https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Meta-analytical insights into organic matter enrichment in the surface microlayer

Amavi N. Silva[1], Surandokht Nikzad[2,3], Theresa Barthelmeß[1], Anja Engel[1], Hartmut Hermann[4], Manuela van Pinxteren[4], Kai Wirtz[3,5], Oliver Wurl[6] and Markus Schartau[1]

[1]GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, 24143, Germany
[2]Trent University, Peterborough, ON K9L 0G2, Canada
[3]Institute of Coastal Systems, Helmholtz-Zentrum Hereon, Geesthacht, 21502, Germany
[4]Leibniz Institute for Tropospheric Research (TROPOS), Atmospheric Chemistry Department (ACD), Leipzig, 04318, Germany
[5]Christian-Albrechts University of Kiel, Kiel, 24148, Germany
[6]Carl von Ossietzky University of Oldenburg, Oldenburg, 26129, Germany

Correspondence to: Amavi N. Silva (asilva@geomar.de)

Abstract. The surface microlayer (SML), the uppermost ~1 mm water layer at the air-water interface, plays a critical role in mediating Earth system processes, yet current knowledge of its composition and organic matter enrichment remains scattered across disciplines. Here, we present the first known meta-analysis of SML studies that quantitatively assesses the distributional characteristics of selected organic compounds, including organic carbon and nitrogen, amino acids, fatty acids, transparent exopolymer particles, carbohydrates, lipids and proteins, through probability density estimates, central tendency metrics and correlations analyses. Our results confirm a preferential enrichment of nitrogen-enriched, particulate organic matter in the SML, highlighting the significance of compound-specific accumulation and selective enrichment patterns. We also observe that the enrichment of a given compound may exhibit notable variability that depends on distinct internal and external conditions. Our evaluation of enrichment factors (EFs) of various measurable compounds provides updated estimates for their typical values and ranges. While delving into the ability of EFs to reflect the partitioning of organic matter within the SML, we also critically examine their limitation in capturing trophic conditions. Based on these findings, we propose that future SML research should incorporate both absolute concentration changes and enrichment capacities in the SML, alongside their relative changes (as denoted by EFs), to more accurately interpret ecological implications. Additionally, our meta-analysis demonstrates the value of logarithmic data transformations and robust central tendency estimates, as essential tools for improving the statistical reliability, comparability, and representation of SML enrichment patterns.

1 Introduction

Approximately 70% of the Earth's surface is covered by a hydrated gelatinous 'skin' known as the surface microlayer (hereafter referred to as 'SML'; note that while this term is commonly used to denote the sea surface microlayer, in this study

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

it refers to the surface microlayer in both marine and freshwater systems), which is of a thickness of 1 – 1000 µm (Astrahan *et al.*, 2016; Hunter, 1980; Liss and Duce, 1997, 1997; Wurl *et al.*, 2009). Situated between the surface waters of all natural water bodies and the atmosphere, this uppermost multi-component layer (Astrahan *et al.*, 2016; Carlucci *et al.*, 1985; Cunliffe *et al.*, 2013) creates a unique microhabitat, mainly consisting of neuston (i.e., living communities in the SML), a relatively enriched complex of organic compounds and strong physico-chemical gradients (Cunliffe *et al.*, 2013; Dietz *et al.*, 1976; Engel and Galgani, 2016; Hunter and Liss, 1977). The formation and the composition of the SML are governed by a number of biological, physical and chemical drivers that interact under varying complex environmental conditions and time scales. As a result, the SML dynamics play a pivotal role in a range of environmental processes such as air-water gas exchange, heat transfer across boundary layers, biogeochemical cycling, microbial interactions and distribution of pollutants (e.g., Engel *et al.*, 2017; Frew, 1997; Liss and Duce, 1997; Upstill-Goddard, 2006). Therefore, continued investigation of the compositional heterogeneity of the SML and of the processes therein is crucial to gain deeper insights into its role in ocean biogeochemistry and its potential climate interactions.

The primary source of compounds in the SML is the underlying waters (hereafter referred to as 'ULW'; Baastrup-Spohr and Staehr, 2009; Chen *et al.*, 2016) from where material is transported through diffusive fluxes, rising bubbles and buoyant particles (Joux *et al.*, 2006; Obernosterer *et al.*, 2005). In addition, wet and dry atmospheric deposition as well as *in situ* production and degradation also lead to concentration changes in the SML (Astrahan *et al.*, 2016; Kuznetsova *et al.*, 2004; Milinković *et al.*, 2022). Compared to the ULW, the SML is often enriched in organics (e.g., Baastrup-Spohr and Staehr, 2009; Gao *et al.*, 2012; Gašparović *et al.*, 2007; Liss and Duce, 1997; Marty and Saliot, 1976; Yang, 1999). Many of these compounds are surface active and are generally known as 'surface-active-agents' or 'surfactants' (Maki and Hermansson, 2020; Wurl and Holmes, 2008). Surfactants tend to adsorb at the air-water interface (Wurl *et al.*, 2009) due to their amphiphilic nature (i.e., presence of both hydrophobic and hydrophilic structural parts; e.g., Marty and Saliot, 1976), so that they reduce surface tension and can form stable interfacial films. When the SML becomes highly concentrated in surfactants, these films transform into thick surface slicks that are visible to the naked eye (Liss and Duce, 1997). Some of the naturally occurring surfactants in the SML include fatty acids, proteins, certain polysaccharides, humic-like substances and lipids (Brinis *et al.*, 2004; Marty and Saliot, 1976). In addition, inorganic ions, which do not preferentially adsorb at the air-water interface, can be also present in the SML due to passive upward transport (Knipping *et al.*, 2000; Petersen *et al.*, 2004). Furthermore, sticky microgels, like transparent exopolymer particles (TEP) that originate from bacteria and phytoplankton (Alldredge *et al.*, 1993), are also found in the SML. Such gel-like particles can form through the coagulation of dissolved polysaccharide (Engel *et al.*, 2004; Mari and Burd, 1998; Schartau *et al.*, 2007), and are capable of incorporating other compounds into a cohesive matrix (Cunliffe *et al.*, 2009; Sieburth, 1983; Wurl and Holmes, 2008), thereby enhancing the structural integrity of surface films (Cunliffe and Murrell, 2009).

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Liss and Duce (1997) and Pereira *et al.*, (2018) argue that the SML can restrict diffusive fluxes across the air-sea interface, substantially contributing to reduced rates of ocean-atmosphere gas exchange. Surfactants can impact air-sea gas exchange of greenhouse gases such as carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O) and dimethyl sulfide (DMS) (Frew, 1997; Upstill-Goddard, 2006). Asher (1997), from laboratory measurements, and Tsai and Liu (2003), from global ocean observations, estimate a reduction of annual net CO_2 flux by ~20% – 50% due to the presence of the SML, while Wurl *et al.*, (2016), from i*n situ* measurements, propose that this decrease can be ~15%. Barthelmeß *et al.* (2021) observed that, in a newly upwelled filament off Mauritania, surfactants can suppress CO_2 gas exchange by 12%. Both lab- and field-based experiments find that natural slicks can reduce air-sea gas exchange by 50 – 60% (Goldman *et al.*, 1988; Salter *et al.*, 2011; Mustaffa *et al.*, 2020), causing the SML to drive an overall reduction of 19% in the CO_2 fluxes, as shown by *in situ* observations (Mustaffa *et al.*, 2020). Supporting earlier findings of Springer and Pigford (1970), McKenna and McGillis (2004) and Sabbaghzadeh *et al.* (2017), who raised concerns about the impact of the SML's surfactants on uncertainties in air-sea gas exchange models, Mustaffa *et al.* (2020) further argue that conventional wind-based models miscalculate CO_2 exchange up to 20% in areas with high surfactant concentrations. Moreover, Kock *et al.* (2012) find that, in the eastern tropical North Atlantic region, offsets between air-sea and diapycnal N_2O fluxes could be explained when surfactant effects were introduced to gas exchange models. Work of Goldman *et al.*, (1988) find that surfactants in the SML can also suppress air-sea gas exchange of oxygen (O_2). Disparities in these studies emphasize the significance of accurately assessing the characteristics of the SML and its processes, as well as integrating this knowledge into climate relevant ocean-atmosphere models (Milinković *et al.*, 2022) in order to reduce uncertainties in global gas flux estimations, particularly given that SML is seldom included in gas exchange models (Cen-Lin and Tzung-May, 2013; Engel *et al.*, 2017).

Although the composition and the concentration of compounds within the SML are thought to be strongly correlated with those of the ULW (Basstrup-Spohr and Staehr, 2009; Chen *et al.*, 2016; Joux *et al.*, 2006; Kuznetsova *et al.*, 2004), certain substances are selectively accumulated at the air-water interface, leading to a pronounced enrichment in the SML. The accumulation of these specific compounds in the SML relative to the ULW is often described by the 'Enrichment Factor' (hereafter referred to as 'EF'). The EF of a compound 'x' is given by the following concentration ratio:

$$EF\ of\ x = \frac{Concentration\ of\ x\ in\ SML}{Concentration\ of\ x\ in\ ULW} \quad (1)$$

This equation proposes that when the x's concentration in the SML is higher than that in the ULW, the corresponding EF value becomes > 1 and vice versa, as discussed in Carlson (1983) and Garabetian *et al.* (1993). However, Basstrup-Spohr and Staehr (2009) observed that non-slick areas in which microbial degradation processes are dominant can also demonstrate higher EF values, resembling those found in slick conditions. Kuznetsova *et al.* (2004) observed that organic matter (OM) dynamics in the SML can be decoupled from the composition of the ULW due to factors such as different remineralization rates and

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

selective adsorption (as opposed to homogenous mixing between the two compartments). Hillbricht-Ilkowska and Kostrzewska-Szlakowska (2004), in their work on lakes, suggest that eutrophic ULW conditions can lead to lower SML enrichment when the waters are concentrated by autochthonous OM (i.e., originate within the same ecosystem they are found) that show a lower affinity to the air-water interface. Hardy, (1997) mentioned that the processes leading to SML enrichment are often similar in both marine and freshwater environments. Knulst *et al.* (1997), Münster *et al.* (1998) and Södergren (1987) found that the freshwater SML tends to be more enriched with organic carbon and nitrogen, total phosphorous, ammonia and phosphate ions, while in marine environments, SML enrichment is stronger for carbohydrates, lipids, proteins and amino acids (Liss and Duce, 1997). Overall, and given the diversity of these results, it is clear that a more holistic view of the applicability of EF as a valid and meaningful indicator of SML enrichment is warranted.

To address these objectives, we adopted a meta-analysis of SML-studies, and conducted a comprehensive analysis to (1) assess OM enrichment in the SML, (2) review existing EF estimates and (3) investigate the relevance of EF values as accurate indicators of SML enrichment. New and novel insights gained from these analyses are intended to establish a robust foundation for future modelling efforts focused on the functions of the SML and their implications for biogeochemistry and climate.

