Authors in purple, comments/reviews in black.

General

A few edits were made without prompt from the editor or reviewers in the process of revising the manuscript. For the sake of clarity minor text edits (e.g., wording, spelling, grammar) are not listed here, but are evident from the track-changes-version of the manuscript.

For the sake of readability any format changes in the text (e.g., italic, sub/superscript, dynamic-reference-to-static) and format changes/corrections of Figure, Table, Section, and Appendix references were not marked up.

One major change is a new subsection in *Section 4 Discussion*, that accounts for the request to discuss more and better nuanced the way of applying our findings of sampling efficiency to field samples (RC1 Specific Comments - 1.), as well as to consistently address the three posed study questions (1)-(3) in the *Introduction* of the manuscript, of which the third was not represented well in the logical structure of the manuscript.

- *Section 4.2 Is a meaningful analysis of glass plate samples possible?*
  (1) Are SML samples from the glass plate meaningful for trace gases?
- *Section 4.3 What drives sampling efficiency?*
  (2) If so, what drives the associated losses of those volatile compounds?
- *Section 4.4 How to assess trace gas enrichment in the field?*
  (3) Can we derive a method to assess the SML enrichment accurately in experiments and in the field?

Please refer to Sect. 4.4 in the revised manuscript for this new subsection.

Another major change is that the complete Appendix was moved to Supplements (as advised by the editor). A track-changes-version of the Supplement is available upon request from the corresponding author, as it cannot be uploaded in the submission form for major revisions.

Point-by-point reply

Editor's comment

The referees have made important, sometimes identical, criticisms of the paper. The reponses the authors give to these criticisms are generally satisfactory, except where the approach taken (e.g. only a single surfactant used) does not allow a full answer. The authors should diligently revise the paper on the lines they propose and try to see whether more justification can be given to address the common issues raised by both referees.

We diligently updated the manuscript with the changes we proposed to the reviewers's suggestions, see (the updated) reply to RC1 and RC2 below. As this was the major point addressed in the reviews and the editor's comment, we'd like to specifically mention here that we added specifics/nuances or elaborated further with respect to the single-surfactant usage in the following sections (details can be found in the replies to RC1 and RC2):

- *Section 2.1 Experimental setup*: why TX-100 and how does it compare to natural surfactant pools
- *Section 4.1.1 Mixing in the tank*: discussing how the added surfactant might have affected our premise of well-mixedness
- *Section 4.1.3 Influence of surfactants addition*: added other properties than *SA*, in view of better representing natural surfactant pools
- *Section 4.3 What drives sampling efficiency?*: limitations of *SA* concept and how to extend it to tease apart the actual drivers present in natural surfactant pools
- *Section 4.4 How to assess trace gas enrichment in the field?*: new subsection, also includes a brief recap of effect of surfactants on sampling efficiency, but repeats that (as of now) application of sampling efficiency in the field cannot account for any effects of surfactants
- *Section 5 Conclusion and outlook*: repeat that natural surfactant pools were beyond the scope of this study, list briefly the shortcomings of a single surfactant, conclude that nevertheless the study sheds insights, though transfering to the field is limited

We also revised the language in the manuscript to better reflect the nuances requested from the reviewers when presenting the surfactants treatment and oversaturation. Consequently, we updated the *Abstract* as well. We hope that these changes sufficiently complete the presentation of the single surfactant used. The manuscript was similarly revised in view of the oversaturation as experimental limitation, yet, the changes were less extensive than wrt to the surfactants and we refer to the replies to RC1 and RC2 below or the track-changes-version of the manuscript for those.

RC1

RC1: 'Comment on egusphere-2025-5361', Anonymous Referee #1, 08 Dec 2025

Comments on “Glass plate sampling efficiency for trace gases in the sea surface microlayer” EGUsphere-2025-5361

General Comments

This manuscript addresses an important methodological gap in sea surface microlayer (SML) research by quantifying the sampling efficiency of the glass plate technique for short-lived trace gases (DMS, isoprene, CS_2). The topic is highly relevant for SML biogeochemistry and air–sea exchange studies. The manuscript is generally well structured, and the authors make a commendable effort to analyse their data comprehensively.

