 (standard deviation $< 0.241 \text{ mg L}^{-1}$ TX-100 equiv.), whereas SA_{SML} varied much more with standard deviations from 1.226 to 3.303 mg L^{-1} TX-100 equiv., which in turn cause high standard deviation in the EF as well.

310



Table 2 Surface activity (SA) measurements and corresponding enrichment factors (EF) in experiment C per day in the laboratory in treatment artificial seawater (AS) with surfactants. Each day a different amount of TX-100 was added, increasing over time. For $n = 2$ minimum and maximum values are depicted instead of mean and standard deviation.

lab day	SA								
	ULW			SML			Enrichment factor		
	[mg L ⁻¹ TX-100 equiv.]			[mg L ⁻¹ TX-100 equiv.]					
2023	<i>n</i>	mean ± SD	(median)	<i>n</i>	mean ± SD	(median)	<i>n</i>	mean ± SD	(median)
10 August	2	0.136; 0.185	(0.160)	4	1.876 ± 1.226	(2.144)	4	11.7 ± 7.6	(13.4)
28 August	4	0.762 ± 0.099	(0.774)	6	2.785 ± 1.139	(3.211)	6	3.7 ± 1.5	(4.2)
31 August	3	1.442 ± 0.241	(1.515)	5	6.776 ± 3.303	(8.244)	5	4.7 ± 2.3	(5.7)
all	9	0.855 ± 0.522	(0.778)	15	3.873 ± 2.926	(3.187)	15	6.1 ± 5.2	(5.0)

3.1.3 Sampling duration, sample volume and number of dips

The duration of sampling was recorded for SML samples from right before the first dip until the vial was closed with the rubber stopper as a potential driver for sampling efficiency variability. Sampling duration ranged from 54–235 s over all experiments. Only a few durations were recorded in experiment B in the AS treatment, amounting to 135 s on average. In experiment C, the shortest sampling duration was recorded for the AS treatment with a mean of 112 s. FW sampling took longer with 145 s on average. While the SA was increased each day, the sampling duration decreased from 183 s initially (10 August), to 139 s on the second day (28 August) and 129 s on the last day (31 August). Sampling duration greatly depends on the number of dips, which in turn depends on the targeted sample volume. Sampling the treatment with surfactants overall took longer to sample than the AS treatment, because 15 mL were targeted instead of 10 mL for AS.

ULW samples were taken with a pipette, resulting in exact sample volumes of 10 mL and 15 mL respectively. SML samples, on the other hand, showed variation depending on how much water would stick to the glass plate per dip. FW and AS samples were mostly slightly larger than the targeted 10 mL (10.8 ± 1.0 , $n = 30$). The samples with the lowest TX-100 added were mostly slightly less than the targeted 15 mL (14.3 ± 0.7 , $n = 4$). For the next increase of TX-100, sample volumes were mostly slightly more than the targeted 15 mL (15.9 ± 1.0 , $n = 6$). For the highest concentration of TX-100, all samples were below the targeted 15 mL (12.5 ± 1.3 , $n = 6$). This is linked to how much sample volume can be gathered per dip (Fig. Figure 1, Fig. Figure 2). In the treatment with the highest TX-100 concentration, another (full) dip would have resulted in too much volume for GC–MS analysis (i.e., not enough headspace). Note, that on 28 August (medium concentration of TX-100) also



half dips were captured ($n = 2$), i.e., only the wipe of the first side of the glass plate was filled into the vial. The second side was discarded.

The number of dips was recorded to calculate the operational SML thickness (Eq. (1)) and as a potential, additional source of sampling efficiency variability besides sampling duration. To fill about 10 mL of sample without artificial surfactants, three to five dips were required, whereas two to four dips were required to fill about 15 mL with artificial surfactants added (Fig. Figure 1). FW and AS are similar, also across experiments (FW and AS in experiment B 2.4 and 2.9 mL per dip, in experiment C 3.3 and 3.2 mL per dip respectively). In the treatment with artificial surfactants significantly more volume (5.1 mL per dip) was collected per dip, on average, while the volume collected per dip increases with the amount of TX-100 present in the SML (measured as SA), as shown in Fig. Figure 2.

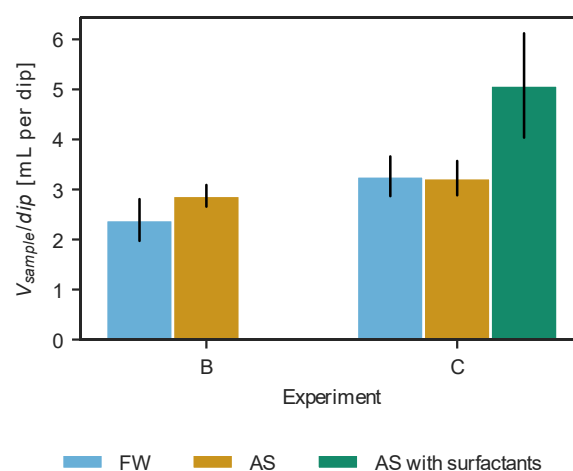


Figure 1 Collected sample volume V_{sample} for SML samples per dip of the glass plate in the three treatments. Black lines indicate standard deviation.

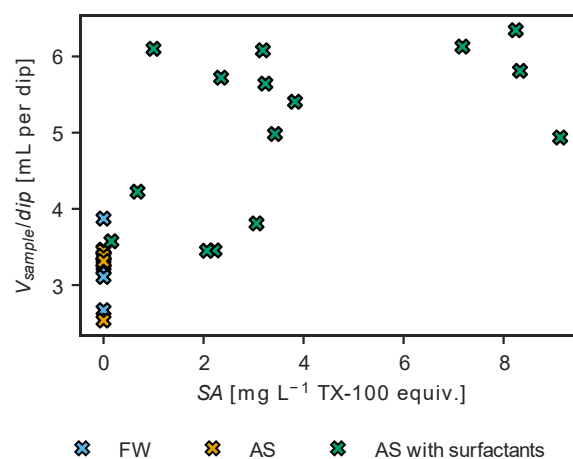


Figure 2 Sample volume collected per dip in experiment C with the glass plate increases with SA.



An increased sample volume per dip indicates that a thicker layer was sampled (Eq. (1)). Layer thickness ranged from 25–63 μm . In experiment B, FW and AS differ at $40 \pm 7 \mu\text{m}$ and $48 \pm 4 \mu\text{m}$, while the sampled layer was thinner and more similar in FW and AS in experiment C, with $33 \pm 4 \mu\text{m}$ and $32 \pm 3 \mu\text{m}$ respectively. In the treatment with artificial surfactants, the thickness increases to an average of $51 \pm 10 \mu\text{m}$.

3.2 Mixing in the tank

Our experiments are based on the premise that SML and ULW have the same trace gas concentration, achieved by mixing the tank before each sampling. In order to test our assumption, a reference ULW sample (PA_{ref}) was taken and compared to samples from other depths and locations in the tank (Fig. Figure 3). PA was either taken at reference depth, but at a different location, or at the surface. Almost all ratios (PA/PA_{ref}) lie within the range of uncertainty of the measurements (grey shaded area) for all three trace gases. Four sample sets (indicated as grey points) were removed from further calculations (i.e., sampling efficiency), as the additional QC did not indicate poor quality, nor were there deviations from the general sampling procedure noted in the protocols. Four points for isoprene-d5 and three for $^{13}\text{CS}_2$ from 20 April and 08 August were outside the uncertainty range, but were not removed from the data set. Mean values of the ratios (excluding grey points) are 1.00 ± 0.05 , 1.01 ± 0.10 , and 1.01 ± 0.08 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ respectively. Standard deviations vary between trace gases, with the largest variation for isoprene-d5. Ratios are independent of sampling location, depths, day or experiment (one-way ANOVA failed to reject null hypothesis in all cases).

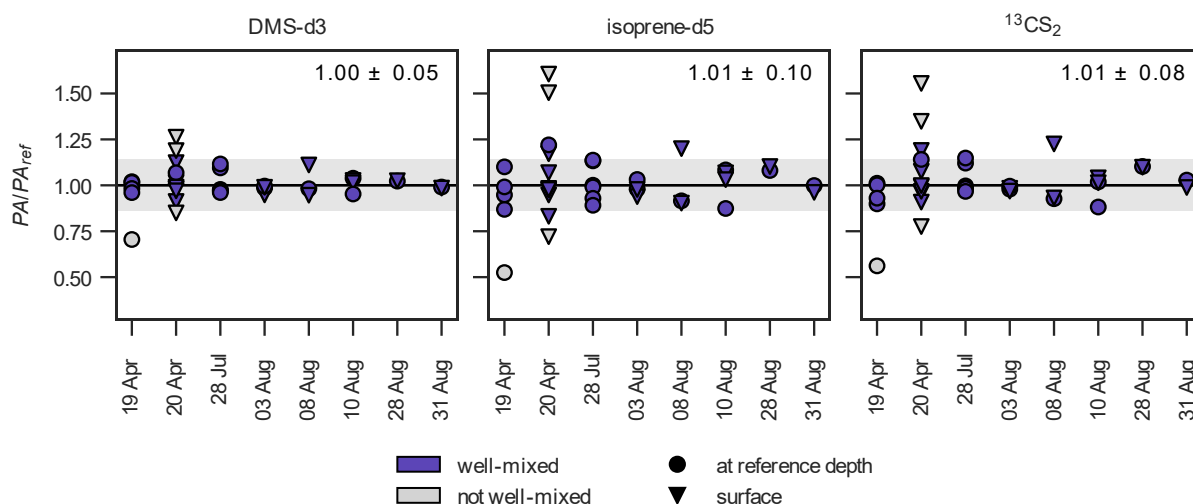
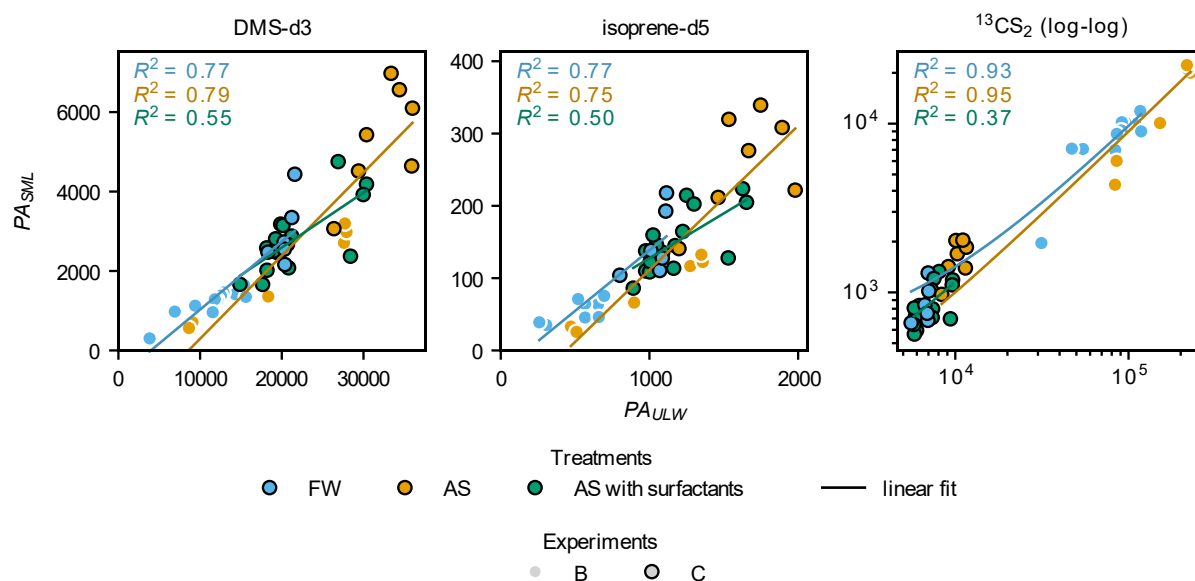


Figure 3 Test for uniform trace gas concentration in the tanks (well-mixed). The ratios PA/PA_{ref} ($n = 37$) are shown against date for DMS-d3 (left), isoprene-d5 (middle) and $^{13}\text{CS}_2$ (right). Black horizontal line depicts $PA = PA_{ref}$ (i.e., well-mixed). Grey shaded area indicates $\pm 14.1\%$ uncertainty, calculated by error propagation. Grey points ($n = 4$) were identified as not well-mixed. Mean and standard deviation for PA/PA_{ref} is indicated in upper right of each subplot (excluding grey points). April samplings belong to experiment B, all other samplings are experiment C. All three treatments (FW, AS, AS with surfactants) are included.