2 Methodology

The work presented here synthesizes findings from multiple studies on the SML and employs a quantitative meta-analysis. Such systematic reviews can provide a more precise and accurate understanding of overarching trends, even when individual studies report inconsistent results (Crocetti, 2016). Mengist *et al.*, (2020) highlight the importance of meta-analyses by stating that *"Systematic reviews with meta-analysis represent the gold standard for conducting reliable and transparent reviews of literature."*

2.1 Data collection and compilation

The primary dataset consists of 2055 data points, extracted from 30 peer-reviewed publications (hereafter referred to as 'reference studies') identified through a comprehensive and systematic literature search of scholarly articles published between 1967 and 2022. From these studies, data containing simultaneously collected SML and ULW concentrations (hereafter referred to as '$[C]_{SML}$' and '$[C]_{ULW}$', respectively) were extracted for twelve different observational types of organic compounds (hereafter known as 'target compounds'): total organic carbon (TOC expressed in mg L^{-1}), particulate organic carbon (POC in mg L^{-1}), dissolved organic carbon (DOC in mg L^{-1}), total organic nitrogen (TON in mg L^{-1}), particulate organic nitrogen (PON in mg L^{-1}), dissolved organic nitrogen (DON in mg L^{-1}), amino acids (AA in µmol L^{-1}), fatty acids (FA in µg L^{-1}), transparent exopolymer particles (TEP in µg Xeq L^{-1}), carbohydrates (CHO in µmol L^{-1}), lipids (in µmol L^{-1}) and proteins (in µmol L^{-1}). TOC pool includes all forms of organic carbon, thus comprising both POC and DOC. Similarly, the TON pool combines both PON and DON. In general, the particulate pool constitutes a minor fraction of the total pool. The major classes of biopolymers

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

are proteins, CHO and lipids, with AA serving as the monomers of proteins. Depending on the elemental composition of these biopolymers, they contribute to both, the organic carbon and/or organic nitrogen pool. While the ratio of these biopolymers is higher in the particulate pool, it usually declines to only a few percent in the dissolved pool. TEP is composed of polysaccharides (i.e., CHO) with a major fraction contributing to POC, while a minor fraction exists at the interface between the dissolved and particulate phases (Verdugo *et al.*, 2004).

The EF values for these target compounds were systematically calculated from corresponding $[C]_{SML}$ – $[C]_{ULW}$ pairs, using Eq. (1). In this study, $[C]_{SML}$, $[C]_{ULW}$ and EF data are collectively referred to as 'primary data'. Auxiliary information associated with the primary data, (i.e., sampling factors and environmental variables) were also extracted when reported. They are referred to as 'secondary data'. All the data, in general, were collected either (1) directly from the source when presented, or else (2) through digitization of graphs and plots using PlotDigitizer (https://plotdigitizer.com) and GraphClick v3.0 (https://graphclick.en.softonic.com/mac) or, (3) via author correspondence when data was not available in the published materials. The resulting compiled database is herein referred to as 'Surface Microlayer Data (SMD)'. Supplementary Table S1 provides an overview of the reference studies on which SMD is based.

2.2 Statistical analyses

Given that the SMD ranges over several orders of magnitudes, when the dataset is handled in linear-space (i.e. in its original form), higher values dominate and overshadow the features associated with lower values (Feenstra, 2006). These potential limitations of linear scaling were reduced by transforming our primary data into their logarithmic (log_{10}) counterparts. Hereafter, the term 'linear' refers to the original, untransformed data, while the term 'log' stands for their logarithmic equivalents. The following sections describe the subsequent analyses conducted in our work.

2.2.1 Probability distributions

Making inferences based on ratios such as EFs requires careful consideration, as changes in the numerator and the denominator often affect these ratios asymmetrically (Keene, 1995). In the context of this study, while reductions in $[C]_{ULW}$ can lead to unusually high EF values that can approach infinity (i.e., stretched towards higher values), increases in $[C]_{ULW}$ may produce EFs decreasing down to 0 (i.e., compressed towards lower values). This results in distributions that significantly deviate from Gaussian (i.e. normally distributed) shape. Therefore, distributional characteristics of the primary data were examined through probability distributions.

Probability density functions (hereafter referred to as 'PDF') of the EF values were examined by applying non-parametric Kernel Density Estimates (hereafter referred to as 'KDE'; Parzen, 1962; Silverman, 1986; Wegman, 1972). KDE employs a normalized weighting function – known as 'Gaussian kernel' – which is centered at each datapoint. The sum of these kernels produces a smooth and continuous PDF that fits the underlying data. Selection of the width of a kernel – known as 'bandwidth'

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

– is an integral part of the KDE approach, as bandwidths too small or too large lead to overfitting and underfitting of data, respectively, failing to capture the true patterns in distributions. Following this, optimal bandwidths for linear KDEs were computed based on Härdle *et al.* (2004). For log KDEs, a fixed optimal bandwidth was applied.

Robustness of the KDE method decreases at low sample size. Since the SMD contains variables with sample sizes as low as 16 (for proteins), a bootstrap resampling approach was adopted where 67% of the original data (i.e., 2/3 of the sample) were randomly subsampled and an individual KDE was generated at each iteration. The process was repeated 1000 times, each time with a different random subsample, generating a set of KDE. These were then averaged to produce an ensemble mean, from which the final PDFs were derived.

Additionally, cumulative distribution functions (hereafter referred to as 'CDF') were determined for $[C]_{SML}$ and $[C]_{ULW}$ from the ensemble means of the bootstrapped KDEs. Appendix A provides further information on the KDE method.

2.2.2 Summarization, comparison and correlation estimates of distributions

For describing, comparing and relating the resulting PDFs and CDFs, we used standard statistical measures. Their mathematical expressions are given in Appendix B.

(1) To describe the central tendencies, mode (hereafter referred to as 'x_m'), median (hereafter referred to as '$\tilde{x}$'), arithmetic mean (hereafter referred to as '$\overline{x_a}$') and geometric mean (hereafter referred to as '$\overline{x_g}$'), were computed.

(2) The values at 5th and 95th percentiles of each distribution (hereafter referred to as 'upper threshold: UT' and 'lower threshold: LT', respectively) were also estimated in order to determine their central 90% range (i.e., degree of spread).

(3) To numerically compare the $[C]_{SML}$ and $[C]_{ULW}$, Integrated Quadratic Distance (Hereafter known as 'IQD') values of their CDFs were approximated based on Eq. (B3), which measure how different the two distributions are with regard to symmetry and multimodality.

(4) To investigate and quantify potential relationships between $[C]_{SML}$ and $[C]_{ULW}$ of each target compound, their linear correlation was analysed by employing both parametric Pearson and non-parametric Spearman's tests (both methods were applied for cross-validation purposes; agreement between the two correlation coefficient values increases the confidence in the robustness of the observed relationship).

3 Results

Unless otherwise stated, all analyses were performed on log scale. Nevertheless, to avoid potential misinterpretation of log scales in data presentation, primarily due to their limited readability among non-expert audiences (e.g., Menge *et al.*, 2018), all results are presented on linear scale (unless otherwise stated).

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

3.1 Characterizing EF distributions

Figure 1 compares the KDE-derived PDFs of the EF values for the carbon-enriched (in blue) and nitrogen-enriched (in orange) organic compounds (Hereafter known as 'PDF_C' and 'PDF_N', respectively). PDF_C was derived from EF values for TOC, DOC, POC, FA, TEP and CHO. The remaining target compounds derive PDF_N. In their linear version (Figure 1(a)), both PDFs demonstrate positive skewness (i.e., right-skewness) with the two x_m values being 1.10 and 1.15, respectively. Nevertheless, the peak probability density of the PDF_C (i.e., the height of the PDF = ~1) is more than twice that of the PDF_N (~0.4). $\tilde{x}$ of the two PDFs vary substantially, with PDF_C and PDF_N yielding values of 1.25 and 1.75, respectively. The values for $\overline{x_a}$ (2.60 and 5.39, respectively) and $\overline{x_g}$ (1.51 and 2.34, respectively) further reflect this divergence. In contrast, their log-transformed versions (Fig. 1(b)) approximate normal distributions, with PDF_C estimating the (exponentials of) x_m = 1.11; $\tilde{x}$ = 1.24 and $\overline{x_a}$ = 1.51. The PDF_N yields corresponding values of 1.33, 1.76 and 2.34. Their peak probability densities also reflect that the PDF_C (~2.6) is twice as high as that of PDF_N (~1.5).

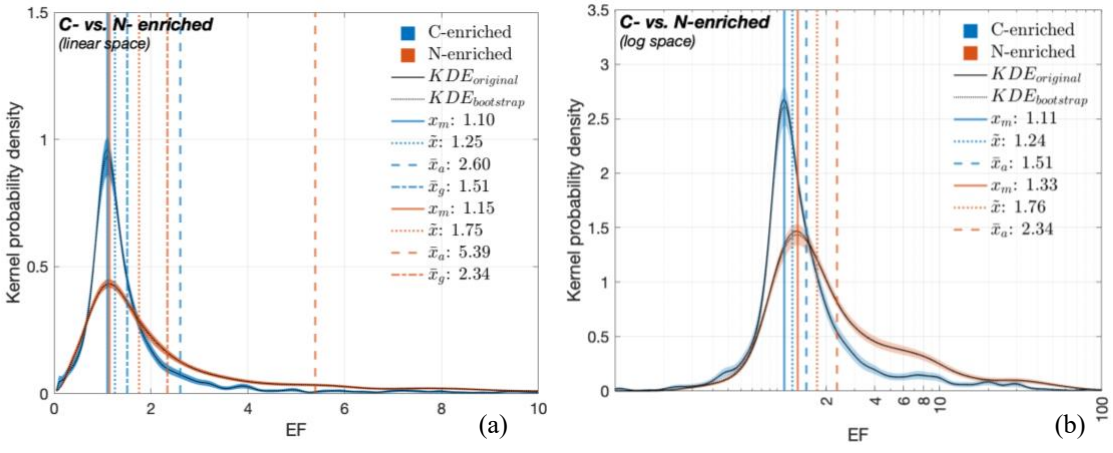

Figure 1: PDFs of the EF values for carbon-enriched (blue) and nitrogen-enriched (orange) compounds. PDFs of the (a) untransformed (i.e., linear) and (b) log-transformed EF values. The solid black line indicates the KDEs derived from original data while the dashed black line represents the ensemble mean of bootstrapped KDEs. Central tendency metrics (mode [$\boldsymbol{x_m}$], median [$\tilde{\boldsymbol{x}}$], arithmetic mean [$\overline{\boldsymbol{x_a}}$], geometric mean [$\overline{\boldsymbol{x_g}}$]) given in panel (b) are the exponentials of the corresponding estimates on the log scale.