However, there are several limitations that prevent the study from supporting some of its conclusions in the current form. The experimental configuration is constrained by practical conditions (homogeneous tank, no natural organic enrichment, oversaturated conditions) and does not fully represent the complexity of natural SMLs. The key assumption that SML and ULW have identical concentrations ($C_{SML} = C_{ULW}$) requires further clarification, especially under oversaturation and in surfactant treatments. Moreover, differences between experiments are confounded by changes in tank material, location, and artificial sea salt formulation. In addition to the major points above, several specific issues in the manuscript should also be addressed.

Thank you for the thorough comment! We have carefully revised our manuscript with the suggestions given. Next to improving the presentation of methods and conclusion as pointed out in the suggestions, we especially revisited the *Abstract* and *Section 5 Conclusion and outlook* to pose the experimental constraints more clearly. We agree with the list of limitations provided by the reviewer, and improved the manuscript as follows:

1. limitation: experimental configuration is constrained
   We revised the *Abstract* and *Sect. 5 Conclusion and outlook* to include the limitations more clearly. As changes were mostly a few words long, please refer to the track-changes-version of the revised manuscript. Additionally, we also revised *Section 4.2 Is a meaningful analysis of glass plate samples possible?* as we realized that the same reasoning applies there (line 581, 657). Further, we added a longer part in *Sect. 5* (line 674) to include a more detailed conclusion about the surfactant treatment; we added after
   “Given the consistency we conclude that meaningful analysis of SML trace gas samples in the field is possible”: “ for DMS, isoprene, and CS2. The complexity of natural surfactant pools was beyond the scope of this study. The approximation with an artificial surfactant (mostly exceeding naturally occurring concentrations) revealed only a mild (negative) effect of SA on sampling efficiency, for which the mechanisms could not finally

be identified. The relatively low effect might be different in the field, as not SA itself, but the variety of physicochemical properties present in natural surfactant mixes may also play a role that was neither captured nor studied in this study. In view of the relatively low (negative) effect on sampling efficiency, despite the mostly unnaturally high SA in this study, we propose that the sampling efficiency determined without surfactants can be used as a first approximation for field samples (e.g., independent of the presence of surfactants)."

2. limitation: experimental configuration does not fully represent the complexity of natural SML

   We improved the presentation of the surfactant configuration in *Section 2.1 Experimental setup* (*Methods*) by rephrasing the second research objective (line 118) and by adding the reasoning for the choice of a single surfactant (line 124ff): "The wet (soluble) surfactant Triton X-100 (TX-100) was used as an artificial surfactant to approximate the effect surfactants have on the water matrix, e.g., lowering of surface tension, physicochemical interactions of molecules. In order to keep the experiment as controlled as possible we chose a single, artificial surfactant, previously used as model surfactant in other SML studies because it is soluble. In addition, it is the substance used to calibrate voltammetry measurements. Consensus of how to best model the complexity of natural surfactant pools in lab experiments is still lacking and determination of this was beyond the scope of this study.".

   We also made minor revisions to *Section 5 Conclusion and outlook* to emphasize better the physicochemical focus in the surfactant treatment, as well as the empirical nature of our results (e.g., lines 668, 669, 670, see track-changes-version of manuscript). We added a new sentence (line 693):

   "Classification of natural surfactant pools is one of the goals of a potential continuation of the BASS project (application under review)."

   Additionally, we rephrased to "an artificial surfactant" (as opposed to "artificial surfactants") in the whole manuscript, where appropriate, to highlight the limitation of the single surfactant.

   These changes also address a comment in RC2.

3. limitation: identical concentrations require further clarification

   We revised the manuscript carefully to identify where further clarification is needed. We left line 96ff (*Sect. 2 Methods*) untouched (i.e., no mention of oversaturation or surfactant added as in Methods only the pure concept of the experiment is explained), as well as line 350ff (*Sect. 3.2 Mixing in the tank*) was left as it is because results should not include discussion. We added a new paragraph at the end of *Section 4.1.1 Mixing in the tank (Discussion)* (after line 517) to clarify the key assumption in view of oversaturation and surfactants:

   "Though close to the surface, the samples shown in Fig. 3 might overlook gradients in the SML, especially for the strong oversaturation present and in the absence of wind shear. We can only hypothesize as there are no sensors to measure the trace gas concentration of DMS, isoprene, or CS2 (or their isotopically labelled forms) in the SML. To assess the strength of a potential gradient forming in the SML in the 2 min waiting time after mixing, we estimated the diffusive boundary layer losses (Sect. 4.1.2). The addition of the artificial surfactant further complicates the hypotheses. The surfactant did not mix well into the SML with the chosen mixing approach (i.e., mostly EF ≠ 1, Table 2), potentially indicating that the mixing was also less effective for the trace gases as compared to treatments without surfactants. As a consequence, the gradients in the SML that have formed during the rest times before mixing (usually >45 min) are not fully homogenized, causing PASML < PAULW (and not PASML = PAULW), decreasing sampling efficiency artificially as a result (as observed in Fig. 6). On the other hand, the presence of surfactants might also cause an accumulation of trace gases in the SML, smoothening the gradients present, by reducing gas exchange, balancing a potential less effective mixing. Furthermore, physicochemical interactions between trace gases and the artificial surfactant are possible, affecting gradients and mixing positively or negatively. Given the uncertainty we interpret the results in the treatments with an artificial surfactant under the assumption that trace gases (but not the artificial surfactant) were as well-mixed as in the treatments without surfactants, and discuss accordingly the limitations."

4. limitation: differences between experiments are confounded
   We agree that the difference in experimental setup make the analysis more difficult. We considered this early on and decided to still include experiment A (as Supplement), as we believe that the results are nevertheless meaningful, and experiment B, as there was no indication that necessitatet (allowed) omission. At the end of *Sect. 4.1 Experimental setup (Discussion)* we added a new paragraph that discusses the significance of the differences between experiment B and C (including the tank material and locations), as well as the implications:
   "Each experiment was conducted at a different location (and time of year, Table 1). The tanks were placed indoors (the roof at ICBM in experiment B closes all around), to make sure that results are independent of location (and time of year), and consequently comparable between experiments. Air properties (e.g., temperature, humidity) likely varied (though not measured), but due to the strong concentration gradient we deem their effect negligible. The material of the tanks was either glass or plastic (Table 1). Contamination was ruled out, since we were using isotopically labelled trace gases. Neither of the materials used is known to react with the studied isotopically labelled trace gases or the added artificial surfactant. Adsorption to the surfaces likely differs between glass and plastic, as well as the exchange of heat across the material. These effects are accounted for by measuring water temperature and surface activity.
   The above noted differences in setup between experiment B and C make the interpretation of our results more difficult, given that no clear factor driving the differences between the experiments could be identified. The effect of the differences does not seem to overrule the overall trends (as seen in the combined data set of B and C in Fig. S6 of the Supplement). Throughout this manuscript we kept information about the source experiment visible (except for Table 7)."

Specific Comments

Thank you for the specific comments.

1. The manuscript proposes to use sampling efficiencies of 0.13±0.01 (DMS/isoprene) and 0.12±0.01 (CS_2) as thresholds to identify cases of net production in the SML. However, this recommendation requires significant caution. The experimental setup represents a specific scenario: a stagnant, homogeneous system with trace gas oversaturation. The steep concentration gradient driven by this oversaturation likely promotes rapid outgassing during the sampling process, potentially resulting in a lower sampling efficiency than might occur under near-equilibrium field conditions. Conversely, the lack of wind-driven turbulence in the lab might overestimate efficiency compared with rough seas. Therefore, these "thresholds" should be framed as system-specific empirical values derived from a high-outgassing regime, rather than general operational criteria.
   We rephrased the statement in the *Abstract* to indicate that more detail can be found in the *Discussion*, but refrained from further expanding the thought to avoid making the Abstract too long:
   "Instead, we discuss using~~suggest to use~~ the identified sampling efficiencies of 0.13 ± 0.01 (± standard error) for DMS and isoprene, and 0.12 ± 0.01 for CS2 as thresholds to identify cases of net ~~production~~enrichment in the SML."
   To frame the values in Table 7 (*Sect. 4.3 What drives sampling efficiency?*) as empirical, the following was added to the caption:
   "Treatment with an artificial surfactant is omitted here, since it remains uncertain whether our premise (well-mixed) was met. All trace gases were oversaturated. Isotopically labelled trace gases were used as analogues for the gases listed here."
   A completely new subsection was added: *Sect. 4.4 How to assess trace gas enrichment in the field?* that discusses the approach. While revising, we realized that the third of our questions raised at the end of the Introduction ("(3) Can we derive a method to assess the SML enrichment accurately in experiments and in the field?") is not accounted for in the logical structure of the manuscript, which we have done now by adding *Sect. 4.4*. Please refer to the respective section in the revised manuscript.
   In *Sect. 5 Conclusion and outlook* we made only minor changes (e.g., lines 678ff, see track-changes-version of manuscript). At the end of the paragraph in line 681 we added two sentences to hypothesize about the success of applying the proposed threshold-method:
   "The mesocosm study showed strong oversaturation and was conducted under a roof, therefore the application of sampling efficiency to identify net production is promising. In the case of the field cruises success is tentative in view of the experimental constraints under which sampling efficiency was derived."
2. The central definition of sampling efficiency assumes $C_{SML} = C_{ULW}$ so that any difference between SML and ULW samples reflects pure sampling artefacts. This assumption is particularly challenging under the experimental conditions of strong oversaturation and during surfactant treatment, where ongoing outgassing and the sampling sequence may introduce gradients not fully accounted for by the "well-mixed" test. It is therefore recommended that the Discussion more explicitly acknowledge this limitation, clarifying that the reported sampling efficiencies likely reflect a combination of true sampling artefacts and contemporaneous water-air exchange.
   At the end (line 596) of *Sect. 4.2 Is a meaningful analysis of glass plate samples*