3.3 Measurements of trace gases



370 **Figure 4 Correlation of SML and ULW PA for DMS-d3 (left), isoprene-d5 (middle) and $^{13}\text{CS}_2$ (right), grouped by treatments fresh water (FW) and artificial seawater (AS) without and with surfactants. Linear fits for $^{13}\text{CS}_2$ were performed on the non-logged data, log-log scale is only used to better visualize the very low values.**

Figure 3 shows measured PA from treatments FW, AS and AS with surfactants, plotted as SML against ULW, for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$. SML is always much lower than ULW. The order of magnitude for PA varies between trace gases due to different amounts of trace gas in the spike. DMS-d3 is in the order of 10^4 , isoprene-d5 at 10^3 and $^{13}\text{CS}_2$ up to 10^5 . $^{13}\text{CS}_2$ shows two distinct clusters of points, as the two spikes used were significantly different in the amount of $^{13}\text{CS}_2$, whereas the amounts for DMS-d3 and isoprene-d5 were similar. For all three gases, a clear linear, positive correlation is visible between SML and ULW. R^2 of the linear fit (SciPy, linregress()) for DMS-d3 and isoprene-d5 are both 0.77 for FW, 0.79 and 0.75 for AS, and much lower at 0.55 and 0.50 for AS with surfactants respectively. R^2 for $^{13}\text{CS}_2$ is 0.93 for F and 0.95 for AS, and much lower than for the other trace gases for AS with surfactants at 0.37. Note, that the latter data points all fall exclusively into one cluster only, indicating a bad fit without the influence of the large spread present for FW and AS. These values, together with slope and intercept as well as mean PA with standard deviation are reported in Table 3. Mean values and standard deviation visualize the order of magnitude, as within each treatment the amount of spike added differed slightly, causing high standard deviation. This is especially visible for $^{13}\text{CS}_2$, where standard deviation is close to the mean value for FW and AS (includes both clusters), but standard deviation is much lower than the mean for AS with surfactants (only lower value cluster). Slopes differ slightly between gases and treatments. Slopes for DMS-d3 are 0.17 (FW), 0.21 (AS) and 0.14 (AS with surfactants). Slopes for isoprene-d5 are similar, with 0.17 (FW), 0.20 (AS) and 0.12 (AS with surfactants). Slopes for $^{13}\text{CS}_2$ are lower, with 0.09 (FW and AS) and 0.10 (AS with surfactants). Intercepts are not zero, but all are much smaller than the respective mean



PA_{ULW} and therefore considered negligible. Correlation of all data points without distinguishing treatments yields R^2 of 0.80 (DMS-d3), 0.76 (isoprene-d5) and 0.96 ($^{13}CS_2$).

Table 3 Linear regression of PA of SML and ULW samples for DMS-d3, isoprene-d5, and $^{13}CS_2$. All R^2 show a strong fit, except when marked with * (moderate fit).

Trace gas	Treatment	n	PA				
			ULW	SML	Linear Regression (OLS)		
			mean $\pm$ SD		slope	intercept	R^2
DMS-d3	FW	17	14606 $\pm$ 5240	1817 $\pm$ 1019	0.17	−679	0.77
	AS	13	26547 $\pm$ 9157	3755 $\pm$ 2153	0.21	−1796	0.79
	AS with surfactants	16	21559 $\pm$ 4685	2805 $\pm$ 873	0.14	−175	0.55*
	all	46	20399 $\pm$ 7970	2708 $\pm$ 1575	0.18	−891	0.80
isoprene-d5	FW	16	719 $\pm$ 282	91 $\pm$ 54	0.17	−29	0.77
	AS	13	1334 $\pm$ 475	178 $\pm$ 109	0.20	−88	0.75
	AS with surfactants	16	1183 $\pm$ 238	150 $\pm$ 42	0.12	4	0.50*
	all	45	1062 $\pm$ 422	137 $\pm$ 78	0.16	−34	0.76
$^{13}CS_2$	FW	17	56749 $\pm$ 44288	5727 $\pm$ 4236	0.09	500	0.93
	AS	13	81274 $\pm$ 90906	7326 $\pm$ 8266	0.09	114	0.95
	AS with surfactants	16	7159 $\pm$ 1385	865 $\pm$ 235	0.10	128	0.37*
	all	46	46431 $\pm$ 61976	4488 $\pm$ 5674	0.09	331	0.96

3.4 Sampling efficiency for fresh water and artificial seawater

The sampling efficiency of the trace gases is calculated using the ratio of measured PA in SML over measured PA in ULW (Eq. (5)). Due to mixing in the tank the theoretical maximum value for the sampling efficiency is, thus, 1.0 (or 100 %), i.e., when $PA_{SML} = PA_{ULW}$.

Figure 5 shows sampling efficiency for the studied trace gases in FW and AS treatment, separated by experiment. Overall the three trace gases are similar, but there are significant differences between the experiments (fails to reject null hypothesis with $p < 0.001$, Table 5). In experiment B sampling efficiency is lower than in C. For all three trace gases sampling efficiency decreases in experiment B with the addition of salt, whereas, this reverses for experiment C and sampling efficiency increases with the addition of salt.

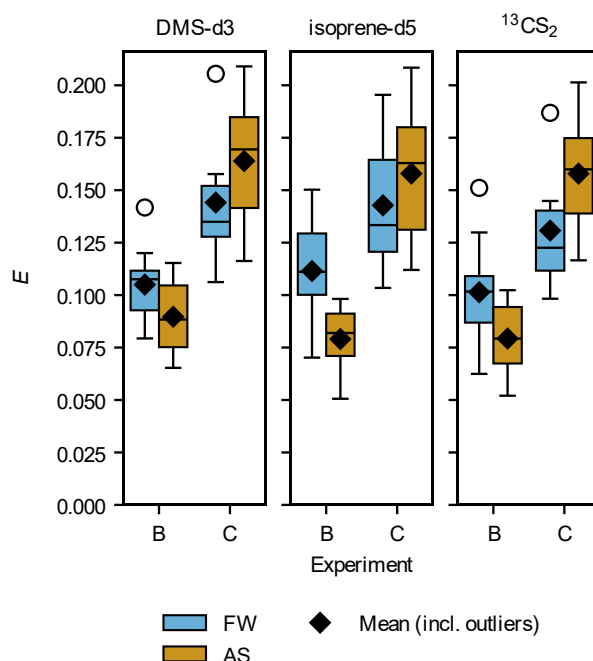


Figure 5 Sampling efficiency for DMS-d3 (left), isoprene-d5 (middle) and $^{13}\text{CS}_2$ (right), separated by experiment, coloured by treatments fresh water (FW, blue) and artificial seawater (AS, orange). Box width shows 25 to 75 percentile, centre line denotes the median, diamonds denote the mean, open circles denote outliers, and whiskers show minimum and maximum. The number of data points in experiment B is $n_{FW} = 11$ ($n_{FW} = 10$ for isoprene-d5) and $n_{AS} = 6$, in experiment C $n_{FW} = 6$ and $n_{AS} = 7$.

Mean sampling efficiency for DMS-d3 in experiment B is 0.105 ± 0.018 ($n = 11$) in FW, decreasing to 0.090 ± 0.020 ($n = 6$) in AS, though the difference is not significant ($p = 0.15$, outliers in Fig. Figure 5 included). In experiment C, the mean sampling efficiency for DMS-d3 in both treatments is higher than in experiment B, with 0.144 ± 0.034 ($n = 6$) in FW increasing to 0.164 ± 0.033 ($n = 7$) in AS. The difference between FW and AS within experiment C is not significant ($p = 0.32$). This overall picture repeats for isoprene-d5 and $^{13}\text{CS}_2$. In experiment B, mean sampling efficiency for isoprene-d5 is 0.112 ± 0.025 ($n = 10$) in FW, reducing to 0.079 ± 0.018 ($n = 6$) in AS, whereas in experiment C sampling efficiency is overall higher, with 0.143 ± 0.035 ($n = 6$) in FW and then further increasing to 0.158 ± 0.036 ($n = 7$) in AS. For $^{13}\text{CS}_2$, sampling efficiency in experiment B averages to 0.101 ± 0.025 ($n = 11$) in FW and lowers to 0.079 ± 0.019 ($n = 6$) in AS, whereas it is overall higher in experiment C with 0.131 ± 0.032 ($n = 6$) in FW and rising to 0.158 ± 0.031 ($n = 7$) in AS. Similar to DMS-d3, the differences within both experiments between treatments are not significant (Table 5), except for isoprene-d5 in experiment B ($p = 0.009$). Standard deviations of sampling efficiency are slightly larger in experiment C than in B for all three trace gases and differ significantly between treatments FW and AS (Levene's test for equal variances failed to reject null hypothesis with p 0.045, 0.0497 and 0.033 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$, SciPy, `levne(...,centre='mean')`). Standard errors are almost two times larger in experiment C than in B. SE in FW and AS for DMS-d3 is 0.0074 and 0.013, for isoprene-d5 is 0.0080 and 0.014 and for $^{13}\text{CS}_2$ is 0.0073 and 0.013. SE overall ($n = 30$) are 0.0070, 0.0074 and 0.007 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$.



There is no significant difference between FW in experiment B and C for any of the three trace gases (except for DMS-d3, $p = 0.038$), but in AS the differences are significant for all three trace gases ($p < 0.001$), see Table 5. The medium used in FW was ultrapure water, from different devices, but same models. The media used for AS, on the other hand, were different in the composition of the salts and other substances added, though also S_p slightly varied.

Table 4 Summary statistics of sampling efficiency for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ in experiment B and C in fresh water (FW) and artificial seawater (AS) without and with surfactants treatment, as visualized in Fig. Figure 5 and Fig. Figure 6. † indicates in which group one data point had to be removed for isoprene-d5 ($PA_{SML} < \text{LOD}$).

Summary Statistics								
Experiment	Treatment	n	DMS-d3		isoprene-d5		$^{13}\text{CS}_2$	
			mean $\pm$ SD	(median)	mean $\pm$ SD	(median)	mean $\pm$ SD	(median)
B	FW	11 [†]	0.105 $\pm$ 0.018	(0.108)	0.111 $\pm$ 0.025	(0.111)	0.101 $\pm$ 0.025	(0.102)
	AS	6	0.090 $\pm$ 0.020	(0.088)	0.079 $\pm$ 0.018	(0.082)	0.079 $\pm$ 0.019	(0.079)
C	FW	6	0.144 $\pm$ 0.034	(0.135)	0.143 $\pm$ 0.035	(0.133)	0.131 $\pm$ 0.032	(0.123)
	AS	7	0.164 $\pm$ 0.033	(0.169)	0.158 $\pm$ 0.036	(0.163)	0.158 $\pm$ 0.031	(0.160)
	AS with surfactants	16	0.130 $\pm$ 0.025	(0.131)	0.127 $\pm$ 0.024	(0.124)	0.121 $\pm$ 0.024	(0.120)

Sampling efficiency as a function of water temperature, salinity, spike volume per litre added (i.e., proportional to trace gas concentration) and number of dips was investigated using MLR models ($n = 19$ complete observations of 30). Note, that 10 of 11 data points in experiment B treatment FW were excluded, because of missing temperature measurements, and one of seven in experiment C treatment AS, because of unrecorded number of dips. The MLR models were significant for all three trace gases ($F(4,14) > 5.6$, $p < 0.01$) and explain about 60 % of the variance observed in sampling efficiency (adjusted R^2 is 0.63, 0.60 and 0.65 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ respectively). The models predicted average ($\pm$ SE) sampling efficiency of 0.129 ± 0.007 , 0.127 ± 0.008 and 0.120 ± 0.007 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$, when all predictors were at their mean values ($p < 0.001$). Spike volume per litre is a significant, positive predictor for sampling efficiency for all three trace gases ($p < 0.01$, β is 0.0342, 0.0404 and 0.0353 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$, $s_{\text{spike}} = 1.4 \mu\text{L L}^{-1}$). For DMS-d3 and $^{13}\text{CS}_2$ also water temperature is a significant, but negative predictor ($p < 0.05$, β is -0.0223 and -0.0219 , $s_{T_w} = 1.8 \text{ }^\circ\text{C}$), whereas, for isoprene-d5 the number of dips is a significant, positive predictor ($p = 0.018$, $\beta = 0.0213$, $s_{\text{dips}} = 0.6$). Salinity was not a significant predictor for any of the trace gases (p between 0.448 and 0.608). See Appendix D Results linear and multiple linear regressions in fresh water and artificial seawater from Sect. 3.4 for details of MLR, linear regression results and additional plots.



Table 5 Test statistics of sampling efficiency for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ in experiment B and C in fresh water (FW) and artificial seawater (AS) treatment, as visualized in Fig. Figure 5. [†] indicates in which group one data point had to be removed for isoprene-d5 ($PA_{SML} < \text{LOD}$). *, ** and *** mark where the null hypothesis was rejected, with p -values <0.05 , <0.01 and <0.001 respectively.