We also compared EF-based PDFs (Figure 2) for dissolved (PDF_D, in purple) and particulate (PDF_P, in green) OM where we refer to a filter size of 0.22 μm (Gao *et al.*, 2012). At a linear scale (Figure 2(a)), the PDFs are again right-skewed for the two clusters, with characteristics: (1) 1.20 (both PDF_D and PDF_P) for x_m; (2) 1.50 and 2.25 for $\tilde{x}$; (3) 4.39 and 4.41 for $\overline{x_a}$ and, (4) 1.91 and 2.53 for $\overline{x_g}$, respectively. The peak probability density of the PDF_D (~0.6) exceeds that of the PDF_P (~0.3) by nearly a factor of two. The log PDF_D and PDF_P (Fig. 2(b)) approximate normal distributions alongside the following exponentiated central values, respectively: (1) x_m = 1.20 and 1.88; (2) $\tilde{x}$ = 1.48 and 2.29; (3) $\overline{x_a}$ = 1.91 and 2.53. Their peak probabilities compare between ~1.8 (for PDF_D) and ~1.0 (for PDF_P).

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

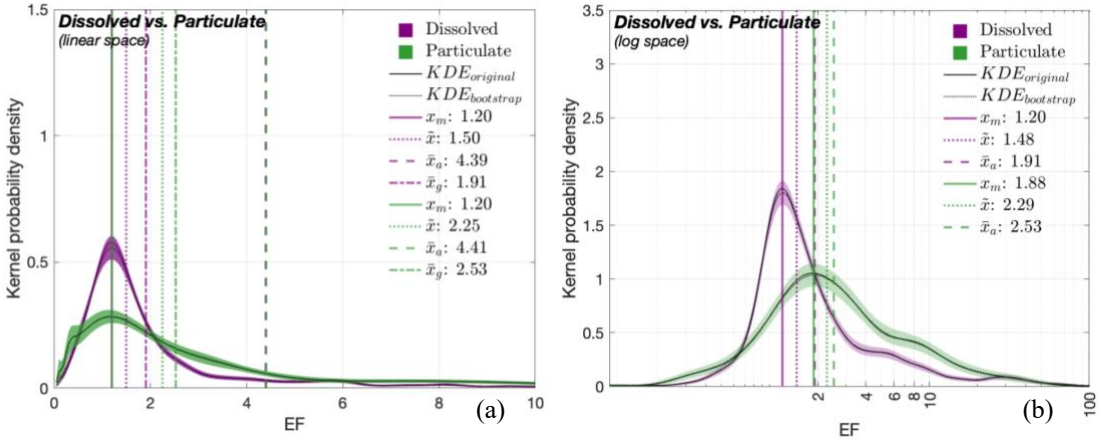

Figure 2: PDFs of the EF values for dissolved (purple) and particulate (green) compounds. PDFs of (a) linear and (b) log EF values. See Fig. 1 caption for details on KDEs and central tendency metrics.

Figure 3 displays the PDFs of the EF values for the target compounds. All the distributions exhibit nearly log-normal characteristics, nevertheless they vary in their degrees of spread. Here, only $\tilde{x}$ and $\overline{x_a}$ values estimate the central tendency of each distribution (the rationale for this approach is discussed in section 4.2). The values of $\tilde{x}$ (dotted pink line) and $\overline{x_a}$ (dashed pink line) are closely aligned in magnitude. According to these derived estimates, median and geometric mean EFs are largest for POC (Fig. 3(b): $\tilde{x}$ = 3.09; $\overline{x_a}$ = 3.22) across all the target compounds, with PON (Fig. 3(e)) and DON (Fig. 3(f)) following closely, each exhibiting $\tilde{x}$ and $\overline{x_a}$ values > 2. Although proteins (Fig. 3(l)) also show higher central tendency estimates, it should be noted that they have the smallest sample size (= 16), followed by lipids (sample size = 20). Therefore, the results of these two compounds should be interpreted with caution due to their lower statistical robustness. A comparison of threshold metrics (i.e., LT and UT; see section 2.2.2) reveals that the EF distributions for FA (Fig. 3(h)) and POC (Fig. 3(b)), exhibit the highest UT values (14.5 and 13.3, respectively) along with the greatest distributional variability. TOC (Fig. 3(a)) and TON (Fig. 3(d)) show the least variability among all target compounds. While some compounds exhibit well-defined unimodal EF distributions (e.g., POC, PON), few others (e.g., TON, AA) display polymodal patterns.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

3.2 Comparing SML and ULW concentrations

Figure 4 presents the CDFs of the ULW (in red) and SML (in blue) concentrations for the target compounds. A CDF exhibits how probability accumulates across a range of values (in the current context, $[C]_{SML}$ and $[C]_{ULW}$ data). All CDFs (both $[C]_{ULW}$ and $[C]_{SML}$) exhibit a characteristic sigmoidal shape: a slow initial rise (i.e., lag phase), followed by a steep rise (i.e., exponential phase), eventually reaching a plateau (i.e., stationary phase). CDFs for TOC display two distinctive plateaus indicating bimodal concentration distributions for both SML and ULW.

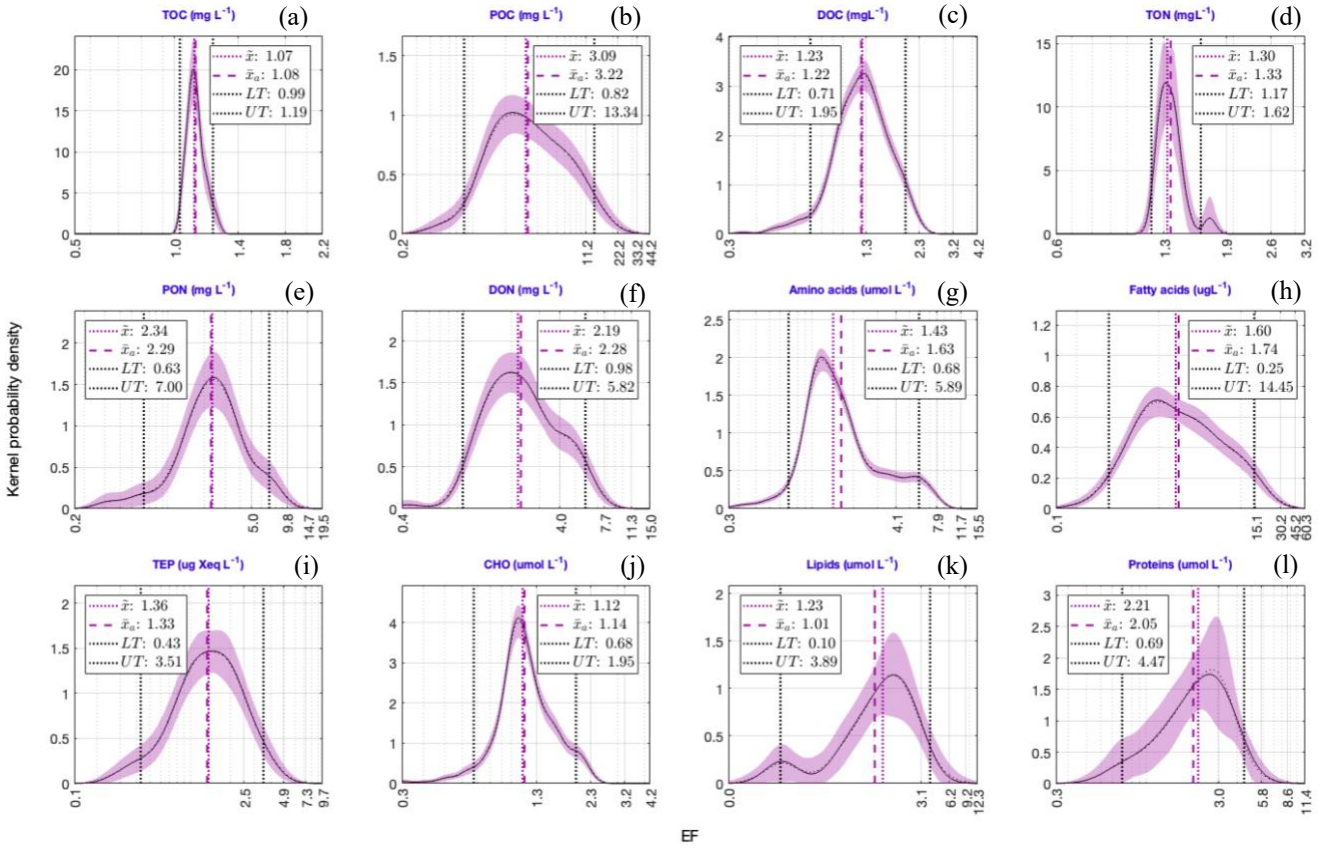

Figure 3: PDFs of the EF values for the twelve target compounds. The lower and upper thresholds of each distribution (dashed black lines) are defined by 5th and 95th percentiles of each PDF (see section 2.2.2). The values of these thresholds, along with the central tendency metrics given in each panel, are the exponentials of the corresponding estimates in the log space.

Additionally, despite the homogeneity in the general shape and trend of these CDFs, their corresponding IQDs (Fig. 5) reveal that the magnitudes of the offsets between $[C]_{ULW}$ and $[C]_{SML}$ distributions vary substantially across the target compounds. Lower IQD values indicate greater similarities between the CDFs, while higher values document clear distinguishability and thus also document a more robust enrichment signal. The lowest IQD is reported for the CDFs of TOC and CHO (0.005) while

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

that of DON yield the highest in value (0.184). In addition, lower CDFs (i.e. IQD < 0.05) are observed for lipids (0.008), DOC (0.012), TEP (0.018), AA (0.039) and TON (0.041), whereas POC (0.18) and proteins (0.15) exhibit a greater divergence (i.e. IQD > 0.15) between $[C]_{ULW}$ and $[C]_{SML}$. The color intensity of each bar reflects the sample size (n) of each target compound (i.e., smaller the sample, lighter the color).