possible? we added a few sentences to discuss that the sampling efficiency is mainly driven by pure sampling artefacts, and list which other effects are (a negligible) part of it: “The aim of this study was to capture pure sampling error (as sampling efficiency), for which we introduced the premise of a fully homogenized water body (i.e., CSML = CULW). The verification of the premise worked well, but in the absence of alternatives lacked the ability to account for gradients in the SML. In the 2 min waiting time after mixing and during sampling, gradients could have formed due to turbulent or diffusive fluxes through the air-water interface, both enhanced by the strong oversaturation present. As all experiments were performed indoors wind was not present, therefore the only potentially significant source for turbulent fluxes was the sampling itself, which we consider as part of the sampling efficiency. Diffusive losses were negligible. These two aspects are discussed in more detail in Sect. 4.1.2 and 4.1.3. Despite the discussed low contribution of the turbulent and diffusive fluxes, sampling efficiency as determined here remains a combination of pure sampling error and other on-going air-water exchange processes, to which the latter might even contribute as much as the uncertainty (Table 7) associated with sampling efficiency.”
With respect to the potentially overlooked gradients in the well-mixed test a new paragraph was added in *Sect. 4.1.1 Mixing in the tank* (as reply to “3. limitation”, see above), that discusses this more in depth.

3. The decision to omit explicit calibration and rely only on PA ratios is reasonable in principle, provided the GC-MS response is linear with negligible intercept over the concentration range used. However, the manuscript does not show any supporting calibration data, yet draws conclusions that depend on distinguishing efficiencies around 0.12–0.13. Even small deviations from linearity or a non-zero intercept could affect the ratio PA_SML/PA_ULW, especially if SML and ULW concentrations differ by a factor of ~8–10. A brief demonstration of linearity (in Supplementary Material) or a more explicit discussion of this assumption would strengthen the quantitative claims.
   We prepared a new section in the Supplements (*Sect. S2 Calibration curve examples gas chromatography–mass spectrometry*). A reference to this was added in *Sect. 2.3 Sample analysis trace gases* (line 192): “Example calibration curves are shown in Sect. S2 of the Supplement.”. Please refer the respective section in the Supplement.

Minor Comments

We went over the whole document for typos and mistakes. Please refer to the track-changes-version of the manuscript for any changes not suggested here.

Page 1, line 18 – “in the field. . Instead,” → remove extra period.
Done.

Page 4, Line 121: Two types of artificial seawater (Tropic Marin® Pro-Reef Sea Salt vs. ordinary aquarium salt) are used. It would help to explain the rationale for these choices and how their compositions differ (e.g. trace metals, alkalinity, additives). Since sampling efficiency differs between experiments B and C, the role of salt composition should be discussed more concretely.

At the first mention (*Sect. 2.1 Experimental setup (Methods)*, line 121) we added the link to the website describing the salt. Tropic Marin® Pro-Reef Sea Salt contains all trace metals, except the radioactive ones. It also contains calcium at around 430 mg/L, PO4 at 0.01 mg/L, and nitrate 0.7 mg/L. Nothing else than that salt was added to the water. Further, (line 121) “Depending on the availability, [...]” was added to make clear why two different salts were used. To the caption of Table 1 was added:

“The treatments with artificial seawater were prepared with either of two salts, depending on their availability. Details about composition are only available for the medium with Tropic Marin® Pro-Reef Sea Salt.”