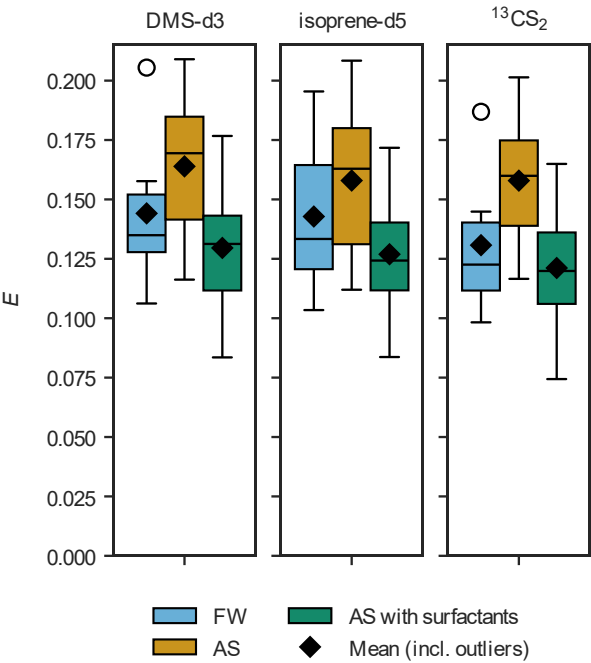
			Welch's t-test		
Groups		n	DMS-d3 p -value	isoprene-d5 p -value	$^{13}\text{CS}_2$ p -value
B		17 [†]	<0.001***	<0.001***	<0.001***
C		13			
FW		13 [†]	0.48	0.92	0.52
AS		17			
B	FW	6 [†]	0.15	0.009**	0.063
	AS	11			
C	FW	7	0.32	0.46	0.15
	AS	6			
FW	B	6 [†]	0.038*	0.090	0.083
	C	11			
AS	B	7	<0.001***	<0.001***	<0.001***
	C	6			

3.5 Sampling efficiency with artificial surfactants

To mimic in-situ conditions more closely, the experiments were extended to include a treatment of AS with surfactants. When artificial surfactants (TX-100) were added to AS (Fig. Figure 6), the mean sampling efficiency for all of the three trace gases reduced even below the FW treatment, DMS-d3 to 0.130, isoprene-d5 to 0.127 and $^{13}\text{CS}_2$ to 0.121 (Table 4). Standard deviations are also smaller than in FW and AS treatment (0.025 for DMS-d3, 0.024 for isoprene-d5 and $^{13}\text{CS}_2$). The range (difference from minimum to maximum) is still similar to the range observed in FW and AS for each of the three trace gases. Sampling efficiency decreases with increasing SA_{SML} (Fig. Figure 7). It decreases fast from $SA = 0$ until about $SA = 3.0$. Above $SA = 3.0$ the values seem to arrange around a constant value of sampling efficiency. Linear fit is weak (R^2 between 0.26 and 0.28 for all three trace gases). The slope is negative and small with about -0.005 per mg L^{-1} TX-100 equiv. and intercepts are between 0.146 to 0.154 for the three trace gases (



460 Table 6). Though small, the slope is significantly different from no slope ($p < 0.05$ for all three trace gases). An exponential fit was only possible for DMS-d3 and $^{13}\text{CS}_2$, for which it explains slightly more variability than the linear fit (R^2 between 0.42 and 0.44). When treatment FW was included exponential fit was only possible for DMS-d3 (not shown). The exponential fit approaches 0.12 and 0.11 for DMS-d3 and $^{13}\text{CS}_2$ with a small amplitude of 0.05 for both at a decay rate of -0.68 and -0.88 . The exponential fit is added for visualization in Fig. Figure 7.



465 **Figure 6** Sampling efficiency for DMS-d3 (left), isoprene-d5 (middle) and $^{13}\text{CS}_2$ (right) when artificial surfactants are added to artificial seawater (AS with surfactants) from experiment C ($S_p = 31.1 \pm 2.9$). Fresh water (FW) and artificial seawater without surfactants (AS) are depicted as well for reference. Box width shows 25 to 75 percentile, centre line denotes the median, diamonds denote the mean, open circles denote outliers, and whiskers show minimum and maximum. The number of data points is $n_{FW} = 6$, $n_{AS} = 7$ and $n_{SA} = 16$.

470

Table 6 Curve fitting for sampling efficiency against SA_{SML} in experiment C, treatments with AS only.

Trace gas	n	slope $\pm SE$	Linear Fit			Exponential Fit	
			intercept $\pm SE$	R^2	p -value	fitted function	R^2
DMS-d3	22	-0.0054 ± 0.002	0.154 ± 0.008	0.26	0.015*	$0.048e^{-0.68x} + 0.12$	0.42
isoprene-d5	22	-0.0055 ± 0.002	0.151 ± 0.008	0.28	0.012*	fit not possible	
$^{13}\text{CS}_2$	22	-0.0054 ± 0.002	0.146 ± 0.008	0.27	0.013*	$0.047e^{-0.88x} + 0.11$	0.44

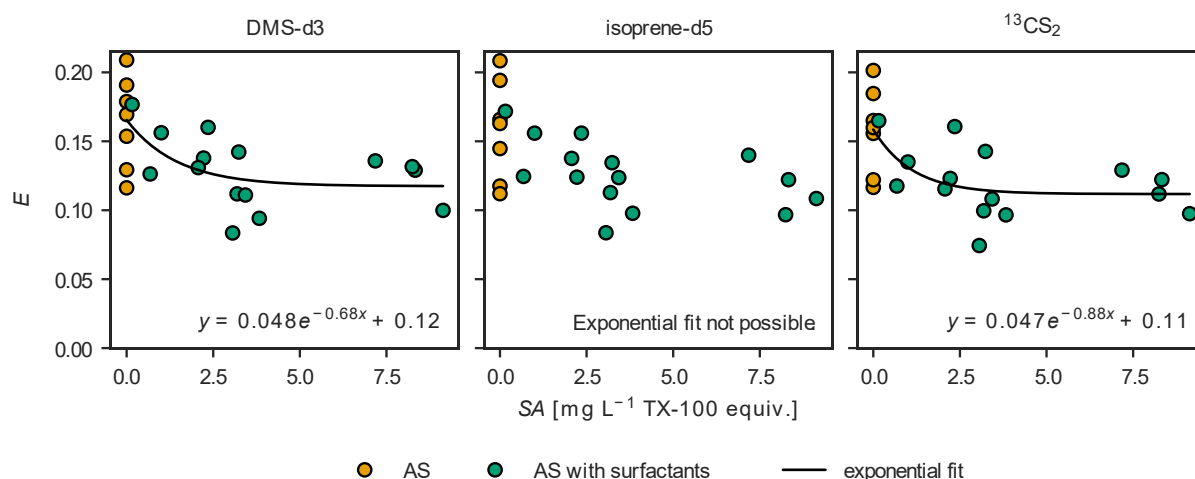


Figure 7 Sampling efficiency for DMS-d3 (left), isoprene-d5 (middle) and $^{13}\text{CS}_2$ (right) against surface activity (SA) in experiment C. Exponential fit is added as aid for visualization.

Sampling efficiency as response to water temperature, salinity, spike volume per litre added (i.e., proportional to trace gas concentration), number of dips and SA in the SML was investigated using MLR models for each trace gas ($n = 27$ complete observations of 29). The MLR models were not significant for any of the three trace gases ($F(5,21) > 1.9$, $p > 0.09$) and explained less than 25 % of the variance observed in sampling efficiency (adjusted R^2 is 0.25, 0.22 and 0.23 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ respectively). The model predicted average ($\pm SE$) sampling efficiency of 0.139 ± 0.006 , 0.137 ± 0.006 and 0.131 ± 0.006 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$, when all predictors were at their mean values ($p < 0.001$). Spike volume per litre is a significant, positive predictor for sampling efficiency for only DMS-d3 anymore ($p = 0.029$, $\beta = 0.0213$, $s_{\text{spike}} = 0.94 \mu\text{L L}^{-1}$). Instead, water temperature is now a significant, still negative predictor for all three trace gases ($p < 0.05$, β is -0.0324 , -0.0309 and -0.0321 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$, $s_{T_w} = 1.4^\circ\text{C}$). SA in the SML was a significant, negative predictor for $^{13}\text{CS}_2$ only ($p = 0.041$, $\beta = -0.0137$, $s_{SA,SML} = 2.91 \text{ mg L}^{-1}$ TX-100 equiv.), though DMS-d3 is on the limit ($p = 0.05$), while isoprene-d5 is non-significant ($p = 0.07$). Neither number of dips ($p > 0.8$), nor salinity (p is 0.247 and 0.206 for DMS-d3 and isoprene-d5, but much lower at 0.066 for $^{13}\text{CS}_2$) were a significant predictor for any of the trace gases. See Appendix E Results linear and multiple linear regressions with artificial surfactants from Sect. 3.5 for details of MLR, linear regression results and additional plots.

4 Discussion

4.1 Experimental setup

Water temperature in this study covered a range of $\Delta T = 5.9^\circ\text{C}$. Sea surface temperature in the field encompasses a wider range. At Boknis Eck station, for example, the range studied here is only observed during the summer season (Laß et al., 2013).



In our treatments, intermediate salinity was excluded, focussing on fresh water and seawater. We do not expect to see any improvement in our results by including intermediate salinity, as salinity was a non-significant predictor for sampling efficiency. However, the choice of salt (Tropic Marin® Pro-Reef Sea Salt vs ordinary aquarium salt) may have had an effect, which is attributed to either the composition of the salt (affecting the solubility, e.g., (Weisenberger and Schumpe, 1996)), other substances or the alkalinity of the water, rather than the salinity. The increase in targeted sample volume from 10 mL (FW, AS) to 15 mL (AS with surfactants) did not affect PA_{ULW} . However, we could not assess whether PA_{SML} differed between 10 and 15 mL. The increase in volume is achieved by dipping the glass plate more often, potentially affecting the sampling efficiency. Since the number of dips was a non-significant predictor for sampling efficiency in the MLR, the effect on PA_{SML} and consequently on sampling efficiency, is considered negligible. Finally, a range of saturation states was not investigated. We only performed experiments with oversaturated conditions in the water, but we expect that the concentration gradient (magnitude and direction) has an effect on sampling efficiency.

4.1.1 Mixing in the tank

Figure 3 shows that the tank was well-mixed on all days, because almost all the ratios PA/PA_{ref} are within the acceptable uncertainty range. The four points from 20 April and 08 August, which were not acceptable for isoprene-d5, and the three for $^{13}CS_2$, are still considered as well-mixed, since the ratios are acceptable for DMS-d3. The mean values of the ratios (excluding grey points) are all close to 1.0 (i.e., $PA = PA_{ref}$), supporting that the tank was well-mixed without any physically driven gradients present or forming, as then the mean would likely show an offset. This is further supported by the ratios with surface samples on 20 April, sampled as close as possible to the surface. Due to the oversaturation in the tank and near-zero atmosphere, we hypothesize that $PA_{surface} < PA_{ULW}$ should develop as result of the outgassing, causing ratios with surface samples to systematically drop below 1.0. However, in Fig. Figure 3 they distribute equally around 1.0. The increased standard deviation of isoprene-d5 as compared to the other two trace gases, can be attributed to increasing integration error as raw, non-normalized PA of isoprene-d5 were generally close to the LOD, whereas, DMS-d3 and $^{13}CS_2$ were usually sufficiently above the LOD for the integration error to diminish. Our assumption may not hold for treatment AS with surfactants, see Sect. 4.1.3 Influence of surfactants addition.

4.1.2 Dilution by diffusive boundary layer

The concentration gradient between near-zero atmosphere for the studied trace gases and the trace gas concentration in the tank drives diffusive losses through the water surface (i.e., outgassing), creating a gradient in the water towards the surface. As a result, a diffusive boundary layer (DBL) forms at the surface. The gradient defines the thickness of DBL, which dilutes our glass plate sample. After mixing had stopped, the DBL had up to 120 s to develop with a corresponding thickness of up to 2.1 mm. Accordingly, in treatments without surfactants, the dilution by the DBL reduced the PA sampled by the glass plate to about 0.79 of the concentration in the ULW in both experiments. In order to calculate the diffusive losses, several conservative assumptions had to be made. We, therefore, expect that the actual diffusive losses are even smaller. The time for the DBL to



develop is effectively shorter, since the turbulent motion from mixing took about 1 min to subside to a level when diffusion starts to dominate. This also causes the DBL thickness to stay thinner. Additionally, the geometry of the tank limits the maximum thickness of the DBL, whereas in our calculation we assume a monotonic increase of the thickness over time. For our setup a thickness of 0.5 mm is more reasonable. This is still much larger than the thickness that was sampled by one dip of the glass plate, though with the limited tank surface it might be more appropriate to compare the DBL thickness with the sample volume distributed over the full tank surface, which yields that about 0.1 mm of the surface that is sampled away by one SML sample. Note, that the 0.79 reflects the concentration in the DBL at the instant of time right before the first dip of the glass plate. Our calculations are not valid anymore when the sampling disturbs the tank. The box model used is explained in detail in Appendix F.

The estimate of the concentration including diffusive losses is much higher than our sampling efficiency, which indicates that the DBL dilution is not dominating our sampling efficiency estimates. This is especially true, since the diffusive losses are overestimated under the assumptions we made in our calculations.

4.1.3 Influence of surfactants addition

The decrease of sampling efficiency with increasing TX-100 concentration (Fig. Figure 7) in our lab experiments is unexpected. It is hypothesized that air-sea gas exchange of trace gases decreases with increasing SA_{SML} , resulting from (1) the wave dampening effect of surfactants that in turn reduces the turbulence, and, thus, the turbulence driven gas exchange, and (2) the surfactants enriched layer acting as a physicochemical barrier (Garbe et al., 2014, 2.2.7 Surface Films). The decrease in gas exchange is a net effect of these two factors and depends on SA. Although there was no wind on our lab experiments, some turbulence was introduced by mixing and sampling, leading to the hypothesis that sampling efficiency would increase with SA. It should be noted within this discussion that natural slicks have been related to exceeding a threshold of 1 mg L^{-1} TX-100 equiv. up to and beyond 3 mg L^{-1} TX-100 equiv. (Barthelmeß and Engel, 2022). 12 SML samples were $>1 \text{ mg L}^{-1}$ TX-100 equiv. and 9 were $>3 \text{ mg L}^{-1}$ TX-100 equiv. (of 15), indicating that we reached beyond the natural range of SA in the SML. Natural surfactant pools consist of a mixture of substances, which vary greatly depending on the biological and chemical processes on-going in the SML and in their physicochemical effect (Engel et al., 2017). Though TX-100 as a wet surfactant is considered a good choice for a model surfactant, it is not ideal. Adenaya et al. (2021), for example, show that DBL thickness with TX-100 is up to 30 % thicker than with natural surfactants for $SA = 2 \text{ mg L}^{-1}$ TX-100 equiv., therefore, diluting the concentration on the glass plate. Furthermore, we could not rule out a change in the trace gas solubility or diffusivity in surfactants instead of (sea)water. Insufficient purge time caused by a potential decrease in purge efficiency with increasing SA (i.e., mostly affecting PA_{SML}) was ruled out as reason for the decrease in sampling efficiency. Also micelle formation was ruled out, as SA_{SML} was much lower than critical micelle concentration (Mukerjee and Mysels, 1971).