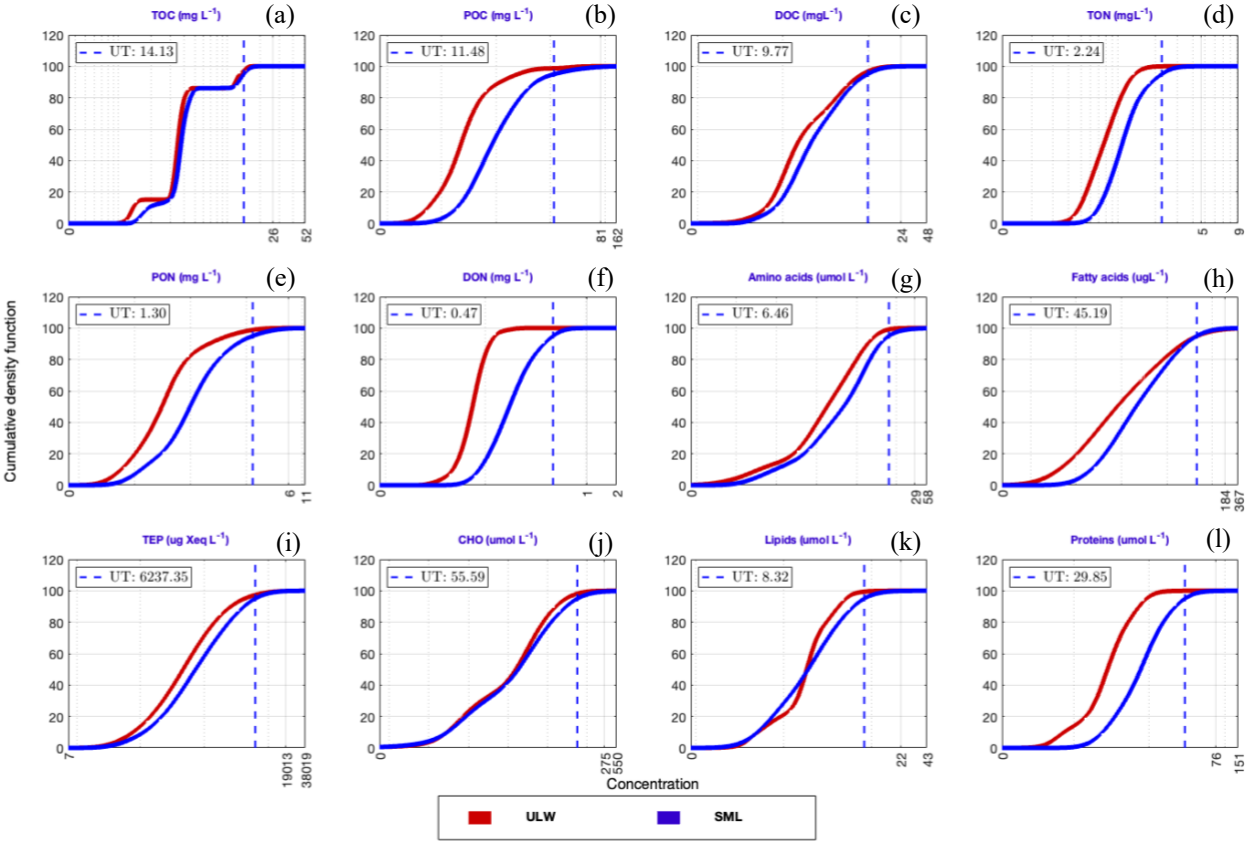

Figure 4: CDFs of the ULW (red) and SML (blue) concentrations for the target compounds. The upper thresholds for $[C]_{SML}$ (UT; given by blue dashed lines) are defined by 95th percentiles of the corresponding CDF. The values of these thresholds are the exponentials of the corresponding estimates in the log space. Their corresponding IQDs are given in Figure 5.

Correlations between $[C]_{ULW}$ and $[C]_{SML}$ of the target compounds were statistically estimated using liner correlation, as presented in Figure 6. The coefficients 'ρ' and 'r' stand for the correlation values derived from non-parametric Spearman's and parametric Pearson correlation tests, respectively. For all target compounds, except for DON, lipids and proteins, we found strong correlations between their SML and ULW concentrations (ρ and r > 0.5) with robust positive relationships. Individual datapoints for TOC, DOC and CHO (Figs. 6(a), (c) and (j)) closely fall on the 1:1 reference line where $[C]_{SML} = [C]_{ULW}$ (dashed

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

black line). In contrast, those for POC, TON, AA and FA are notably shifted towards the y-axis, suggesting higher $[C]_{SML}$ values relative to $[C]_{ULW}$ that corresponds to potentially enriched (depleted) SML (ULW) concentrations against ULW (SML) concentrations (see inset plot in Fig. 6(a)). Although TEP shows a slight enrichment in the SML, the effect is not particularly pronounced (Figure 6(i)). In addition, all the datapoints (regardless of whether they display copulation or not) were further color-coded according to their respective EFs. The results reveal an overall consistency in EFs across concentration ranges irrespective of their magnitudes. For example, in Fig. 6(c), EF values remain below 5, both when $[C]_{ULW}$ and $[C]_{SML}$ are < 0.5 mg L^{-1} and > 5 mg L^{-1}. This pattern holds across nearly all the target compounds.

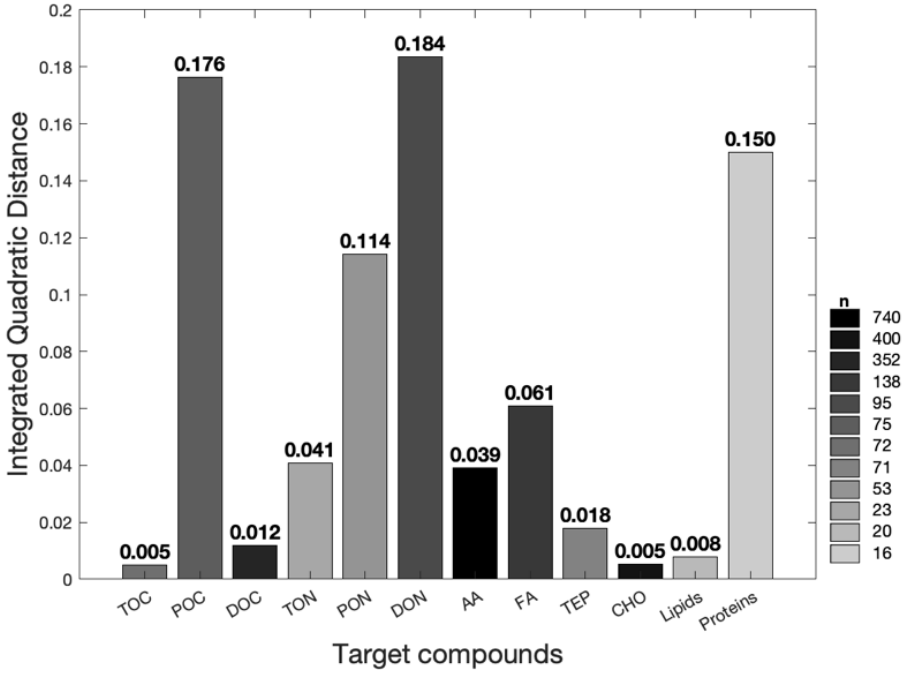

Figure 5: IQD values quantifying the divergence between ULW and SML concentrations for each target compound. The IQD represents the squared difference between ULW- and SML-based CDFs shown in Figure 4. Higher IQD indicates greater divergence between the two distributions and vice versa. Bar color intensity corresponds with the sample sizes.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

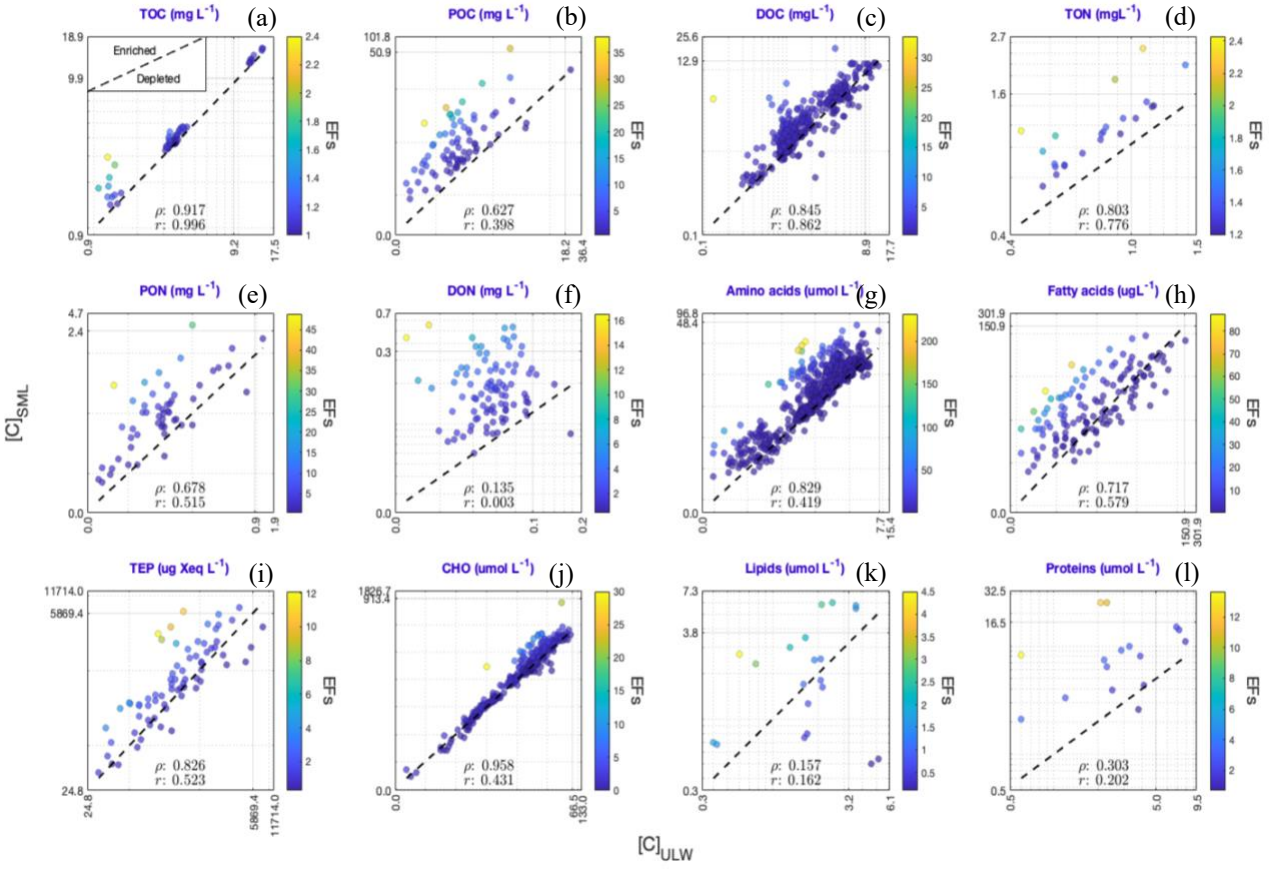

Figure 6: Linear correlation between $[C]_{ULW}$ and $[C]_{SML}$ for the target compounds. The datapoints are color-coded based on their corresponding EFs. Dashed black line indicates 1:1 line when $[C]_{ULW}$ (x-axis) = $[C]_{SML}$ (y-axis). Inset plot in panel (a) exhibits the relevant implications of the figure: Correlations above the 1:1 line correspond to selective SML enrichment and vice versa. The values of 'ρ' and 'r' give Spearman's and Pearson correlation coefficients, respectively.