In *Sect. 4.1 Experimental setup (Discussion)* changes were made to line 496ff and a few sentences added to discuss the implications of the water composition on the differences observed between experiment B and c:

“Additionally, we did not expect that~~ However,~~ the type~~choice~~ of salt (Tropic Marin® Pro-Reef Sea Salt vs ordinary aquarium salt) ~~may~~would 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~~. Differences in the FW treatments between experiments were not statistically significant, while for the AS treatment they were (Table 5). We do not attribute this to the composition of the salt (affecting the solubility differently than ordinary aquarium salt, e.g., (Weisenberger and Schumpe, 1996)), other substances, or the alkalinity of the water (see Table 1 for details of the salts), as neither are known to affect the isotopically labelled trace gases used in this study. Additionally, they are not expected to alter the physicochemical properties enough to affect sampling efficiency as they are dosed in natural concentrations. As the three trace gases exhibit a similar effect, something more general appeared to change between the AS experiments. It is, therefore, more likely that the differences are attributable to the combined effect of changes in (environmental) parameters between experiments, which were not ideally overlapping across all treatments (see Fig. S6 of the Supplement), e.g., changes in number of dips, water temperature, different glass plate size and wiper, mixing in of the trace gases, degree of oversaturation. In that respect, the differences between B and C might also be explained by a smoother sampling routine in C due to the practice gained in B.”

Page 6, Line 140: The text notes that a single glass plate dip yields 1.8–6.3 mL of medium, leading to substantial variability in sample volume per dip. Could you clarify what factors drive this large range in collected volume per dip? Understanding these drivers is important for assessing whether volume variability introduces bias to sampling efficiency.

In *Sect. 2.2 Sampling* (line 140f) we split the range “1.8-6.3 mL” into two, as the presence of surfactants is the main driver: “A single dip provided between 1.8–4.0 mL of medium when no surfactants were present, and 3.4–6.3 mL when an artificial surfactant was added.”. We added a few sentences in *Sect. 4.1.3 Influence of surfactant addition (Discussion)* to discuss the variation in volume per dip:

“The presence of the artificial surfactant increased the yield per dip significantly from about 3 mL to about 5 mL on average (Fig. 1). A larger yield per dip means less dips to reach the targeted sample volume, consequently reducing the sampling duration. This could lead to an increase in sampling efficiency, since the vial itself is left open for a shorter period. But since with the addition of the artificial surfactant also the targeted sample volume was increased to 15 mL (to satisfy voltammetry requirements), the sampling duration actually increased as compared to the treatments with 10 mL targeted. We, thus, hypothesize that sampling efficiency should decrease in consequence. Though the effect is deemed negligible as sampling duration was not identified as significant in the MLR."

Page 7, Line 190: The decision to use peak area ratios without GC-MS calibration is stated, but the assumption of a linear relationship between peak area and concentration is not validated. Including a brief note on how this assumption was tested or acknowledging its limitations would enhance transparency.
See "Specific Comments - 3."

Page 13 Line 337: Correct figure reference formatting. Several places refer to "Fig. Figure X" (e.g. Figure 1, Figure 2, Figure 3, Figure 5, Figure 6, Figure 7), which is clearly a typesetting artefact.
All references to figures are now static and corrected to the requested formatting.

Page 21, Lines 490–495: The total water temperature range is modest (ΔT ≈ 5.9 °C). You might add that this limited range likely contributes to the relatively weak apparent temperature effect in the raw data, despite statistical significance in the MLR.
We revised *Sect. 4.3 What drives sampling efficiency?* and added (line 627):
"Although, the trend is captured there is only a weak effect of water temperature on sampling efficiency, despite the statistical significance in the MLR, due to the modest range of water temperature covered in this study." "Although, the trend is captured there is only a weak effect of water temperature on sampling efficiency, despite the statistical significance in the MLR, due to the modest range of water temperature covered in this study."

Citation: https://doi.org/10.5194/egusphere-2025-5361-RC1

RC2

RC2: 'Comment on egusphere-2025-5361', Anonymous Referee #2, 07 Jan 2026

Review EGUsphere Lange et al.