We did observe that surfactants were not well-mixed (EF on average >1.0 , Table 2). At about $SA = 1.5 \text{ mg L}^{-1}$ TX-100 equiv. the surface of the water body is covered 100 % with TX-100, beyond which the coverage cannot increase more. Any excess TX-100 mixes further into the ULW, increasing the SML thickness (personal communication Falko Asmussen-Schäfer). EF



of surfactants for the same level of TX-100 added in our experiments varied strongly, including $EF < 1.0$ for three cases, indicating that the mixing and re-establishment of a TX-100 gradient was not consistent. There is a slight trend towards lower EF with increasing SA (Table 2), which is related to the fact that with increasing SA_{ULW} relatively less surfactants go to the surface (personal communication Falko Asmussen-Schäfer). Furthermore, sampling during this treatment leads to depletion of surfactants. The heterogeneous, non-steady SA likely complicates the mixing of the trace gases in the tank, as well as their interaction with the surfactants.

Since we cannot definitively identify the causes for the unexpected decrease in sampling efficiency with SA , we exclude the treatment AS with surfactants from our final sampling efficiency recommendation. Since sampling efficiency only decreases slightly when surfactants are added (Fig. Figure 6) and the mean of sampling efficiency in the treatment with surfactants falls within the range of the sampling efficiencies without surfactants, we conclude that the effect of SA on sampling efficiency is minor, especially in view of natural concentrations.

4.2 Is a meaningful analysis of glass plate samples possible?

Yang (1999) and Yang et al. (2001) used the glass plate to sample DMS from the SML in the field. The enrichment factors range from 1.21–3.08 and 0.41–1.18 respectively, highlighting that measurements are possible, and not all DMS is lost in the process of sampling. But whether those values and their variations are meaningful or mainly driven randomly has yet to be investigated. To accompany their field data interpretation, Yang et al. (2001) investigate, what drives differences in C_{SML} in the lab, but they do not derive a sampling efficiency. Our results show evidence—against the accepted notion that glass plate sampling is not meaningful for volatile compounds (e.g., Engel et al., 2017)—that the glass plate is a reliable sampling method for trace gases, yet associated with high losses (i.e., low sampling efficiency). Our results for FW and AS are more consistent than expected across the parameter space we investigated. Additionally, sampling efficiency does not vary significantly between the tested trace gases. However, the variance cannot be explained fully by the recorded parameters, therefore, any correction approach of C_{SML} should be qualitative and not quantitative, i.e., Eq. (7) cannot be used at this time. The consistency of the results might also be related to the limited parameter space (i.e., water temperature and trace gas concentration in the ULW) we achieved in this study. For example, Yang et al. (2001, their Table 2) showed in a lab setup that an enrichment of DMS correlates negatively with water temperature (0–28 °C), starting from an enrichment factor of 0.58 and reducing to 0.11. It is, therefore, likely that the effect water temperature would have on sampling efficiency is simply not visible in our study. There are more parameters, that will affect sampling efficiency, that have not been covered in this study. Wind influence on the glass plate will likely decrease sampling efficiency for the short-lived trace gases with increasing wind speed, until a point $U_{10,max}$, when all trace gas content is removed towards the near-zero atmosphere (personal communication with Theresa Barthelmeß). Furthermore, local wind variations in space and time on the scale of the glass plate size and on the order of the sampling duration will likely increase the variability of sampling efficiency (for same U_{10}). Extending the study to long-lived trace gases with a non-zero atmospheric background will change the sampling efficiency, as well. While for short-lived, highly reactive trace gases, as studied here, the sampling efficiency is relatively low, it is probably much higher for long-lived, less



reactive trace gases, like N_2O . A small, test data set (data not shown) of N_2O glass plate and ULW samples (12 SML–ULW pairs) that we collected using artificial seawater ($S_P = 30.2$, $T_{water} = 22.9 \pm 0.7$ °C) in a tank (equilibrated with lab air over more than 24 h), showed mean sampling efficiency of 0.97 ± 0.12 , indicating that the ΔC and the corresponding equilibration
 595 behaviour of the trace gas play a major role. Unfortunately, this effect was not covered in this study, as the three chosen trace gases were too similar to capture this difference.

4.3 What drives sampling efficiency?

Sampling efficiency in general is driven by turbulence effects, the equilibration between water on the glass plate and the surrounding air, potential adsorption effects (Walker et al., 2016) and photochemical processes (Zemmelink et al., 2005). Note,
 600 that chemical and biological reactions are excluded from this study and not considered to be factors for driving sampling efficiency here. The dilution of SML samples by ULW water due to varying sampling thicknesses, is also excluded from our concept of sampling efficiency, because $C_{SML} = C_{ULW}$. Turbulence effects and equilibration can both reduce and increase sampling efficiency, depending on the concentration gradient from the glass plate towards the surrounding air. Photochemical processes heavily depend on the trace gas, light intensity and duration of exposure. Sampling efficiency will decrease with
 605 degradation and increase with production processes, however, the increase would be artificial (i.e., not due to sampling effects). All experiments were performed indoors and UV light, therefore, was filtered out. Adsorption to the glass plate, the wiper and the funnel always reduces sampling efficiency by removing molecules from the sample, though, the effective amount is hard to quantify. Adsorption losses are deemed negligible in our setup, because of the regular cleaning of all used materials. Turbulence increases the effect of equilibration many times and we kept this at a minimum by omitting wind influence. The
 610 presence of substances other than trace gases, e.g., surfactants, can have an enhancing, reversing or neutral effect on turbulence and equilibration. In this study the concentration gradients were always going towards the near-zero atmosphere concentration, therefore, we hypothesize that the turbulence introduced by the sample handling during glass plate sampling decreases sampling efficiency. Additionally, the sampling efficiency is expected to further decrease, due to fast paced equilibration along the high concentration gradient (from oversaturated water to near-zero atmosphere) and the short-lived, highly reactive species
 615 with low solubility that we use.

The MLR (explains about 60 % variance) identifies spike volume per litre (positive) and water temperature (negative) as significant predictors for DMS-d3 and $^{13}CS_2$ sampling efficiency in FW and AS. For isoprene-d5 it is spike volume per litre (positive) and the number of dips (positive). This changes with the addition of surfactants. The MLR (explains less than 25 % of variance). Spike volume per litre (positive) and water temperature (negative) are significant predictors for DMS-d3 sampling
 620 efficiency in FW and AS without and with surfactants. SA_{SML} is almost a significant predictor (negative). The only significant predictor for isoprene-d5 sampling efficiency is water temperature (negative). Water temperature (negative) and SA_{SML} are significant predictors for $^{13}CS_2$ sampling efficiency. It is reasonable that the spike volume per litre is correlated with sampling efficiency, as it is proportional to the concentration of the trace gases in the tank, which in turn is proportional to the ΔC (as $C_{air} = \text{constant} \approx 0$). The identified, negative relationship of sampling efficiency and water temperature matches the



625 expectations and the findings by. Most likely this effect relates to the negative relationship of solubility and water temperature, highlighting that the solubility is a potential candidate to be used to predict sampling efficiency. Furthermore, water temperature increases the speed of molecular diffusion, which in turn would also further reduce sampling efficiency. Salinity is not significant for any of the trace gases. Although, salinity has an effect on solubility (Weisenberger and Schumpe, 1996), it is much lower than the effect of temperature. There seems to be an effect from the kind of water used though, as the two AS
 630 media used (coral reef seawater and ultrapure water with ordinary aquarium salt) differ significantly, e.g., sampling efficiency for DMS-d3 in AS is almost two times higher in experiment C than in experiment B. Within this study, it cannot be determined whether this can be attributed to the salt composition, other substances mainly present in the coral reef seawater or any of the other properties of the coral reef seawater, like the alkalinity. Surprisingly the number of dips was only significant in one case, with a predicted positive relationship, which would imply a larger sampling efficiency with more dips. This is counter to the
 635 hypothesis, that the sampling duration (which is a multiple of the number of dips) decreases sampling efficiency, and this result is therefore deemed a mathematical construct. Although each dip has its individual sampling efficiency, it is to be expected that the number of dips affects the overall variability of our sampling efficiency estimate, but not the mean sampling efficiency. However, the number of dips is closely linked to sampling duration. The sampling duration affects how long the vial is left open, potentially resulting in a change of the mean sampling efficiency. However, our lab experience has shown that leaving
 640 the vial open has a negligible effect and non-significant results for the number of dips from the MLR supports this. SA_{SML} was included in MLR as the factor driving sampling efficiency, though SA_{ULW} seemed to have an effect on PA as well. We expected a positive relationship with sampling efficiency, since field studies show a decrease in gas exchange for increased SA_{SML} , i.e., more retained in SML. Instead, the model and linear regressions (Appendix E) show negative correlations. Most recent gas exchange reduction publications look into wind-driven effects, which is why those are not directly applicable here, but Schmidt
 645 and Schneider (2011) show a decrease in gas exchange for stirring samples with surfactants only, highlighting that a positive correlation is still more likely. The observed negative correlation might be related to the differing sampling thickness (less dips required with surfactants, i.e., larger thickness sampled), though with the mixing we should have achieved a homogeneous concentration throughout the tank (Fig. Figure 3). Even if there had been a negative gradient towards the SML (e.g., DBL, due to outgassing to the near-zero atmosphere), the increased sampling thickness would have put proportionally more ULW into
 650 the sample, i.e., increasing C_{SML} and causing an (artificial) increase in sampling efficiency (i.e., positive correlation). However, Adenaya et al. (2021) show that the thickness of the DBL increases with the addition of artificial surfactants (as opposed to natural surfactants), which would counteract the effect of the sampling thickness. The final assessment, though, is difficult, as the increase of the DBL cannot be quantified in our setup. Though SA_{SML} was much higher than natural values, it was still very far from the critical micelle concentration of TX-100, which is why the negative correlation can also not be explained by
 655 micelle formation. It is possible that the surfactants affect the homogeneous mixing of the trace gases. If surfactants decrease the trace gas concentration with increasing SA , this would explain the negative correlation and the decrease in mean sampling efficiency. The trace gases would mix less into the SML than in the ULW (as $SA_{SML} > SA_{ULW}$), decreasing sampling efficiency.

Finally, the data is relatively scattered and the linear correlation is weak, which might be an indication that we are not capturing a single, physical relationship, but a net effect of entangled factors.

660 **Table 7 Sampling efficiency summarized from treatment FW and AS. MLR is multiple linear regression, T_{water} is the water temperature and C_{ULW} the concentration of trace gas in the ULW. ⁽¹⁾Values copied from solubility compilation by Sander (2023). ⁽²⁾Value calculated for $T = 25\text{ °C}$ and $S_P = 0$ from Saltzman et al. (1993). Diffusivity in (sea)water is not published (n.p.) for isoprene or CS_2 . ⁽³⁾Due to their analogous structures, CO_2 diffusivity can be used as an approximation for CS_2 , i.e., $D_{CO_2} = 1.88 \times 10^{-9}\text{ m}^2\text{ s}^{-1}$ at $T = 25\text{ °C}$ and $S_P = 0$ (Mazarei and Sandall, 1980). Global estimates of atmospheric mixing ratios are taken from ⁽⁴⁾Liss et al. (2014), 665 ⁽⁵⁾Ferracci et al. (2024), Southern Ocean only, and ⁽⁶⁾Lennartz et al. (2020).**

Trace gas	Sampling efficiency						Solubility ⁽¹⁾	Diffusivity	Atmospheric
	Data set			MLR		Drivers	$\times 10^{-3}$	$\times 10^{-9}$	mixing ratios
	n	mean $\pm SE$	min–max	n	intercept $\pm SE$		[mol m ⁻³ Pa ⁻¹]	[m ² s ⁻¹]	[ppt]
DMS	30	0.12 $\pm$ 0.01	0.07–0.21	19	0.13 $\pm$ 0.01	T_{water}, C_{ULW}	5.3	1.35 ⁽²⁾	~100 ⁽⁴⁾
isoprene	29	0.12 $\pm$ 0.01	0.05–0.21	19	0.13 $\pm$ 0.01	C_{ULW} , dips	0.13	n.p.	42.4 ⁽⁵⁾
CS_2	30	0.12 $\pm$ 0.01	0.05–0.20	19	0.12 $\pm$ 0.01	T_{water}, C_{ULW}	0.61	n.p. ⁽³⁾	~50 ⁽⁶⁾

5 Conclusion and outlook

In this study we set up experiments to determine sampling efficiency of the glass plate technique for the short-lived trace gases DMS, isoprene and CS_2 in the SML. To understand the drivers of sampling efficiency we varied water temperature, salinity and trace gas concentration. To mimic in-situ conditions, we added three different levels of surfactants in one treatment. 670 Sampling efficiency ($\pm SE$) is low at about 0.12 ± 0.01 , but consistent for all three trace gases, while the drivers only have a low impact in view of the overall uncertainty of the sampling efficiency, though water temperature and trace gas concentration were significant for almost all cases. Note, that surfactants in the SML exceeded natural values by far, and had an unexpected negative correlation with sampling efficiency.