3.3 Investigating concentration-dependent enrichment dynamics

Informed by the observations drawn from Fig. 6, Fig. 7 presents a more detailed investigation into the interrelationships among $[C]_{SML}$, $[C]_{ULW}$ and EFs in the environment. The analysis is restricted to $[C]_{SML}$ values (x-axis) that exceed the $\tilde{x}$ (i.e., median) of their respective distributions (median is the most stable central tendency metric of a distribution. Discussed further in section 4.2). These elevated $[C]_{SML}$ are compared against the corresponding $[C]_{ULW}$ (color scale) and EF (y-axis) values. The results reveal following covariation trends:

(i) TOC reports a generally low range of EF values comparable at both low and high concentrations of SML and ULW (Fig. 7(a))

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

(ii) DOC displays relatively consistent EF values regardless the magnitudes of $[C]_{SML}$ and $[C]_{ULW}$ (Fig. 7(c)), but also slightly points towards higher EF values in association with low $[C]_{ULW}$

(iii) DON presents an ascending EF gradient, positively correlated with $[C]_{SML}$ (Fig. 7(f)), revealing more enrichment to be well reflected in the concentrations found in the SML

(iv) AA shows a similar correlation dependence to that of DON, but also reveals a much clearer trend toward higher EF values to be found at lower $[C]_{ULW}$ concentrations (Fig. 7(g))

FA (Fig. 7(h)), despite their larger sample sizes, fail to display conspicuous consistent trends in the $[C]_{SML}$ – $[C]_{ULW}$ – EF triad.

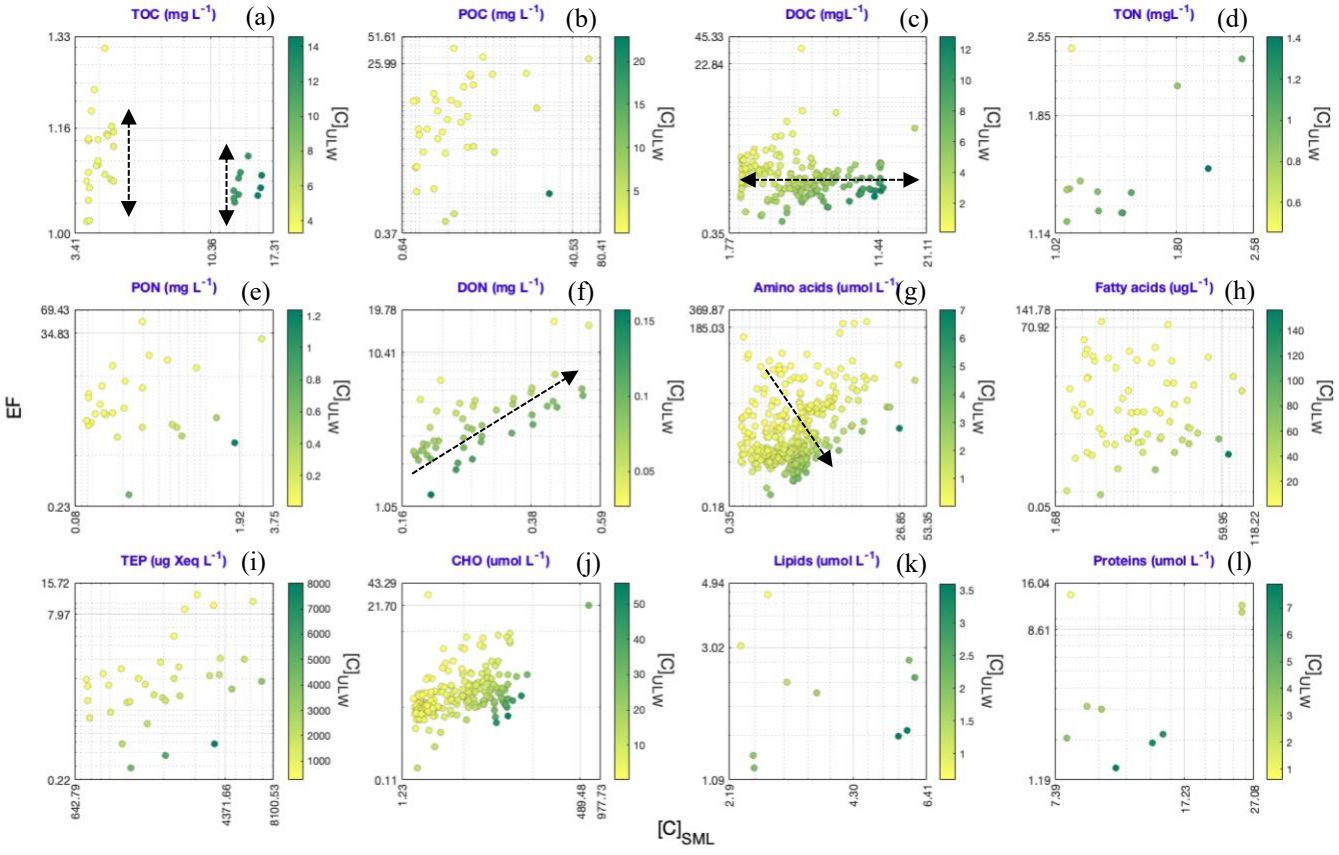

Figure 7: Interdependent relationship of $[C]_{SML}$ values with the corresponding $[C]_{ULW}$ and EF values. The analysis is restricted to $[C]_{SML}$ values that exceed the corresponding $\tilde{x}$ values. The x-axes give the observed $[C]_{SML}$ values against their corresponding EF values on y-axes. Datapoints are color-coded based on corresponding $[C]_{ULW}$ values. The black arrows indicate identified enrichment patterns.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

4. Discussion

By employing a meta-analytical approach, our study presents the first comprehensive overview of the enrichment dynamics in the SML, based on existing literature. A quantitative assessment of the distribution of the reference studies (i.e., the studies on which the SMD is based) across key domains of SML research provides insights into the most frequently studied aspects (Supplementary Figure S1); research on the SML has increased from about 15 publications per year in the early 2000s to approximately 50 per year by 2016 (Engel *et al.*, 2017). However, our work highlights the potential understudied areas in SML research that call for more in-depth analysis. For instance, majority of the reference studies have been conducted in oceanic and coastal regions (~76% of data) and predominantly during warmer months (~77% of data) with a significant mismatch observed for data collected under low and high wind regimes (~81% vs. 19%, respectively). In light of these research gaps, the following sections interpret the main findings revealed by our analysis and discuss their implications for understanding SML enrichment.

4.1 Overarching trends in SML enrichment

4.1.1 Generalized enrichment patterns

Comparison of KDE-derived PDFs for the EF values of (1) carbon-enriched vs. nitrogen-enriched organic compounds (Fig. 1) and (2) dissolved vs. particulate organic compounds (Fig. 2) yield the following key implications:

(1) All the estimated original (i.e., linear scale) PDFs (Figs. 1(a) and 2(a)) display higher probability densities for lower EF values and extended tails towards higher EF values (i.e., right-skewness), suggesting that under natural conditions, modest SML enrichment is far more common in general, while extreme enrichment events are rare.
(2) Nevertheless, variation in the peak probability densities among the PDFs indicate that extreme SML enrichment events are relatively more frequent in nitrogen-enriched compounds (Fig.1: orange PDF) and particulate forms (Fig. 2: green PDF), compared to carbon-enriched compounds (Fig.1: blue PDF) and dissolved forms (Fig.2: purple PDF)
(3) Nitrogen-enriched compounds and particulate forms exhibit a broader EF variability (i.e., higher mode, median, mean values) compared to carbon-enriched compounds and dissolved forms with a relatively more consistent spread (i.e., lower central tendency metrics)

These differences in peaks and central tendency metrics persist in log-transformed PDFs as well (Figs 1(b) and 2(b)). This validates that these variations are not caused by statistical artifacts but reflect real, natural variability in enrichment behavior. Overall, these findings from our meta-analysis indicate that the OM accumulation in the SML is more effective for (1) nitrogen-enriched than for carbon-enriched compounds and (2) particulate than for dissolved forms.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Our findings contradict some earlier works including Baastrup-Spohr and Staehr (2009), Liss and Duce (1997), Yang (1999) who suggest that the SML is similarly enriched for both particulate and dissolved organic (and inorganic) compounds, but align with other studies that report opposing results: Dietz *et al.* (1976) provide evidence for enhanced accumulations of particulate matter in the SML through particle aggregation at the surface. The work further links high abundances of living bacteria in the near-surface to higher availability of POC in the SML. Studies of Carlucci *et al.* (1985), Henrichs and Williams, (1985), Kuznetsova *et al.* (2004), Kuznetsova and Lee (2002) and Reinthaler *et al.* (2008) report that POC and PON tend to be more enriched in the SML than DOC. Engel *et al.* (2017) state that the SML has been shown to be enriched in particulate organic matter, particularly in proteinaceous compounds. These disparities, highlight potential uncertainties in employing carbon-based parameterization (instead of nitrogen-based) to estimate surfactant-suppressed CO_2 fluxes (e.g., Barthelmeß *et al.*, 2021; Li *et al.*, 2024). Overall, these overarching trends of SML enrichment underscore the importance of resolving compound-specific accumulation in the SML, while distinguishing between selective and non-selective enrichment. Cumulative probability comparison results for the $[C]_{ULW}$ and $[C]_{SML}$ (Figs. 4 and 5) and their corresponding linear correlations (Fig. 6) provide a meta-analytical perspective on how compounds are distributed and accumulated between the two compartments. Here, results concerning lipids and proteins are excluded due to apparent randomness in their distributions, potentially caused by smaller sample sizes.

4.1.2 Compound-specific enrichment patterns

Significant correlations between $[C]_{ULW}$ and $[C]_{SML}$ of nearly all the target compounds (ρ and $r > 0.5$) are consistent with the overall understanding that the SML's composition is linked to the availability of material in the underlying sub-surface waters (Chen *et al.*, 2016; Joux *et al.*, 2006; Kuznetsova *et al.*, 2004). Contrary to this general pattern, Kuznetsova *et al.* (2004) suggest that certain OM fractions in the SML and ULW may show lack of correlation, potentially due to constraints such as varying mineralization rates between the two layers and surface adsorption processes. Consistent with this view, linear correlation results for DON indicate such decoupling (Fig. 6(f)), though the underlying causes remain unexplored in this study. Early studies also suggested that the variations in the SML concentrations are typically larger than those in the ULW (Reinthaler *et al.*, 2008). In agreement, CDFs of the $[C]_{ULW}$ and $[C]_{SML}$ demonstrate faster probability accumulation for ULW than SML (Fig. 4), implying generally smaller magnitudes and lower variability in ULW concentrations compared to SML concentrations. Conversely, Carlson (1983) argues that, in certain occasions, OM variability in the SML and ULW may not significantly differ across temporal and spatial scales. The CDFs for TOC, DOC, TEP and CHO which exhibit the lowest IQD values (Fig. 5), support this but is contradicted by those of the other compounds, with higher IQD values (indicating substantial differences between the two concentrations).