This manuscript addresses an important topic, trace gas sampling in the Surface Microlayer (SML), through a systematic laboratory approach with the aim to characterize sampling losses when the SML is sampled using a glass plate. The manuscript is generally well written and the authors have well-constructed their argumentation and reasoning. However, the manuscript also demonstrates that the mechanisms regulating air-sea exchange are very complex and not fully understood yet. Therefore, several limitations prevent a wide application of the evidenced sampling efficiencies.

In particular the influence of surfactants is complex and the choice of one surfactant is limiting the general representativeness of the sampling efficiencies, particularly in cases where surfactants are present (as is generally the case in natural environments).

Thanks for the thorough comment! We observed much overlap between RC1 and RC2 with respect to the points addressed and also the indications on how to resolve them, which strengthens the arguments given in both reviews. We indicated which of our replies were similar in RC1. We reworked the *Abstract* and *Sect. 5 Conclusion and outlook* in view of the comments made in this review. We carefully revised our manuscript with the suggestions given and changed it as follows.

There seems to be a known and recognised bias in the experimental design where the oversaturation of the trace gas in the aqueous phase and zero-air above, will introduce a very strong gradient and may not uphold the well-mixed assumption. While this is discussed in the manuscript, this should be mentioned earlier on, e.g. in the experimental set-up, to better frame the limitations of the sampling efficiencies that are found.
We revised *Sect. 2.1 Experimental setup* and *Sect. 2.2 Sampling (Methods)* to address potentially developing gradients early on.
Line 123ff: “, but in all cases the trace gases in the water were oversaturated, causing a steep gradient towards the near-zero atmosphere.”.
Line 133: “to homogenize any gradients that had formed towards the air side due to the oversaturation”
Line 157 to highlight the limitation of the well-mixed assumption: “These tests were necessary, as the concentration difference between the oversaturated water and the near-zero air is large and might re-establish gradients in the SML quickly.“

While the manuscript clearly motivates the use of the peak area and not the concentrations of the gases, it would be good to indicate at least the LoD for each gas and the system’s sensitivity for each. This will help interpret the magnitude differences for PAs and their ratios discussed and give a better estimation of the environmental relevance of the method. Indeed, without these information, the differences in PA between trace gases explained as different amounts of

trace gas in the sample, could also be due to different trapping and ionisation efficiencies between these very different gasses (as for example shown also in Wang et al., 2023, DOI: 10.1016/j.marchem.2023.104206)Therefore, the information on how sensitive the purge and trap GC-MS system is, seems necessary.
We prepared a new Sect. in the Supplements (*Sect. S2 Calibration curve examples gas chromatography–mass spectrometry*). A reference to this was added in *Sect. 2.3 Sample analysis trace gases* (line 192): “Example calibration curves are shown in Sect. S2 of the Supplement.”. Please refer the respective Sect. in the Supplement. This was requested in RC1. Figure S3 gives an order of magnitude of concentrations.
We carefully revised the full manuscript and do not find any mention of comparing PAs of one trace gas to another, as we are aware that this comparison is not valid. We now state this explicitly in *Sect. 3.2 Mixing in the tank* (line 374):
“[The order of magnitude for PA varies between trace gases due to different amounts of trace gas in the spike] and different trapping and ionisation efficiencies. Therefore, while comparisons of PA within the same trace gas are valid, direct intercomparisons across trace gases are not possible due to the skipped calibration.”

Although the influence of the surfactant on the sampling efficiency seem not very high, it remains a relatively badly understood system. The use of the surface activity might be a convenient choice, but the SA might not be the best indicator for the exchange of gases, as it will also highly depend on the nature of the surfactant, which this method does not qualify. Different (mixes) of surfactants will behave differently towards different gas species, depending on their physicochemical nature. As is mentioned in the conclusion, the influence of the surfactants needs to be further studied, but not only with regard to their concentration (which was attempted here) but also their nature (ionic, nonionic). It would be nice to add this nuance in the text, or in the conclusion.
We agree now, with hindsight, that the surface activity is insufficient when addressing the influence of surfactants on trace gas properties in the medium. Unfortunately, we cannot fix this problem for this study, but we will address this as relevant for future studies. In *Sect. 4.3 What drives sampling efficiency? (Discussion)* we added to nuance the discussion better in view of the complex nature of natural surfactant mixes (line 657):
“Furthermore, SA is a good choice to understand how likely (mixes of) substances are to accumulate at the surface, yet, it overlooks other physicochemical processes potentially affecting sampling efficiency, like ionic/non-ionic characteristics. In view of the relatively low effect on sampling efficiency, this highlights that a single and artificial surfactant, though considered a good model surfactant, might not have been enough to represent the complexity of natural (mixed) surfactant pools.”
In the new subsection in *Discussion* (*Sect. 4.4 How to assess trace gas enrichment in the field?*) we included:
“In this study we find that the addition of an artificial surfactant decreases sampling efficiency, likely due to a larger surface area of the medium on the glass plate, though this remains speculative. The decrease might also relate to specific physicochemical properties of TX-100 (e.g., non-ionic, soluble) and the usage of a different surfactant might alter the effect on sampling efficiency.”