Given the consistency we conclude that meaningful analysis of SML trace gas samples in the field is possible. However, we 675 do not recommend to apply the sampling efficiency quantitatively to correct concentrations of SML samples, given the uncertainty present in our sampling efficiency estimates added to by the uncertainty of factors in the field that have not been attributed in this study (e.g., wind). Instead we propose to use the intercept from the MLR fit with the SE of the sampling efficiency as a limit to categorize ratios of measured SML over ULW concentrations for enrichment (Table 7) identifying those above the sampling efficiency as cases of net production. Our results will be applied to samples of DMS, isoprene and CS_2 680 from a mesocosm study (Bibi et al., 2025) and from two cruises, one in the North Sea (2024) and one in the Baltic Sea (2025), to further investigate the potential to explain SML trace gas processes and link them to air-sea exchange.

Our findings are a first step towards utilizing glass plate sampling for trace gases, yet, they also highlight the need for further investigation. We studied highly reactive, short-lived trace gases here, for which the concentration gradient usually is large. A preliminary test that we performed at presumably equilibrium conditions for the relatively inert, long-lived trace gas N_2O have



685 shown a contrasting near 100 % sampling efficiency, highlighting the role of the concentration gradient and divergence from
equilibrium on sampling efficiency. We, therefore, strongly recommend to extend our results to less reactive and long-lived
trace gases, like the prominent greenhouse gases N_2O and CH_4 to complement our findings. As in field studies of gas exchange
wind is an important driver, there is also the need to investigate how sampling efficiency is affected by wind in combination
with the trace gas concentration gradient between sample water and air, and whether a threshold of wind speed exists above
690 which glass plate samples do not contain the trace gas anymore. The mesh screen is also widely used. Repeating the study with
the mesh screen, could increase the availability of trace gas measurements of the SML. Finally, the relationship between
sampling efficiency on the one hand, and water temperature and surfactants on the other hand, needs more in-depth
investigation. By unravelling the underlying physical relationships, it might become possible to not only categorize enrichment
of trace gases in the SML, but to correct the individual measurements of trace gases in the SML.

695 **Data availability**

All data will be archived and made available to the scientific community at the PANGAEA database upon doi assignment. In
the meantime, data are available from the authors upon request.

Author contribution

Lea Lange: Conceptualization, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing
700 (original draft preparation), Writing (review and editing)
Dennis Booge: Conceptualization, Supervision, Writing (review and editing)
Hendrik Feil: Investigation, Writing (review and editing)
Josefine Karnatz: Investigation, Writing (review and editing)
Ina Stoltenberg: Supervision, Writing (review and editing)
705 Hermann W. Bange: Conceptualization, Funding acquisition, Resources, Supervision, Writing (review and editing)
Christa A. Marandino: Conceptualization, Funding acquisition, Resources, Supervision, Writing (review and editing)

Competing interests

HWB is a member of the editorial board of Biogeosciences.

Acknowledgements

710 We would like to thank Edgar Cortés, Alisa Wüst, Carola Lehnert, Mariana Ribas-Ribas and Oliver Wurl for providing us
with support, facilities and consumables during experiment B at ICBM, Carl von Ossietzky Universität Oldenburg. We would



like to thank our student assistant Fenja Möller for supporting us with sampling during experiment C at GEOMAR Helmholtz Centre for Ocean Research Kiel. We would also like to thank our student assistant Samira Linder for helping with surfactants measurements at GEOMAR Helmholtz Centre for Ocean Research Kiel. This is a contribution to the international Surface Ocean-Lower Atmosphere Study (SOLAS), which receives funding and support from the Scientific Committee on Oceanic Research (e.g., NSF Grant OCE-1840868 and 2140395), the iCACGP, the WCRP and Future Earth, and the State Key Laboratory of Marine Environmental Science, China.

Financial support

This research was supported by the project "Biogeochemical processes and Air-sea exchange in the Sea-Surface microlayer (BASS)", which was funded by the German Research Foundation (DFG) under Grant No 451574234.

References

- Adenaya, A., Haack, M., Stolle, C., Wurl, O., and Ribas Ribas, M.: Effects of Natural and Artificial Surfactants on Diffusive Boundary Dynamics and Oxygen Exchanges across the Air-Water Interface, *Oceans*, 2, <https://doi.org/10.3390/oceans2040043>, 2021.
- The Python Language Reference: <https://docs.python.org/3.11/reference/index.html>, last access: 6 October 2025.
- Barthelmeß, T. and Engel, A.: How biogenic polymers control surfactant dynamics in the surface microlayer: insights from a coastal Baltic Sea study, *Biogeosciences*, 19, 4965–4992, <https://doi.org/10.5194/bg-19-4965-2022>, 2022.
- Bibi, R., Ribas-Ribas, M., Jaeger, L., Lehnert, C., Gassen, L., Cortés, E., Wollschläger, J., Thölen, C., Waska, H., Zöbelein, J., Brinkhoff, T., Athale, I., Röttgers, R., Novak, M., Engel, A., Barthelmeß, T., Karnatz, J., Reinthaler, T., Spriahailo, D., Friedrichs, G., Schäfer, F., and Wurl, O.: Biogeochemical Dynamics of the Sea-Surface Microlayer in a Multidisciplinary Mesocosm Study, *EGUsphere*, 1–53, <https://doi.org/10.5194/egusphere-2025-1773>, 2025.
- Booge, D., Marandino, C. A., Schlundt, C., Palmer, P. I., Schlundt, M., Atlas, E. L., Bracher, A., Saltzman, E. S., and Wallace, D. W. R.: Can simple models predict large-scale surface ocean isoprene concentrations?, *Atmospheric Chemistry and Physics*, 16, 11807–11821, <https://doi.org/10.5194/acp-16-11807-2016>, 2016.
- Ćosović, B. and Vojvodić, V.: The application of ac polarography to the determination of surface-active substances in seawater, *Limnology & Oceanography*, 27, 361–369, <https://doi.org/10.4319/lo.1982.27.2.0361>, 1982.
- Ćosović, B. and Vojvodić, V.: Voltammetric Analysis of Surface Active Substances in Natural Seawater, *Electroanalysis*, 10, 429–434, [https://doi.org/10.1002/\(SICI\)1521-4109\(199805\)10:6<429::AID-ELAN429>3.0.CO;2-7](https://doi.org/10.1002/(SICI)1521-4109(199805)10:6<429::AID-ELAN429>3.0.CO;2-7), 1998.
- Crow, S. A., Ahearn, D. G., Cook, W. L., and Bourquin, A. W.: Densities of bacteria and fungi in coastal surface films as determined by a membrane-adsorption procedure, *Limnology and Oceanography*, 20, 644–646, <https://doi.org/10.4319/lo.1975.20.4.0644>, 1975.



- Cunliffe, M. and Wurl, O. (Eds.): Guide to best practices to study the ocean's surface, Plymouth, UK, 118 pp., 2014.
- Engel, A., Bange, H. W., Cunliffe, M., Burrows, S. M., Friedrichs, G., Galgani, L., Herrmann, H., Hertkorn, N., Johnson, M., Liss, P. S., Quinn, P. K., Schartau, M., Soloviev, A., Stolle, C., Upstill-Goddard, R. C., van Pinxteren, M., and Zäncker, B.: The Ocean's Vital Skin: Toward an Integrated Understanding of the Sea Surface Microlayer, *Front. Mar. Sci.*, 4, <https://doi.org/10.3389/fmars.2017.00165>, 2017.
- Ferracci, V., Weber, J., Bolas, C. G., Robinson, A. D., Tummon, F., Rodríguez-Ros, P., Cortés-Greus, P., Baccarini, A., Jones, R. L., Galí, M., Simó, R., Schmale, J., and Harris, N. R. P.: Atmospheric isoprene measurements reveal larger-than-expected Southern Ocean emissions, *Nat Commun*, 15, 2571, <https://doi.org/10.1038/s41467-024-46744-4>, 2024.
- Garbe, C. S., Rutgersson, A., Boutin, J., de Leeuw, G., Delille, B., Fairall, C. W., Gruber, N., Hare, J., Ho, D. T., Johnson, M. T., Nightingale, P. D., Pettersson, H., Piskozub, J., Sahlée, E., Tsai, W., Ward, B., Woolf, D. K., and Zappa, C. J.: Transfer Across the Air-Sea Interface, in: *Ocean-Atmosphere Interactions of Gases and Particles*, edited by: Liss, P. S. and Johnson, M. T., Springer, Berlin, Heidelberg, 55–112, https://doi.org/10.1007/978-3-642-25643-1_2, 2014.
- Garrett, W. D.: Collection of Slick-Forming Materials from the Sea Surface, *Limnology and Oceanography*, 10, 602–605, <https://doi.org/10.4319/lo.1965.10.4.0602>, 1965.
- Garrett, W. D.: Comments on “Laboratory comparisons of four surface microlayer samplers” (R. F. Hatcher and B. C. Parker), *Limnology and Oceanography*, 19, 166–167, <https://doi.org/10.4319/lo.1974.19.1.0166>, 1974.
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, *Nature*, 585, 357–362, <https://doi.org/10.1038/s41586-020-2649-2>, 2020.
- Harvey, G. W.: Microlayer Collection from the Sea Surface: A New Method and Initial Results, *Limnology and Oceanography*, 11, 608–613, <https://doi.org/10.4319/lo.1966.11.4.0608>, 1966.
- Harvey, G. W. and Burzell, L. A.: A Simple Microlayer Method for Small Samples, *Limnology and Oceanography*, 17, 156–157, <https://doi.org/10.4319/lo.1972.17.1.0156>, 1972.
- Hunter, J. D.: Matplotlib: A 2D Graphics Environment, *Computing in Science & Engineering*, 9, 90–95, <https://doi.org/10.1109/MCSE.2007.55>, 2007.
- Kock, A., Schafstall, J., Dengler, M., Brandt, P., and Bange, H. W.: Sea-to-air and diapycnal nitrous oxide fluxes in the eastern tropical North Atlantic Ocean, *Biogeosciences*, 9, 957–964, <https://doi.org/10.5194/bg-9-957-2012>, 2012.
- Laß, K., Bange, H. W., and Friedrichs, G.: Seasonal signatures in SFG vibrational spectra of the sea surface nanolayer at Boknis Eck Time Series Station (SW Baltic Sea), *Biogeosciences*, 10, 5325–5334, <https://doi.org/10.5194/bg-10-5325-2013>, 2013.
- Lennartz, S. T., Marandino, C. A., Booge, D., Pozzer, A., Steinhoff, T., Arevalo-Martinez, D. L., Kloss, C., Bracher, A., Röttgers, R., Atlas, E., and Krüger, K.: Direct oceanic emissions unlikely to account for the missing source of atmospheric carbonyl sulfide, *Atmos. Chem. Phys.*, 2017.



- Lennartz, S. T., Marandino, C. A., von Hobe, M., Andreae, M. O., Aranami, K., Atlas, E., Berkelhammer, M., Bingemer, H., Booge, D., Cutter, G., Cortes, P., Kremser, S., Law, C. S., Marriner, A., Simó, R., Quack, B., Uher, G., Xie, H., and Xu, X.: Marine carbonyl sulfide (OCS) and carbon disulfide (CS₂): a compilation of measurements in seawater and the marine boundary layer, *Earth System Science Data*, 12, 591–609, <https://doi.org/10.5194/essd-12-591-2020>, 2020.
- 780 Liss, P. S., Marandino, C. A., Dahl, E. E., Helmig, D., Hintsa, E. J., Hughes, C., Johnson, M. T., Moore, R. M., Plane, J. M. C., Quack, B., Singh, H. B., Stefels, J., von Glasow, R., and Williams, J.: Short-Lived Trace Gases in the Surface Ocean and the Atmosphere, in: *Ocean-Atmosphere Interactions of Gases and Particles*, edited by: Liss, P. S. and Johnson, M. T., Springer, Berlin, Heidelberg, 1–54, https://doi.org/10.1007/978-3-642-25643-1_1, 2014.
- Marandino, C. A., De Bruyn, W. J., Miller, S. D., Prather, M. J., and Saltzman, E. S.: Oceanic uptake and the global
 785 atmospheric acetone budget, *Geophysical Research Letters*, 32, <https://doi.org/10.1029/2005GL023285>, 2005.
- Marandino, C. A., De Bruyn, W. J., Miller, S. D., and Saltzman, E. S.: DMS air/sea flux and gas transfer coefficients from the North Atlantic summertime coccolithophore bloom, *Geophysical Research Letters*, 35, <https://doi.org/10.1029/2008GL036370>, 2008.
- Mazarei, A. F. and Sandall, O. C.: Diffusion coefficients for helium, hydrogen, and carbon dioxide in water at 25°C, *AIChE*
 790 *Journal*, 26, 154–157, <https://doi.org/10.1002/aic.690260128>, 1980.
- Mukerjee, P. and Mysels, K. J.: *Critical micelle concentrations of aqueous surfactant systems*, Washington, 36277 pp., 1971.
- Sabbaghzadeh, B., Upstill-Goddard, R. C., Beale, R., Pereira, R., and Nightingale, P. D.: The Atlantic Ocean surface microlayer from 50°N to 50°S is ubiquitously enriched in surfactants at wind speeds up to 13 m s⁻¹, *Geophysical Research*
 795 *Letters*, 44, 2852–2858, <https://doi.org/10.1002/2017GL072988>, 2017.
- Saint-Macary, A. D., Marriner, A., Barthelmeß, T., Deppeler, S., Safi, K., Costa Santana, R., Harvey, M., and Law, C. S.: Dimethyl sulfide cycling in the sea surface microlayer in the southwestern Pacific – Part 1: Enrichment potential determined using a novel sampler, *Ocean Sci.*, 19, 1–15, <https://doi.org/10.5194/os-19-1-2023>, 2023.
- Saltzman, E. S., King, D. B., Holmen, K., and Leck, C.: Experimental determination of the diffusion coefficient of
 800 dimethylsulfide in water, *J. Geophys. Res.*, 98, 16481, <https://doi.org/10.1029/93JC01858>, 1993.
- Sander, R.: Compilation of Henry’s law constants (version 5.0.0) for water as solvent, *Atmos. Chem. Phys.*, 23, 10901–12440, <https://doi.org/10.5194/acp-23-10901-2023>, 2023.
- Schmidt, R. and Schneider, B.: The effect of surface films on the air–sea gas exchange in the Baltic Sea, *Marine Chemistry*, 126, 56–62, <https://doi.org/10.1016/j.marchem.2011.03.007>, 2011.
- 805 Scholz, F.: *Voltammetric techniques of analysis: the essentials*, *ChemTexts*, 1, 17, <https://doi.org/10.1007/s40828-015-0016-y>, 2015.
- Seabold, S. and Perktold, J.: *Statsmodels: Econometric and Statistical Modeling with Python*, SciPy 2010, <https://doi.org/10.25080/Majora-92bf1922-011>, 2010.
- The Matplotlib Development Team: *Matplotlib: Visualization with python*, Zenodo,