Works of Hunter and Liss (1977) and Kurata *et al.* (2016) discuss the selective enrichment of surfactants in the SML, mainly driven by microbial processes. Hydrophobic compounds tend to show more affinity to the surface compared to hydrophilic substances (Marty and Saliot, 1976). In agreement, our linear correlation results reveal preferential accumulation of AA (Fig.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

6(g)) and FA (Fig. 6(h)), often parts of biosurfactants (e.g., lipopeptides or polypeptides), in the SML. Linear correlation results shown in Fig. 6(i) provide evidence to the view that TEP is generally enriched in the SML compared to the ULW (Cunliffe and Murrell, 2009; Cunliffe *et al.*, 2009; Wurl and Holmes, 2008), although this enrichment is not strongly pronounced in our dataset. Additionally, the nearly overlapping CDFs for TEP in SML and ULW (Fig. 4(i)) along with its low IQD value (= 0.081; Fig. 5) further suggest a surprisingly weak enrichment. However, concentration trend of TEP observed in our data closely aligns with that of CHO (Figs. 4(j) and 5), supporting the prevailing hypothesis that TEP is formed through coagulation of dissolved polysaccharides (Passow, 2000). Thornton *et al.* (2016) observe that TEP and dissolved polysaccharides do not always exhibit significant enrichment in the SML as anticipated.

POC and PON correlation patterns (Figs. 6(b) and 6(e), respectively) where $[C]_{SML}$ significantly exceed $[C]_{ULW}$, and that of DOC (Fig. 6(c)) where $[C]_{SML}$ is nearly equal to $[C]_{ULW}$, provide strong meta-analytical evidence to earlier works that discuss the selective enrichment of POC and PON in the SML over DOC (e.g., Carlucci *et al.*, 1992; Henrichs and Williams, 1985; Kuznetsova and Lee, 2002; Kuznetsova *et al.*, 2004 and Reinthaler *et al.*, 2008). Carlson (1983) suggests that the distribution of some organic fractions between the SML and the ULW may be governed by specific partitioning processes. For instance, while Chen *et al.* (2016) point out the significant role of the ULW in DOC and CHO accumulation in the SML, Dietz *et al.* (1976) observe fairly consistent abundances for these compounds between the two layers. Our experiments also show strong 1:1 correlation for DOC (Fig. 6(c)) and CHO (Fig. 6(j)), suggesting an absence of preferential affinity towards the SML (unlike surfactants), which further indicates that their enrichment is predominantly controlled by the ULW. Although CHO, AA and FA are identified to be the key constituents of the organic carbon pool (Hedges *et al.*, 1994), our correlation results reveal that their partitioning between the SML and the ULW and, their eventual enrichment patterns, may not be consistent (Figure 6).

4.1.3 Influencing factors and current uncertainties

Baier *et al.* (1974), Hunter and Liss (1981) and MacIntyre (1974) argue that the compositional diversity of the SML prevents single compounds from fully representing the dissolved OM class, which further emphasizes the importance of assessing compound-specific accumulation in the SML. Such investigations could shed light on selective and non-selective enrichment dynamics of OM. An analysis of EF-based PDFs for various AA fractions (Figure 8(a)) – Total AA (TAA), Dissolved Free AA (DFAA), Dissolved Combined AA (DCAA) and Particulate AA (PAA) – revealed notable heterogeneity within this compound class, reflecting the chemical diversity and complexity of OM enrichment in the SML. Additionally, consistent with previous studies that investigated the influence on environmental drivers on the enrichment dynamics in the SML (e.g., Asher, 1997; Barthelmeß *et al.*, 2021; Knulst *et al.*, 1997; Kuznetsova *et al.*, 2004; Liu and Dickhut, 1998; Obernosterer *et al.*, 2008; Reinthaler *et al.*, 2008; Tsai and Liu, 2003), our analysis demonstrates that factors such as sampling location (for DOC), sampling season (for CHO) and sampling method (for DOC) (Figs. 8(b) – (d)) play key roles in modulating the enrichment variability of the OM (it should be noted that these specific target compounds are chosen as representative examples because they span all subcategories of secondary data considered in the study (see Supplementary Table S1), and therefore enable a

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

more robust comparison among different settings). Moreover, bimodal CDFs of TOC for the SML and ULW concentrations (Fig. 4(a)) along with the distinct separation of three data clusters in their correlation patterns (Fig. 6(a)), further highlight the significant role of external environmental factors in shaping SML composition. The origins of TOC data used in this study illustrates this variability: Data from (1) a heavily polluted urban lake (concentration range: 12 – 16 mg L^{-1}; Baastrup-Spohr and Staehr, 2009), (2) a forested lake (concentration range: 3 – 5 mg L^{-1}; Baastrup-Spohr and Staehr, 2009), (3) the Arctic Ocean (concentration range: 1 – 3 mg L^{-1}; Gao *et al.*, 2012) and (4) an upwelling filament (concentration range: 3 – 4 mg L^{-1}; Barthelmeß *et al.*, 2021).

The influence on wind speed on SML enrichment remains ambiguous; our comparison of EF values under calm (< 6.6 ms^{-1}; Reinthaler *et al.*, 2008) and rough (> 6.6 ms^{-1}) wind conditions yield inconclusive results (Supplementary Figure S2) with wind speed appearing to have little/no effect on the SML enrichment (e.g., Basstrup-Spohr and Staehr, 2009 and Sabbaghzadeh *et al.*, 2017) or with enrichment persisting even under rough sea conditions (e.g., Kuznetsova *et al.*, 2004; Reinthaler *et al.*, 2008), opposing the general understanding that turbulent conditions may reduce the concentration in the SML (e.g., Carlson, 1983). Nevertheless, it is important to note that imbalanced sampling efforts among these categories, (Supplementary Figure S1), specially with regards to wind speed (Supplementary Figure S2; see the sample sizes), may compromise the robustness and validity of these findings.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

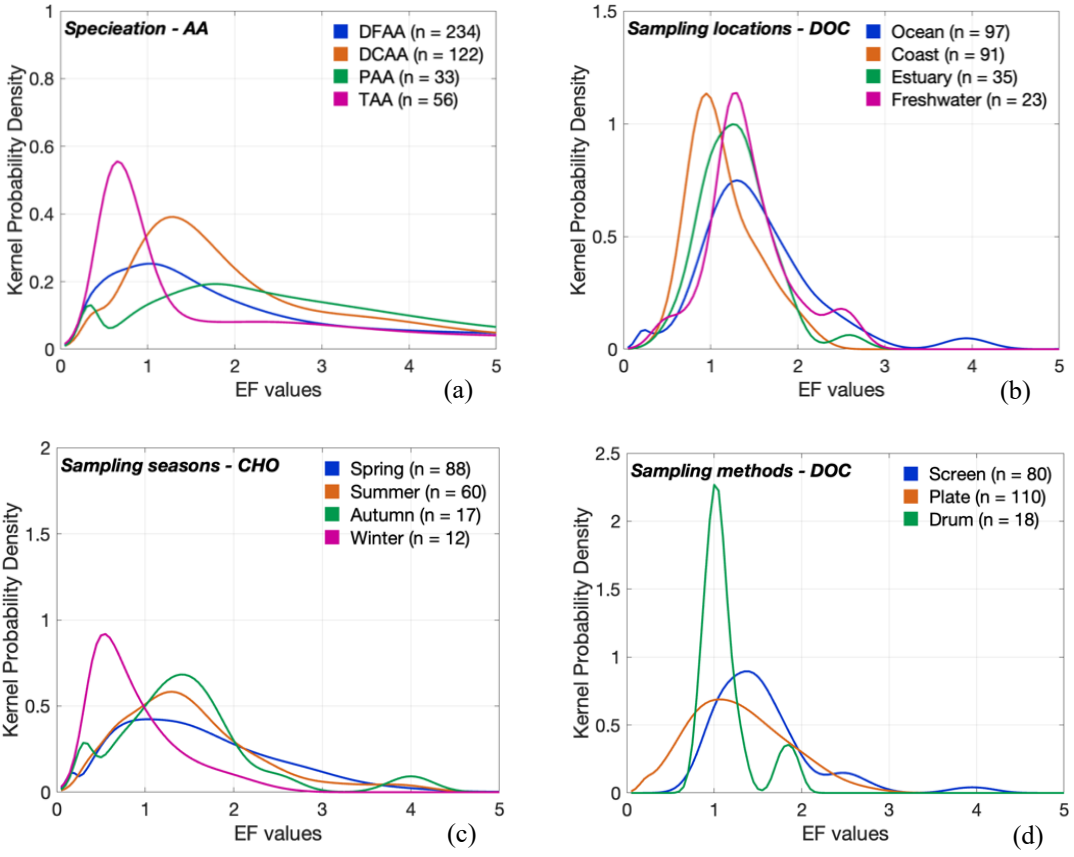

Figure 8: Factor-specific enrichment variability in the SML. PDFs illustrating varying enrichment patterns for (a) AA across chemical forms, (b) DOC across sampling locations, (c) CHO across sampling seasons and (d) DOC across sampling methods. 'n' gives the sample size of each category. Supplementary Table S1 summarizes different sampling locations, sampling seasons and sampling methods observed for the investigated target compounds.

Another major source of uncertainty arises from the variability in sampling depths of the ULW (Supplementary Table S1), which can affect the comparability of different data across multiple studies that would eventually introduce bias into the interpretation of overarching trends. Additional biases which are beyond the scope of this study include the potential influence of diurnal cycles (López-Puertas *et al.*, 2025); OM can be rapidly removed from the SML through photochemical degradation (Obernosterer *et al.*, 2008) and also be affected by reduced bacterial metabolism due to solar radiation (Dietz *et al.*, 1976). Therefore, taken together, our meta-data analysis suggests that, investigating SML enrichment without accounting for these influencing factors may mask true enrichment patterns, limiting the ability to derive meaningful insights. In light of these considerations, our work highlights the need for conducting species-specific and condition-dependent analyses in future SML research that also focus on subsequent environmental parameters, as also proposed by Pereira e*t al.* (2018).

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

4.2 Scale-related biases in EF estimates

Accurate data interpretation is essential to gain precise insights and arrive at substantiated conclusions (Isles, 2020; Menge *et al.*, 2018). This is particularly true for meta-analyses involving continuous environmental data where values may vary by several orders of magnitude (e.g., Vitousek, 2004). In our study, when the PDF_C vs. PDF_N (Fig. 1) and PDF_D vs. PDF_P (Fig. 2) are evaluated on a linear scale (panels (a)), they exhibit right-skewness, whereas their log-transformed versions approximate normal distributions (panels (b)). Comparisons between highly skewed distributions raise uncertainties as their offsets are often dominated by extreme values/outliers. In contrast, when log transformation is applied, the distributions tend to exhibit more symmetric, normalized patterns which enable direct comparisons in shape and spread across different categories (Zuur *et al.*, 2007). Therefore the normality assumption for EF is inappropriate and the computation of an arithmetic mean, a conventional practice adopted in many earlier works (e.g., Gašparović *et al.*, 2007; Gao *et al.*, 2012; Kuznetsova *et al.*, 2005; Williams *et al.*, 1986; Wurl and Homes, 2008; Wurl *et al.*, 2009), can be misleading, likely providing a biased general picture of OM enrichment in the SML.