In *Sect. 4.1.3 Influence of surfactants addition* (*Discussion*) we added the nature of surfactants (soluble/insoluble, ionic/non-ionic, groups like lipids/amino acids/gel-like particles) as potential drivers to nuance better our conclusion.
In *Sect. 5 Conclusion and outlook* a sentence is added to highlight that the classification of natural surfactant pools is a hot topic:
“Classification of natural surfactant pools is one of the goals of a potential continuation of the BASS project (application under review).“
Changes made here were aligned with requested changes in RC1.

Another suggestion for an additional nuance in the conclusion, would be to consider a non-linear approach to this dataset (or an extended) as well seen by? the multiple issues with linear assumptions that are not validated and the poor modelling with surfactants.
We added in *Sect. 5 Conclusion and outlook* (line 672):
“As a consequence of the differences between experiment B and C and the limited number of data points for some of the treatments, it was not possible to perform a meaningful non-linear assessment of the results.”
Further, (line 693) we indicate that the extension to a broader range of environmental factors should be accompanied by non-linear and more complex analysis.

Further minor comments:

L126-127: The text states that ‘Two different tanks were used… ‘, however table 1 lists 3 different tanks. The incubation bath should also briefly be described in the Experimental setup (2.1)
We agree that this is inconsistent usage. We added a footnote in Table 1 that experiment A is further described in *Supplement S1*.

L306: It is mentioned that on 3 occasions the EF was below 1; it is not clear to me if these values were included in table 2? Do the authors have an idea why these samples are different (First of a set? Temperature?...)
In Sect. 3.1.2 Surface activity and enrichment of surfactants we rephrased (line 306ff) to clarify that these three exceptions are not average values and that they are part of the averages presented in Table 2:
“consequently the SML was always enriched with three exceptions where EF < 1.0, once for each level of TX-100 added. Exceptions could neither be explained with sample sequence (i.e., first sample of day or after a long break), how often the tank had been mixed before already, (absolute) SA, amount of TX-100 added, age of the water (i.e., time since filling), contamination from sampling equipment, water temperature, salinity, nor by reported deviations from standard procedure in sampling protocols (e.g., person sampling, mixing times, unexpected events), in GC-MS measurement protocols (because each SML sample was first measured for trace gas content), or in voltammetry measurement protocols (e.g., handling errors, failures). Exceptions were, therefore, not excluded, yielding aon average EF from 3.7–11.”

L524: '…dilution by the DBL reduced the PA sampled by the glass plate to about 0.79 of the concentration in the ULW…' : Please clarify what are the units of this 0.79 as I first assumed this was percentage or a ratio but now think it might be absolute reduction in a.u. of the peak area?
It is a ratio. We rephrased at both occurences to remove confusion. Please refer to the track-changes-version of the manuscript.

Tables and Figures:

Figure 2: difference between different sample types is difficult to see due to small symbols
We changed the symbol.

Table 6: what is the meaning of the asterisks (*) in the table?
The info was missing (significance level). We added the explanation of the asterisk in the caption.

For all figure references: in the text is mentioned 'Fig. Figure x' ; this should normally just read 'Figure x'.
All references to figures are now static and corrected to the requested formatting. Same suggestion made in RC1.

Language and editorial:

L77: it's not clear to me what is meant by undersize here. Could you perhaps find a synonym?
It refers to the effect that particles can be too small to be captured by a technique. This is more relevant for organic material than trace gases actually. Therefore, in this context we removed the word from the list to avoid confusion.

L107 : … experiment A is described in Appendix A… please correct
Done.

L306: … was enriched on average EF from 3.7… this sentence does not seem grammatically correct. Please correct.
Done.

L625: …. expectations and findings by: this sentence seems not grammatically correct, please correct.
Done.