- 810 <https://doi.org/10.5281/zenodo.14464227>, 2024.
- Turner, S. M. and Liss, P. S.: Measurements of various sulphur gases in a coastal marine environment, *J Atmos Chem*, 2, 223–232, <https://doi.org/10.1007/BF00051074>, 1985.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E.,
 815 Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., and van Mulbregt, P.: SciPy 1.0: fundamental algorithms for scientific computing in Python, *Nat Methods*, 17, 261–272, <https://doi.org/10.1038/s41592-019-0686-2>, 2020.
- Walker, C. F., Harvey, M. J., Smith, M. J., Bell, T. G., Saltzman, E. S., Marriner, A. S., McGregor, J. A., and Law, C. S.:
 820 Assessing the potential for dimethylsulfide enrichment at the sea surface and its influence on air–sea flux, *Ocean Science*, 12, 1033–1048, <https://doi.org/10.5194/os-12-1033-2016>, 2016.
- Waskom, M. L.: seaborn: statistical data visualization, *Journal of Open Source Software*, 6, 3021, <https://doi.org/10.21105/joss.03021>, 2021.
- Weisenberger, S. and Schumpe, A.: Estimation of gas solubilities in salt solutions at temperatures from 273 K to 363 K,
 825 *AIChE Journal*, 42, 298–300, <https://doi.org/10.1002/aic.690420130>, 1996.
- Wurl, O., Wurl, E., Miller, L., Johnson, K., and Vagle, S.: Formation and global distribution of sea-surface microlayers, *Biogeosciences*, 8, 121–135, <https://doi.org/10.5194/bg-8-121-2011>, 2011.
- Yang, G.-P.: Dimethylsulfide enrichment in the surface microlayer of the South China Sea, *Marine Chemistry*, 66, 215–224, [https://doi.org/10.1016/S0304-4203\(99\)00042-0](https://doi.org/10.1016/S0304-4203(99)00042-0), 1999.
- 830 Yang, G.-P., Watanabe, S., and Tsunogai, S.: Distribution and cycling of dimethylsulfide in surface microlayer and subsurface seawater, *Marine Chemistry*, 76, 137–153, [https://doi.org/10.1016/S0304-4203\(01\)00054-8](https://doi.org/10.1016/S0304-4203(01)00054-8), 2001.
- Zavarsky, A., Booge, D., Fiehn, A., Krüger, K., Atlas, E., and Marandino, C.: The Influence of Air-Sea Fluxes on Atmospheric Aerosols During the Summer Monsoon Over the Tropical Indian Ocean, *Geophysical Research Letters*, 45, 418–426, <https://doi.org/10.1002/2017GL076410>, 2018.
- 835 Zemmellink, H., Houghton, L., Sievert, S., Frew, N., and Dacey, J.: Gradients in dimethylsulfide, dimethylsulfoniopropionate, dimethylsulfoxide, and bacteria near the sea surface, *Mar. Ecol. Prog. Ser.*, 295, 33–42, <https://doi.org/10.3354/meps295033>, 2005.

Appendix

840 A Preliminary test

Experiment A was conducted on three consecutive days between 13–15 March 2023, using one of our incubation baths as tank, filled with ultrapure water. We did these preliminary tests to learn about handling, measurements, mixing in the tank and sampling efficiency calculation from PA . As no glass plate had been available, we used a mesh screen. Targeted sample volume was 60 mL. Figure A1 depicts the ratios of PA/PA_{ref} in experiment A, analogous to Fig. Figure 3 (experiment B and C). Figure
845 A2 depicts the resulting sampling efficiency.

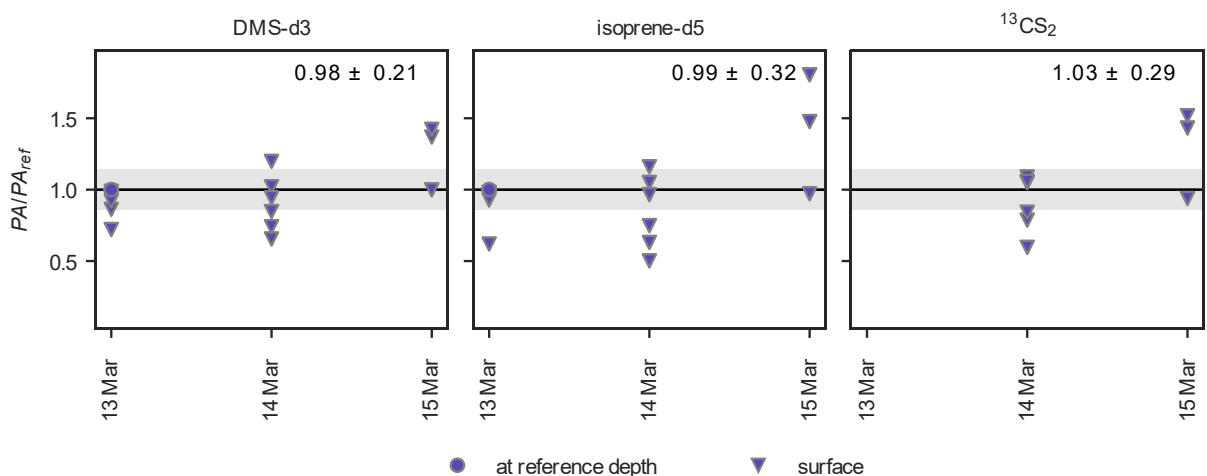


Figure A1 Test for uniform trace gas concentration in the tank (well-mixed) in experiment A with mesh screen. The ratios PA/PA_{ref} ($n = 16$ for DMS-d3 and isoprene-d5, $n = 9$ for $^{13}\text{CS}_2$) are shown against date for DMS-d3 (left), isoprene-d5 (middle) and $^{13}\text{CS}_2$ (right). Black horizontal line depicts $PA = PA_{ref}$ (i.e., well-mixed). Grey shaded area indicates $\pm 14.1\%$ uncertainty. Mean and standard deviation for PA/PA_{ref} is indicated in upper right of each subplot (including all shown points).
850

9 of 16 points lie within the uncertainty range for DMS-d3 and isoprene-d5, and 4 of 9 for $^{13}\text{CS}_2$. Therefore, it seems the tank was overall less well-mixed than in experiment B and C. However, no points are excluded from calculation of sampling efficiency, as experiment A serves as a test. This results in average ratios ($\pm$ standard deviation) of PA/PA_{ref} of 0.98 ± 0.21 , 0.99 ± 0.32 , and 1.03 ± 0.29 for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ respectively. Sampling efficiency is 0.30 ± 0.08 ($n = 13$),
855 0.27 ± 0.09 ($n = 13$), and 0.26 ± 0.09 ($n = 10$) for DMS-d3, isoprene-d5, and $^{13}\text{CS}_2$ respectively. Mean sampling efficiency and its standard deviation are, thus, larger here than with the glass plate in experiment B and C (compare with FW in Table 4). The preliminary test with the mesh screen was less well constrained, which might explain the larger variation. The mesh screen might be more efficient, since wiping is not required. Also, the mesh screen captures more volume per dip, reducing the overall sampling duration. Furthermore, the mechanisms that contain the sample (fine mesh vs glass surface) might inherently be
860 different physically and interact with the samples' properties. In field studies, the mesh screen samples are often associated with more dilution from ULW, as it samples a larger thickness than the glass plate. We do not expect this to have caused the differences in sampling efficiency during our experiment, as the ULW and SML should have been well-mixed. Finally, we did



not remove any data points, thereby assuming that the tank was well-mixed at all times, though Fig. A1 indicates otherwise. An extended QC would presumably change the results and likely decrease the variation, but potentially maintain a similar mean sampling efficiency. We conclude from this preliminary test, that our experimental design seems suited to repeat the sampling efficiency estimation for mesh screen sampling.

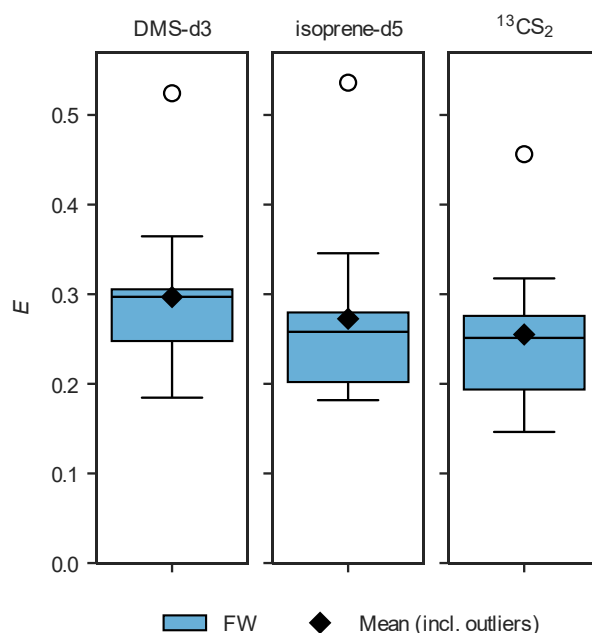
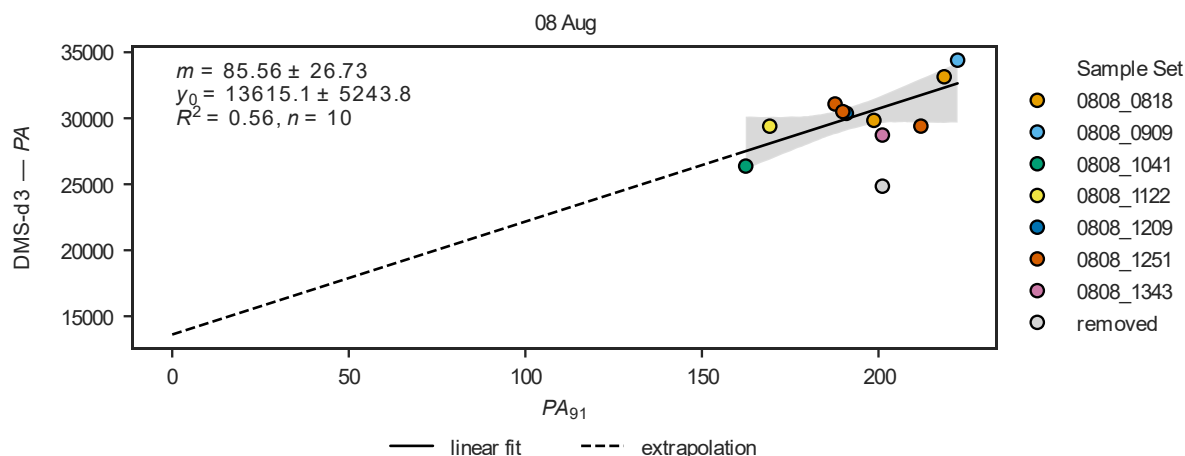


Figure A2 Sampling efficiency for DMS-d3 (left), isoprene-d5 (middle) and ¹³CS₂ (right) in experiment A with the mesh screen in treatment fresh water (FW). Box width shows 25 to 75 percentile, centre line denotes the median, diamonds denote the mean, open circles denote outliers, and whiskers show minimum and maximum. The number of data points in experiment A is $n = 13$ ($n = 10$ for ¹³CS₂).

B Additional quality control using mass 91 applied in the context of Fig. Figure 3



875 **Figure B1 Example from 8 August of linear fit between PA of DMS-d3 and mass 91 to identify poor-quality measurements. All samples are ULW samples. Linear fit excludes grey point.**

Prior to assessing whether the tank was well-mixed (Sect. 3.2 Mixing in the tank), an additional step of quality control was performed on all ULW samples on each of the isotopes. PA of ULW samples that were far from the linear fit between PA of the respective isotope and mass 91 (e.g., grey point in Fig. B1) were flagged as poor-quality measurements and consequently removed from analysis. We assume that the poor-quality is caused by errors in measurement handling or short-term sensitivity drifts in the instrument. If the removed ULW sample was a reference sample, it was replaced with another ULW sample from the same sample set or, in the absence of more ULW samples, the sample set was discarded completely. SML samples formed a different cluster, due to the systematically lower PA compared to ULW (caused by decreased sampling efficiency) and, therefore, this step was not applied to SML samples. The sample indicated as being removed in Fig. B1 (mass spectrum ID "BULK_13CM_CC_06") was the only ULW sample that was removed from the data set this way. It will be indicated with use_for_analysis_flag=False in the published data set.