The here constructed PDFs given in Figures 1 and 2 reveal that both mode (x_m; shown by solid straight lines) and arithmetic mean ($\overline{x_a}$; shown by dashed straight lines) differ between the two scales: The mode reflects the peak of a distribution and is sensitive to the shape of its respective density curve. It varies depending on whether a dataset is in 'skewed' linear space or 'normalized' log space and becomes ambiguous in polymodal distributions (regardless of the scale: e.g. Fig. 3). As a consequence, the mode in general provides an unreliable measure of central tendency. While the linear-arithmetic mean, which is influenced by outliers, result in biases that exaggerate the corresponding central tendency, the log-arithmetic mean prevents the extreme values from being dominant through balanced averaging and hence provides a reliable estimation of central tendency. Nevertheless, geometric mean in linear space ($\overline{x_g}$; straight lines with alternating dots and dashes) is a meaningful measure given that it is equivalent to the exponential of the arithmetic mean in logarithmic space (See Eqs. (B1) and (B2)). Median ($\tilde{x}$; dotted straight lines), on the other hand, remains relatively consistent across both scales as it is a rank-based measure of central tendency that is unaffected by the magnitude of outliers. Accordingly, we suggest that future SML enrichment studies employ a logarithmic scale for data analyses, and adopt either geometric mean and/or median on linear scale or arithmetic mean and/or median on logarithmic scale for reliable trend analysis.

Based on these new insights on scale transformations and central tendency metric considerations, we have redefined the typical EF values of the studied target compounds and their degrees of spread from a meta-analytical perspective, from the estimated $\tilde{x}$, $\overline{x_a}$ and thresholds (i.e. UT and LT) of their PDFs (Fig. 3). To re-establish these EF ranges as generally observed estimates under common conditions, 'the box plot method' (introduced by Tukey, 1977) was applied to the data to detect and remove potentially extreme EF values that rarely occur in nature. This method considers the values at 25th and 75th percentiles (also known as 'Q_1' and 'Q_3', respectively) of each distribution and computes a statistical estimate – Interquartile Range (also known

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

as 'IQR) – by the difference between Q_3 and Q_1 (i.e., $IQR = Q_3 - Q_1$). If a datapoint falls outsides the range of $Q_3 + IQR \times 1.5$ (also known as 'upper bound') and $Q_1 - IQR \times 1.5$ (also known as 'lower bound'), it is considered atypical and therefore excluded from the analysis. This outlier detection approach is widely used across a broad range of research (e.g., Barbato *et al.*, 2011; Carling, 2000; Hoaglin *et al.*, 1986).

4.3 Role of EF in reflecting SML enrichment

While the metric of EF offers a convenient way to assess the accumulation trends in the SML and therefore serves as the basis for many established insights and inferences in existing SML research (see Introduction), its ability to accurately and robustly express the 'true' enrichment nature of the SML has constantly been a question of interest (e.g., Basstrup-Spohr and Staehr, 2009; Hillbricht-Ilkowska & Kostrzewska-Szlakowska, 2004; Knulst *et al.*, 1997; Kuznetsova *et al.*, 2004; Liss and Duce, 1997; Münster *et al.*, 1998; Södergren, 1987). The EF is a ratio that expresses the 'relative' changes in $[C]_{SML}$ with respect to $[C]_{ULW}$ (Eq. (1)), and hence is sensitive to the variations in either layer. Ideally, to effectively reflect conditions of growing SML enrichment, EF values should gradually rise in response to increasing $[C]_{SML}$ and decreasing $[C]_{ULW,}$ which can be visibly observed for DON (Fig. 7(f)), AA (Fig. 7(g)) and CHO (Fig. 7(j)). Nevertheless, our meta-analysis highlights several inconsistencies that challenge the relevance of the EF values as indicators of 'true' SML enrichment. For instance, on the one hand similar EF values can be observed for both oligotrophic and eutrophic environments (referring to the EFs of TOC: Fig. 7(a)), which limits the ability to distinguish the differences in their trophic status, despite them being conspicuous in TOC's absolute concentration range (bimodal CDFs; Fig. 4(a)). On the other hand, high (low) EF values may occur under oligotrophic (eutrophic) conditions leading to over- (under-) estimation of ecological setting (Fig. 7(g)). Furthermore, symmetrical changes in SML and ULW yield near-constant EF values across a wide range of concentrations (Fig. 7(c)), which could cause misinterpretations in key ecosystem shifts. We have also observed consistent EF values, even when SML and ULW concentrations vary over several orders of magnitudes (Figs. 6(g) – (j)), which further raise concerns over the metric's robustness. Therefore, although widely used, EF values should be interpreted with caution and, combined with additional parameters that provide more accurate information about the true enrichment behaviour of the SML.

A complementary parameter would be the typical upper limit of a $[C]_{SML}$ distribution which may reflect the maximum concentration capacity of the SML. Such a measure can serve as a robust concentration estimate of such maximum capacity if approximated from a meta-data derived distribution that includes observations across all diverse environmental conditions. Table 1 summarizes the upper $[C]_{SML}$ threshold estimates (i.e., UT; at 95th percentile) for the target compounds, based on their CDFs (Fig. 4). Although the robustness of these values largely depends on the quality and the scope of the underlying metadata, our bootstrapping approach addresses these potential limitations. Nevertheless, we acknowledge that these estimates remain data-constrained and therefore can improve with the inclusion of more comprehensive, high-resolution datasets across diverse environmental conditions. Measured concentrations beyond these thresholds within the SML must be regarded as exceptionally

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

high and should therefore find special attention, in particular with regard to the associated biogeochemical, oceanographic and weather conditions.

Table 1: A summary of estimated UT values (upper threshold; concentration at 95th percentile) for [C]$_{SML}$ distributions of the target compounds. This metric represents the maximum accumulation capacity of a certain compound in the SML**.** The values are rounded to the nearest whole number.

Compound	TOC mg L^{-1}	POC mg L^{-1}	DOC mg L^{-1}	TON mg L^{-1}	PON mg L^{-1}	DON mg L^{-1}	AA µmol L^{-1}	FA µg L^{-1}	TEP µg Xeq L^{-1}	CHO µmol L^{-1}	Lipids µmol L^{-1}	Proteins µmol L^{-1}
UT value	14.0	11.0	10.0	2.0	1.0	0.5	6.0	46.0	6166.0	56.0	8.0	30.0

Considering absolute changes in the SML concentrations (rather than relative changes) – calculated as the magnitude difference between corresponding SML and ULW concentrations (i.e., $[C]_{SML} - [C]_{ULW}$) – provides complementary insights into the SML's enrichment dynamics. When this metric is compared against the EF values for DOC data where EF > $\tilde{x}$ (= 1.2), resulting Spearman's correlation coefficients (ρ) reveal a stronger relationship (Figure 9(c) ; ρ = 0.32) relative to EF vs. [C]$_{SML}$ (Figure 9(a) ; ρ = 0.08) and EF vs. [C]$_{ULW}$ (Figure 9(c) ; ρ = -0.12) correlations. This implies that although the EFs may have a limited capacity to represent the absolute concentrations of either SML or ULW, they are more responsive to the absolute concentration 'changes' in the two compartments. This analysis reveals that although 'enrichment factor' obscures accurately interpreting the trophic status or the actual enrichment in the SML, it may still hold value as a proxy that reflects the degree of partitioning between the surface microlayer and underlying waters.

Furthermore, normalization of $[C]_{SML} - [C]_{ULW}$ metric to corresponding [C]$_{SML}$ values (i.e., $([C]_{SML} - [C]_{ULW})/[C]_{SML}$) ultimately yields an EF-based metric: $1 - \frac{1}{EF}$. This expresses how much of the SML concentration is above the ULW baseline. Unlike conventional EF values, $1 - \frac{1}{EF}$ only ranges between 0 and 1. It rescales compound-specific variability in EF and is therefore better suited for comparison across all the different observational types; normalization of EF onto a common scale allows direct evaluations without bias from different units, magnitudes or concentration ranges. This better captures true trends, rather than artifacts/effects of scale, while enhancing visualization and communication of results. In addition, when compared against the absolute changes, this metric exhibits stronger correlation ((ρ = 0.52; Fig. 9(d)), likely due to increased robustness to concentration variability obtained through its scale-dependent nature. As a result, when incorporated into modelling efforts, the normalized EF metric can offer distinct advantages such as integration of heterogenous datasets, consistent parameterization, easier comparisons of model predictions and robust sensitivity analyses. Together, these benefits contribute to more reliable and generalizable models of SML processes.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

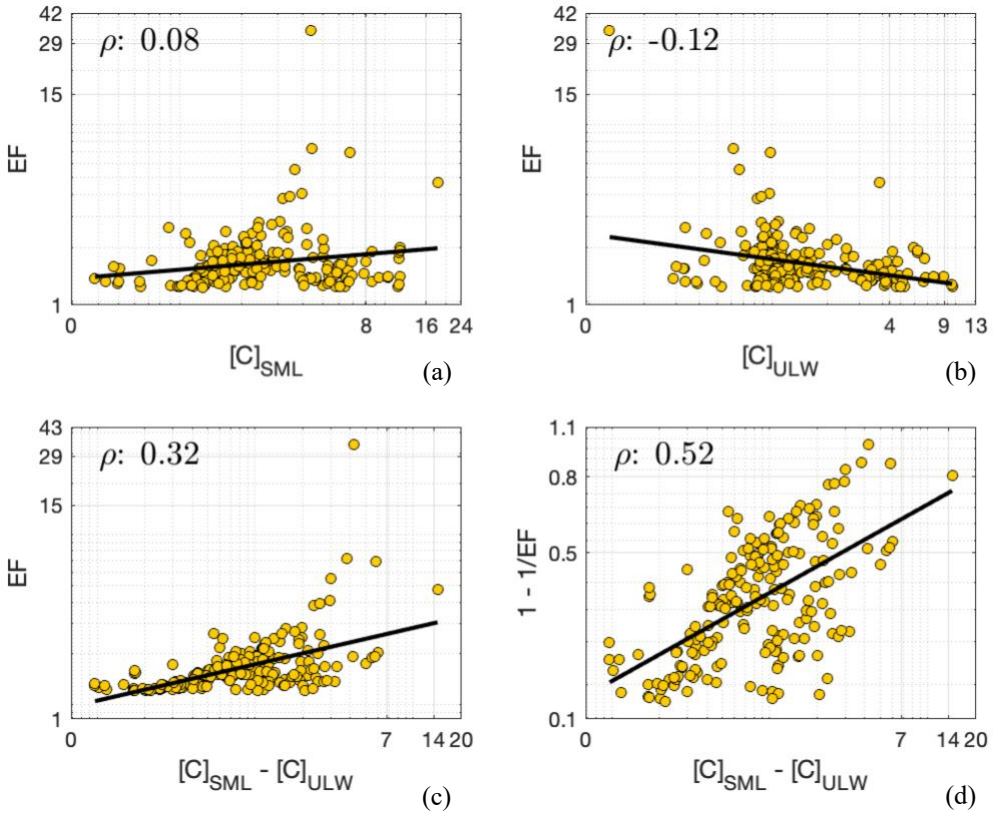

Figure 9: Correlations between EF and (a) $[C]_{SML}$, (b) $[C]_{ULW}$, (c) $[C]_{SML} - [C]_{ULW}$ and, (d) correlation between 1 – 1/EF and $[C]_{SML} - [C]_{ULW}$. These plots were generated for DOC data, in order to investigate the observed lack of correlation among EF – $[C]_{SML}$ – $[C]_{ULW}$, as shown by Fig. 7(c).