C Effect of purging on surface activity measurements

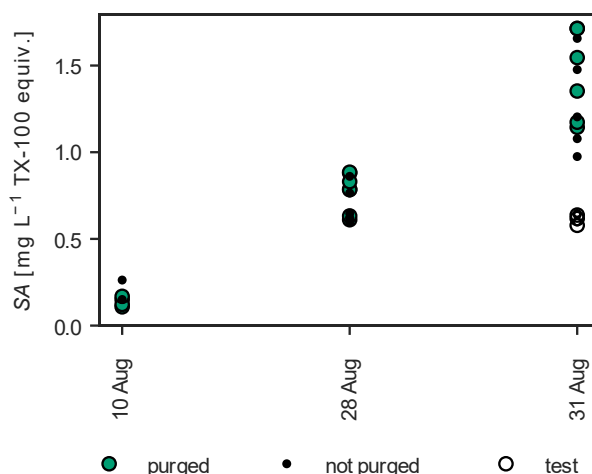


Figure C1 Measured surface activity (SA) of purged and unpurged samples. ULW samples in treatment AS with surfactants ($n = 30$, green and black) are shown and a test sample set ($n = 6$, empty circles).

Due to the limited surface area of the tank used in experiment C ("AZ crate") two SML samples taken after each would not have been appropriate replicates of each other. Therefore, it was tested if the GC–MS sample could be reused for voltammetry analysis. The main differences for a reused GC–MS sample in voltammetry analysis are (1) longer standing time of sample before freezing and (2) being purged with helium for 10 min. These differences could lead to contamination due to the added handling. Figure C1 shows the data points used in the statistics of purged vs unpurged SA (Sect. 3.1.2 Surface activity and enrichment of surfactants). Only ULW samples were used, as we could safely assume that purging is the only source of error. Additionally, we took six subsamples (three for control, three were purged) from a beaker filled with ULW from the tank to specifically test the effect of purging. Surface activity is not affected by purging.



D Results linear and multiple linear regressions in fresh water and artificial seawater from Sect. 3.4

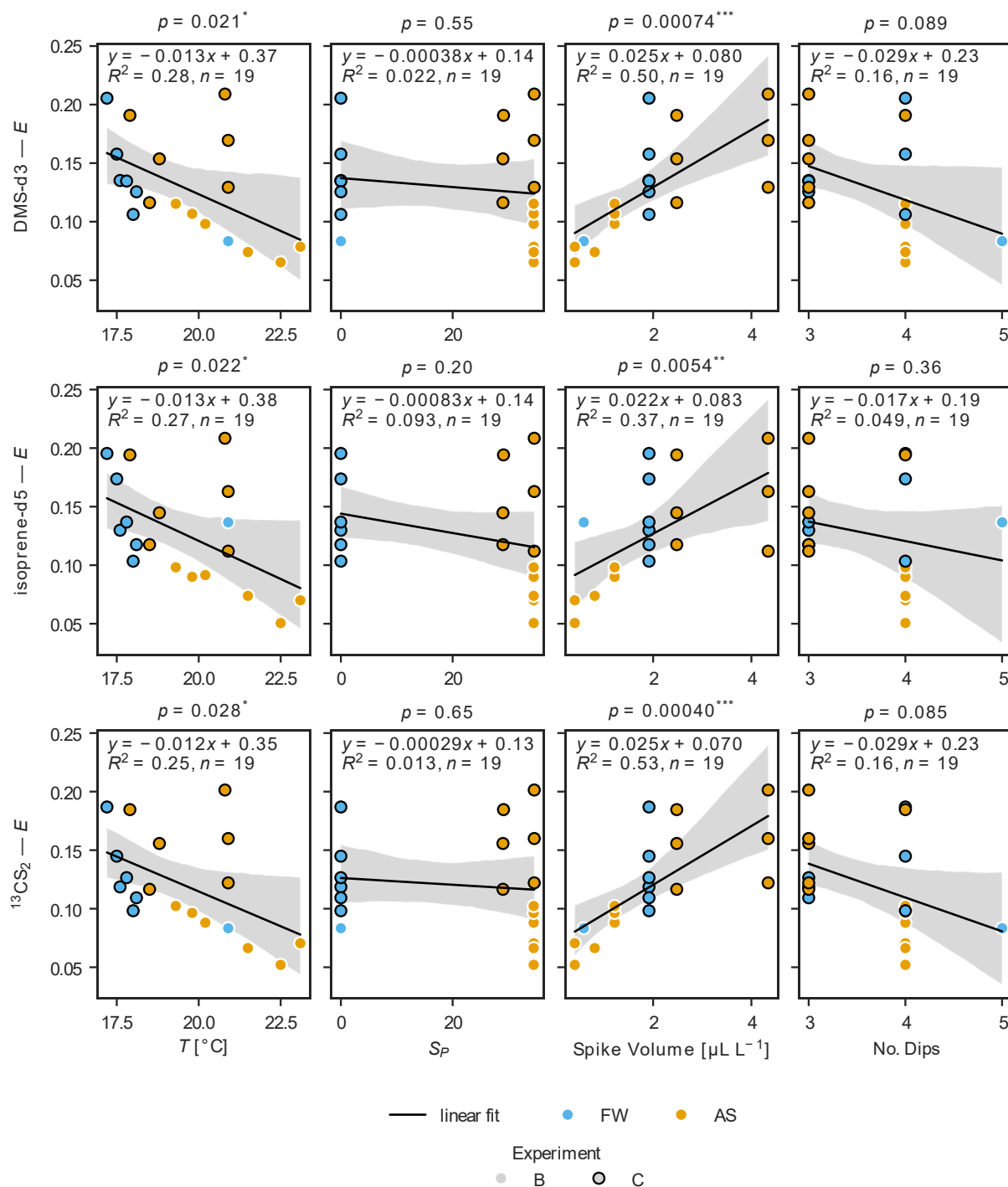


Figure D1 Linear regressions of sampling efficiency for DMS-d3 (top), isoprene-d5 (middle) and $^{13}\text{CS}_2$ (bottom) against water temperature, salinity, spike volume per litre and number of dips in fresh water (FW) and artificial seawater (AS) treatment. Linear fit, R^2 and number of samples (n) are given in top corner of each subplot.

Table D1 Multiple linear regression results for DMS-d3 in fresh water (FW) and artificial seawater (AS), using statsmodels smf.ols(...).fit(cov_type='HC3')

OLS Regression Results						
=====						
Dep. Variable:	ratio_sml_over_ulw_dmsd3	R-squared:	0.709			
Model:	OLS	Adj. R-squared:	0.626			
910 Method:	Least Squares	F-statistic:	5.639			
Date:	Sat, 01 Nov 2025	Prob (F-statistic):	0.00643			
Time:	15:37:10	Log-Likelihood:	45.086			
No. Observations:	19	AIC:	-80.17			
Df Residuals:	14	BIC:	-75.45			
915 Df Model:	4					
Covariance Type:	HC3					
=====						
	coef	std err	z	P> z	[0.025	0.975]

920 Intercept	0.1292	0.007	18.107	0.000	0.115	0.143
temperature_water_degC	-0.0223	0.011	-2.021	0.043	-0.044	-0.001
salinity_psu	0.0041	0.008	0.513	0.608	-0.012	0.020
cummulative_spike_volume_mikrol_per_l	0.0342	0.012	2.769	0.006	0.010	0.058
dips	0.0117	0.009	1.350	0.177	-0.005	0.029
925	=====					
Omnibus:	0.354	Durbin-Watson:	1.951			
Prob(Omnibus):	0.838	Jarque-Bera (JB):	0.074			
Skew:	0.145	Prob(JB):	0.964			
Kurtosis:	2.903	Cond. No.	2.58			
930	=====					

Notes:

[1] Standard Errors are heteroscedasticity robust (HC3)

Table D2 Multiple linear regression results for isoprene-d5 in fresh water (FW) and artificial seawater (AS), using statsmodels smf.ols(...).fit(cov_type='HC3')

OLS Regression Results			
=====			
Dep. Variable:	ratio_sml_over_ulw_isoprened5	R-squared:	0.691
Model:	OLS	Adj. R-squared:	0.602
Method:	Least Squares	F-statistic:	6.803



```

Date:                      Sat, 01 Nov 2025   Prob (F-statistic):      0.00294
Time:                      15:37:10          Log-Likelihood:          43.711
No. Observations:         19                AIC:                    -77.42
945 Df Residuals:          14                BIC:                    -72.70
Df Model:                  4
Covariance Type:          HC3

=====
                                coef      std err          z      P>|z|      [0.025      0.975]
950 -----
Intercept                  0.1267      0.008       16.425     0.000      0.112      0.142
temperature_water_degc    -0.0169      0.012       -1.455     0.146     -0.040      0.006
salinity_psu              -0.0069      0.009       -0.741     0.459     -0.025      0.011
cumulative_spike_volume_mikrol_per_l  0.0404      0.014        2.812     0.005      0.012      0.069
955 dips                   0.0213      0.009        2.367     0.018      0.004      0.039

=====
Omnibus:                   3.169   Durbin-Watson:           1.648
Prob(Omnibus):             0.205   Jarque-Bera (JB):         1.515
Skew:                      -0.648   Prob(JB):                 0.469
960 Kurtosis:              3.486   Cond. No.                 2.58

=====
    
```

Notes:

[1] Standard Errors are heteroscedasticity robust (HC3)

965 **Table D3 Multiple linear regression results for $^{13}\text{CS}_2$ in fresh water (FW) and artificial seawater (AS), using statsmodels smf.ols(...).fit(cov_type='HC3')**

```

                                OLS Regression Results
=====
970 Dep. Variable:      ratio_sml_over_ulw_c13s2   R-squared:              0.729
Model:                  OLS                       Adj. R-squared:         0.652
Method:                 Least Squares             F-statistic:           6.742
Date:                   Sat, 01 Nov 2025          Prob (F-statistic):    0.00305
Time:                   15:37:10                  Log-Likelihood:        45.959
975 No. Observations:   19                       AIC:                   -81.92
Df Residuals:           14                       BIC:                   -77.20
Df Model:                4
Covariance Type:        HC3

=====
                                coef      std err          z      P>|z|      [0.025      0.975]
980 -----
Intercept               0.1202      0.007       17.774     0.000      0.107      0.133
    
```



	temperature_water_degc	-0.0219	0.010	-2.104	0.035	-0.042	-0.001
	salinity_psu	0.0053	0.007	0.758	0.448	-0.008	0.019
985	cummulative_spike_volume_mikrol_per_l	0.0353	0.012	2.932	0.003	0.012	0.059
	dips	0.0125	0.008	1.662	0.096	-0.002	0.027

	Omnibus:	0.054	Durbin-Watson:	1.864
	Prob(Omnibus):	0.973	Jarque-Bera (JB):	0.147
990	Skew:	0.094	Prob(JB):	0.929
	Kurtosis:	2.612	Cond. No.	2.58

Notes:

995 [1] Standard Errors are heteroscedasticity robust (HC3)

E Results linear and multiple linear regressions with artificial surfactants from Sect. 3.5

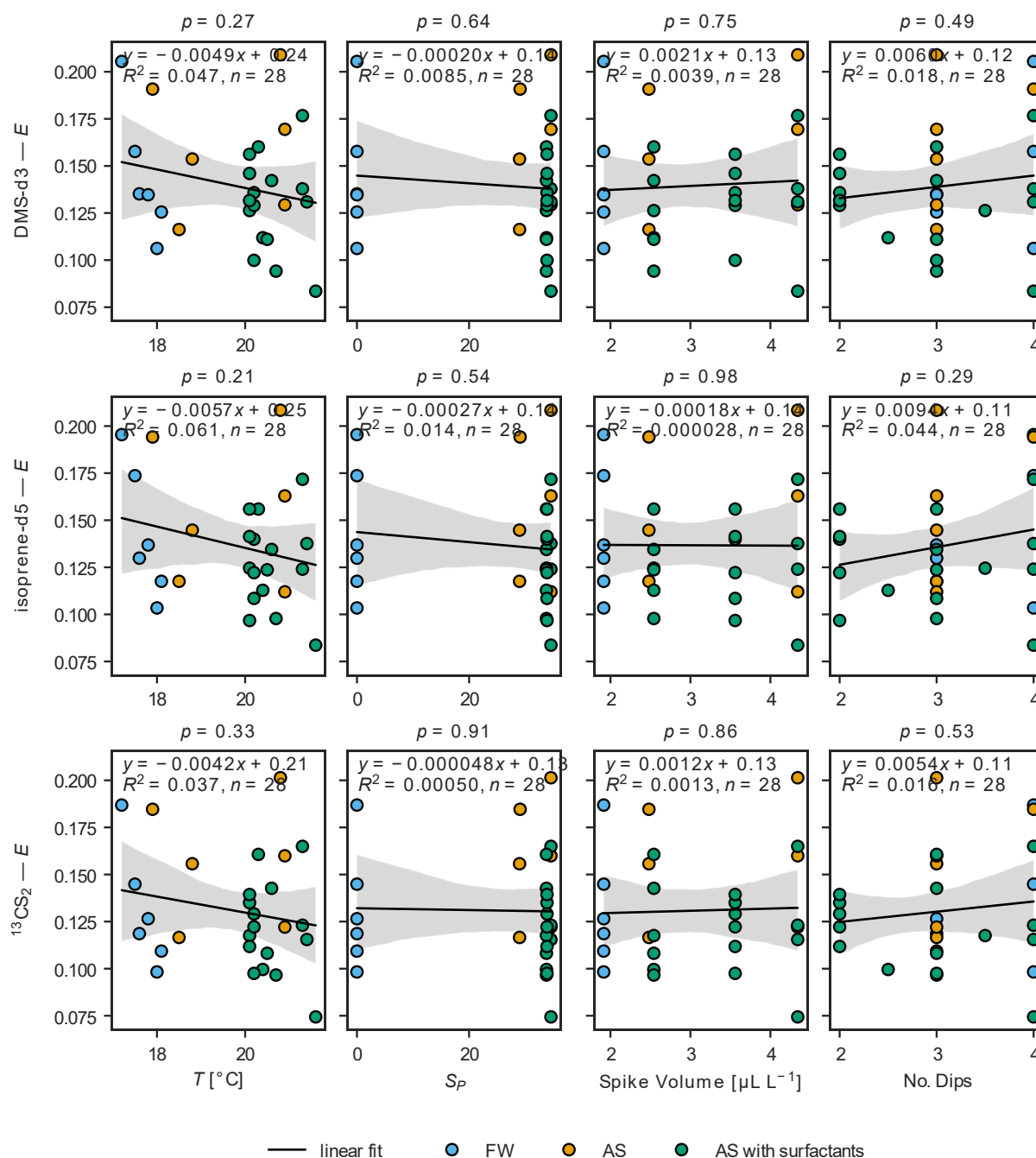


Figure E1 Linear regressions of sampling efficiency for DMS-d3 (top), isoprene-d5 (middle), and $^{13}\text{CS}_2$ (bottom) against water temperature, salinity, spike volume per litre and number of dips in fresh water (FW) and artificial seawater (AS) without and with surfactants treatment. Linear fit, R^2 and number of samples (n) are given in top corner of each subplot. Linear regression of surface activity (SA) is shown in Fig. E2.

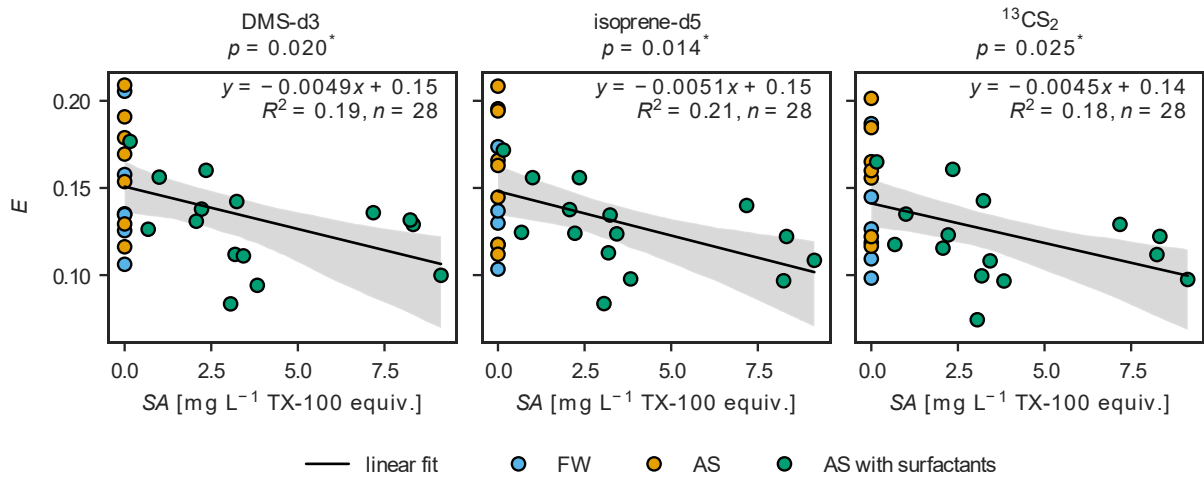


Figure E2 Linear regression of sampling efficiency for DMS-d3 (left), isoprene-d5 (middle), and $^{13}\text{CS}_2$ (right) against surface activity (SA) in SML in fresh water (FW) and artificial seawater (AS) without and with surfactants treatment. Linear fit, R² and number of samples (n) are given in top corner of each subplot. Linear regressions of water temperature, salinity, spike volume per litre and number of dips are shown in Fig. E1.

Table E1 Multiple linear regression results for DMS-d3 in experiment C including treatments fresh water (FW), artificial seawater (AS) without and with surfactants, using statsmodels smf.ols(...).fit(cov_type='HC3')

OLS Regression Results						
=====						
Dep. Variable:	ratio_sml_over_ulw_dmsd3	R-squared:	0.392			
Model:	OLS	Adj. R-squared:	0.247			
Method:	Least Squares	F-statistic:	2.006			
Date:	Sat, 01 Nov 2025	Prob (F-statistic):	0.119			
Time:	16:06:07	Log-Likelihood:	62.118			
No. Observations:	27	AIC:	-112.2			
Df Residuals:	21	BIC:	-104.5			
Df Model:	5					
Covariance Type:	HC3					
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	0.1393	0.006	23.395	0.000	0.128	0.151
temperature_water_degC	-0.0324	0.016	-1.995	0.046	-0.064	-0.001
salinity_psu	0.0152	0.013	1.158	0.247	-0.011	0.041
cummulative_spike_volume_mikrol_per_l	0.0213	0.010	2.184	0.029	0.002	0.040
dips	-0.0016	0.007	-0.242	0.809	-0.015	0.011
SA_mg_per_l_tx100_eq_sml	-0.0138	0.007	-1.960	0.050	-0.028	-3.4e-06
=====						



1030 Omnibus: 0.493 Durbin-Watson: 1.896
 Prob(Omnibus): 0.781 Jarque-Bera (JB): 0.067
 Skew: 0.115 Prob(JB): 0.967
 Kurtosis: 3.081 Cond. No. 5.21
 =====

1035 Notes:
 [1] Standard Errors are heteroscedasticity robust (HC3)

1040 **Table E2 Multiple linear regression results for isoprene-d5 in experiment C including treatments fresh water (FW) and artificial seawater (AS) without and with surfactants, using statsmodels smf.ols(...).fit(cov_type='HC3')**

OLS Regression Results

=====

1045	Dep. Variable:	ratio_sml_over_ulw_isoprened5	R-squared:	0.367
	Model:	OLS	Adj. R-squared:	0.216
	Method:	Least Squares	F-statistic:	1.921
	Date:	Sat, 01 Nov 2025	Prob (F-statistic):	0.133
	Time:	16:06:07	Log-Likelihood:	61.383
	No. Observations:	27	AIC:	-110.8
	Df Residuals:	21	BIC:	-103.0
1050	Df Model:	5		
	Covariance Type:	HC3		

=====

	coef	std err	z	P> z	[0.025	0.975]
1055 Intercept	0.1365	0.006	22.385	0.000	0.125	0.148
temperature_water_degc	-0.0309	0.016	-1.993	0.046	-0.061	-0.001
salinity_psu	0.0169	0.013	1.265	0.206	-0.009	0.043
cumulative_spike_volume_mikrol_per_l	0.0167	0.010	1.700	0.089	-0.003	0.036
dips	0.0016	0.007	0.231	0.817	-0.012	0.015
1060 SA_mg_per_l_tx100_eq_sml	-0.0132	0.007	-1.812	0.070	-0.027	0.001

=====

	Omnibus:	0.008	Durbin-Watson:	1.978
	Prob(Omnibus):	0.996	Jarque-Bera (JB):	0.186
	Skew:	-0.020	Prob(JB):	0.911
1065	Kurtosis:	2.596	Cond. No.	5.21

=====

Notes:
 [1] Standard Errors are heteroscedasticity robust (HC3)

1070



Table E3 Multiple linear regression results for $^{13}\text{CS}_2$ in experiment C including treatments fresh water (FW) and artificial seawater (AS) without and with surfactants, using statsmodels smf.ols(...).fit(cov_type='HC3')

OLS Regression Results						
=====						
1075	Dep. Variable:	ratio_sml_over_ulw_c13s2	R-squared:	0.377		
	Model:	OLS	Adj. R-squared:	0.229		
	Method:	Least Squares	F-statistic:	2.203		
	Date:	Sat, 01 Nov 2025	Prob (F-statistic):	0.0925		
	Time:	16:06:07	Log-Likelihood:	62.526		
1080	No. Observations:	27	AIC:	-113.1		
	Df Residuals:	21	BIC:	-105.3		
	Df Model:	5				
	Covariance Type:	HC3				
=====						
1085		coef	std err	z	P> z	[0.025 0.975]

	Intercept	0.1305	0.006	22.465	0.000	0.119 0.142
	temperature_water_degc	-0.0321	0.015	-2.191	0.028	-0.061 -0.003
	salinity_psu	0.0212	0.012	1.836	0.066	-0.001 0.044
1090	cummulative_spike_volume_mikrol_per_l	0.0159	0.010	1.534	0.125	-0.004 0.036
	dips	-0.0003	0.007	-0.047	0.963	-0.013 0.013
	SA_mg_per_l_tx100_eq_sml	-0.0137	0.007	-2.045	0.041	-0.027 -0.001
=====						
	Omnibus:	0.822	Durbin-Watson:	1.945		
1095	Prob(Omnibus):	0.663	Jarque-Bera (JB):	0.514		
	Skew:	0.333	Prob(JB):	0.774		
	Kurtosis:	2.882	Cond. No.	5.21		
=====						
1100	Notes:					
	[1] Standard Errors are heteroscedasticity robust (HC3)					

F Calculations for diffusive boundary layer from Sect. 4.1.2

1105 A simple box model was used to assess the effect of the diffusive boundary layer (DBL) on C_{SML} . A mass balance (Eq. (11)) was set up for the complete tank with flux (Neumann boundary condition) only through the surface (i.e., no-flux Neumann boundary conditions on all other sides).

$$\frac{dn(t)}{dt} = -J_{surface}(t) \cdot A \quad (11)$$

where n is the moles as (absolute) PA (proportional to mol), t is time in s after mixing stopped, J is the flux in $\text{mol m}^{-2} \text{s}^{-1}$ through the surface (defined as positive going inward), and A is the surface area of the tank in m^2 .

1110 The flux through the surface is derived using Fick's first law of diffusion (Eq. (12)). This assumes that the water in the tank was quiescent, $C_{air} \approx 0 \text{ mol L}^{-1}$, and the air is a perfect sink.

$$J_{surface}(t) = \frac{C_0}{\sqrt{\pi D t}} = \frac{n_0}{V_{tank} \sqrt{\pi D t}} \quad (12)$$

where $C_0 = n_0/V_{tank}$ is the initial concentration at $t = 0 \text{ s}$ in mol L^{-1} , D is the diffusivity in $\text{m}^2 \text{s}^{-1}$, n_0 are the initial moles in mol, $V_{tank} = AH$ is the volume of the water in the tank in m^3 , and H is the depth of the water in the tank in m.

The concentration in the ULW is used as initial condition, i.e., $C_0 = C_{ULW}$.

Analytical solution of the ODE in Eq. (11) by integrating is given in Eq. (13).

$$n(t) = n_0 \exp\left(-\frac{2A}{V_{tank}} \sqrt{\frac{Dt}{\pi}}\right) \quad (13)$$

1115 where $n(t)$ are the moles left in the tank at time t , and t is time in s after mixing stopped.

We calculate n_0 from mean PA per experiment by multiplying with the respective volume of the tank. Water volume in experiment B was $34 \text{ cm} \times 78 \text{ cm} \times 26 \text{ cm}$, and $42 \text{ cm} \times 61 \text{ cm} \times 31 \text{ cm}$ in experiment C. Diffusivity of DMS(-d3) ranged from $0.0000111\text{--}0.0000129 \text{ cm}^2 \text{s}^{-1}$ for the temperature range in this study. For the calculation, $D = 1.29 \times 10^{-9} \text{ m}^2 \text{s}^{-1}$ ($T = 20^\circ \text{C}$, $S_P = 0$, calculated for DMS according to (Saltzman et al., 1993)). Due to missing diffusivity values for isoprene(-d5) and $(^{13}\text{C})\text{CS}_2$, these two are skipped in the calculation. As time we (conservatively) set $t = 120 \text{ s}$, though the formation of a quasi-steady DBL only starts after the turbulent motion from the mixing has subsided, which is expected to take about 1 min. Finally, $n(t = 120 \text{ s})$ yields that $n(t)/n_0 = 99.9 \%$ of the moles are still left within the box (i.e., water volume) after 120 s.

1120 To estimate the dilution from the DBL on the glass plate, we partition this $n(t = 120 \text{ s})$ between ULW and DBL. Under the assumption that within this short time C_{ULW} did not change, i.e., $C_{ULW} = C_0$, we calculate how many moles of $n(t)$ are present in the DBL with Eq. (14).

$$n_{DBL} = n(t) - C_0 V_{ULW} \quad (14)$$

where n_{DBL} are the moles in the DBL, $V_{ULW} = A(H - h_{DBL})$ is the volume of the water below the DBL in the tank in m^3 , and h_{DBL} is the thickness of the DBL in m, approximated with Eq. (15).



$$h_{DBL} = a\sqrt{\pi Dt} \quad (15)$$

where h_{DBL} is the thickness of the DBL in m, $a = 3$ is a common coefficient that defines h_{DBL} in relation to ~99 % of C_{ULW} .

1130 Finally, the concentration in the DBL in relation to the concentration in the ULW, i.e. the concentration sampled by the glass plate is calculated by Eq. (16).

$$\frac{C_{DBL}}{C_{ULW}} = \frac{n_{DBL}}{V_{DBL}C_{ULW}} \quad (16)$$

where C_{DBL} is the concentration in the DBL, C_{ULW} is the concentration in the ULW in mol L^{-1} , and $V_{DBL} = A h_{DBL}$ is the volume of the water in the DBL in m^3 , which equates to 0.79 for both experiments.