5. Conclusion

The here presented extensive data collection of $[C]_{SML}$ and $[C]_{ULW}$ measurements of various observational types of organic matter compounds is unique. Although the physical and biogeochemical properties of the SML have been studied in detail across diverse disciplines of aquatic science, to our knowledge, this study is the first to adopt a meta-analytical approach to bridge between insights of individual findings documented in the literature. Our quantitative assessment on the distributional characteristics of 12 organic compounds in the SML yielded statistically robust results owing to the use of the KDE method as the primary analytical technique, which enabled a coherent comparison of multivariate data.

Our results indicate that nitrogen-rich compounds and particulate OM tend to be more enriched in the SML compared to carbon-rich compounds and dissolved OM. These findings underscore the need for future SML research to focus on the species-specific chemistry of OM, specifically surfactants, along with the variability of external influencing factors, in order to better

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

understand and approximate their role in global gas flux estimates. Informed by this, we explore EF values by re-evaluating the typical ranges, previously defined in individual studies, through a meta-analytical perspective. Our assessment also inquired into the suitability of EF values as indicators of true SML enrichment and suggests that, while the EF, which expresses relative changes in the SML compared to the ULW, can reflect the partition variability between the two layers, it falls short in capturing trophic variability. The latter plays a crucial role in determining ecosystem structure. In light of the foregoing, we propose that EF estimates be complemented with additional parameters such as the absolute concentration differences between the SML and the ULW and the maximum concentration capacities within the SML, to provide a more in-depth understanding of enrichment dynamics.

In addition to the primary outcomes, a noteworthy secondary insight from this study is the importance of selecting an appropriate data transformation scale (i.e. linear or logarithmic) and a robust measure of central tendency (i.e. mode, median, arithmetic mean or geometric mean) to ensure accurate data representation and reliable inference in future SML research. Our analysis provides strong evidence for the following advantages of transforming data into logarithmic scale: (1) it facilitates meaningful ratio comparisons such as the EF by converting multiplicative relationships into additive ones; (2) it better reflects the log-normality characteristics of the dependencies between $[C]_{SML}$ and $[C]_{ULW}$, which improves statistical model performance; (3) it was shown to promote homoscedasticity and (4) it enhances robustness, accuracy and interpretability of central tendency metrics.

Appendix A: KDE method – additional information

Although the most basic non-parametric method to derive a probability distribution is histograms, they present two key limitations for comparative studies: (1) unequal sample sizes across comparative groups restrict the use of uniform binning and, (2) imposing uniform bin sizes potentially mask important distributional characteristics. In contrast, KDE circumvents these issues by accounting a datapoint's exact value rather than assigning it to a particular bin of a certain width. This describes the true underlying distribution of the data and allows more consistent and detailed comparisons of distributions. In this analysis, we use Gaussian kernels – smooth, bell-shaped functions based on normal distribution – that weight observations based on their distance. Chen (2017) and the references therein provide a comprehensive review of the KDE and its recent advances.

In Gaussian kernels, bandwidth is analogous to standard deviation. In this study, the bandwidths for the linear KDEs were computed based on an approach that includes a bias-variance trade-off. Briefly, the bias-variance trade-off represents kernels that have a bandwidth that avoids too much variance in the estimates (i.e., bandwidths are not too small) while it does not introduce too much bias for ranges that actually exhibit no data points (i.e., bandwidths are not too large). Calculations of optimal bandwidth applied herein and an example of a bias-variance trade-off are described in Schartau *et al.*, (2010).

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Nevertheless, in log-space, unlike in linear-space, data are more evenly distributed and hence fixed bandwidths avoid over-smoothing of low values and under-smoothing of high values.

The selection of an optimal bandwidth for KDEs is influenced by sample size; smaller sample sizes lead to sparse and noisy distributions which require more smoothing and therefore larger bandwidths. Excessively large bandwidths can result in underfitting. In contrast, larger sample sizes may allow excessively smaller bandwidths that can lead to overfitting. Bootstrap resampling addresses these potential uncertainties in our analysis and, ensures the robustness and precision of the estimated density distributions. Deviations between the bootstrapped KDEs and their ensemble mean were found to approximate a normal distribution (consistent with Central Limit Theorem). Therefore, these ensemble means can be regarded as reliable representations of the underlying data, supporting valid comparisons of probability distributions across different groups or clusters.

Appendix B: Mathematical expressions of distributional characteristics

If a dataset contains values of 'x_i' with a sample size of 'n', mode (x_m) is the most frequently occurring value in the dataset and therefore, the point where a PDF reaches its highest density. A distribution appears to be the most concentrated at x_m. Median ($\tilde{x}$) returns the value at the 50th percentile of an ascending dataset. It divides the area under a PDF into two equal halves. The arithmetic mean ($\overline{x_a}$), is the average of a distribution, given by the following equation:

$$\Sigma_{i=1}^{n} = \frac{x_i}{n} \quad \text{(B1)}$$

$\overline{x_a}$ gives the point where weighted sum of a PDF is balanced. However, in the case of datasets that range over several orders of magnitudes, the geometric mean (hereafter referred to as '$\overline{x_g}$') is the more preferred central tendency estimate, as it accounts for the relative proportions of values (as opposed to their absolute magnitudes as is the case in $\overline{x_a}$) and hence, is less sensitive to outliers. $\overline{x_g}$ is calculated by the following equation:

$$(\Pi_{i=1}^{n} x_i)^{\frac{1}{n}} \quad \text{(B2)}$$

$\overline{x_g}$ of a linear distribution is mathematically equal to the exponentiated $\overline{x_a}$ of the log-transformed version of the same distribution.

In addition, the following equation, which accounts for the squared differences across all the datapoints of the corresponding CDFs, estimates the discrete form of the integrated quadratic distance (*IQD*, explained in section 2.2.2), with $\Delta x_i = x_i - x_{i-1}$:

$$IQD = \sum_{i=1}^{n} \left(\left(CDF_{[C]_{SML}(x_i)} - CDF_{[C]_{ULW}(x_i)} \right)^2 \times \Delta x_i \right) \quad \text{(B3)}$$

A higher IQD value implies that the divergence is greater and therefore the corresponding CDFs are more different.

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Code availability

The computational codes used in this study are available upon request to A.S.

Data availability

All data used in this study were extracted from previously published peer-reviewed sources. Full citations for all the datasets are provided in supplementary information. No new data were generated for this study.

Author contribution

A. S. – Data Curation, Conceptualization, Methodology, Formal Analysis, Visualization, Writing – Original Draft, Writing – Review & Editing
S. N. – Data Curation, Writing – Review & Editing
T. B. – Data Provision, Writing – Review & Editing
A. E. – Funding Acquisition, Data Provision, Writing – Review & Editing
H. H. – Data Provision, Writing – Review & Editing
M. P. – Data Provision, Writing – Review & Editing
K. W. – Methodology, Writing – Review & Editing
O. W. – Funding Acquisition, Data Provision, Writing – Review & Editing
M. S. – Conceptualization, Methodology, Formal Analysis, Visualization, Funding Acquisition, Supervision, Writing – Review & Editing

Competing interests

Authors A. S., T. B., A. E. and M. S. are affiliated with the same institution as H. B., who serves as an overseeing editor for the special issue *"Biogeochemical processes and air-sea exchange in the sea-surface microlayer"*. Authors A. S., T. B., A. E., H. H., M. P., O. W. and M. S. are collaborators with H. B. on an ongoing research project. These potential competing interests have been fully disclosed to the journal. The authors declare no other competing interests relevant to the submitted work.

Acknowledgements

The authors gratefully acknowledge the support and collaboration of colleagues and institutions involved in this work. We also acknowledge the valuable contributions of the authors of previously published datasets that were used in this study. Their work

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

provided essential input for our analysis. This study was conducted as part of the project "Biogeochemical processes and air-sea exchange in the sea-surface microlayer (BASS)", and we thank all project partners for their support and scientific exchange. Data compilation had been initiated through the large integrated project Surface Ocean Processes in the Anthropocene (SOPRAN, 03F0662A), funded by the German Federal Ministry of Education and Research (BMBF).

Financial support

This work is supported by the project "Biogeochemical processes and air-sea exchange in the sea-surface microlayer (BASS)", which is funded by the German Research Foundation (DFG) under Grant No 451574234.

References

Alldredge, A.L., Passow, U., Logan, B.E., 1993. The abundance and significance of a class of large, transparent organic particles in the ocean. Deep Sea Res. Part Oceanogr. Res. Pap. 40, 1131–1140. https://doi.org/10.1016/0967-0637(93)90129-Q

Asher, W., 1997. The sea-surface microlayer and its effect on global air-sea gas transfer. Sea Surf. Glob. Change 251–286.

Astrahan, P., Herut, B., Paytan, A., Rahav, E., 2016. The Impact of Dry Atmospheric Deposition on the Sea-Surface Microlayer in the SE Mediterranean Sea: An Experimental Approach. Front. Mar. Sci. 3. https://doi.org/10.3389/fmars.2016.00222

Baastrup-Spohr, L., Staehr, P.A., 2009. Surface microlayers on temperate lowland lakes. Hydrobiologia 625, 43–59. https://doi.org/10.1007/s10750-008-9695-3

Baier, R., DW, G., Perlmutter, S., King, R., 1974. Dominant chemical composition of sea-surface films, natural slicks, and foams.

Barbato, G., Barini ,E. M., Genta ,G., and Levi, R., 2011. Features and performance of some outlier detection methods. J. Appl. Stat. 38, 2133–2149. https://doi.org/10.1080/02664763.2010.545119

Barthelmeß, T., Schütte, F., Engel, A., 2021. Variability of the Sea Surface Microlayer Across a Filament's Edge and Potential Influences on Gas Exchange. Front. Mar. Sci. 8, 718384. https://doi.org/10.3389/fmars.2021.718384

Brinis, A., Méjanelle, L., Momzikoff, A., Gondry, G., Fillaux, J., Point, V., Saliot, A., 2004. Phospholipid ester-linked fatty acids composition of size-fractionated particles at the top ocean surface. Org. Geochem. 35, 1275–1287. https://doi.org/10.1016/j.orggeochem.2004.04.009

Carling, K., 2000. Resistant outlier rules and the non-Gaussian case. Comput. Stat. Data Anal. 33, 249–258. https://doi.org/10.1016/S0167-9473(99)00057-2

Carlson, D.J., 1983. Dissolved organic materials in surface microlayers: Temporal and spatial variability and relation to sea state. Limnol. Oceanogr. 28, 415–431. https://doi.org/10.4319/lo.1983.28.3.0415

Carlucci, A.F., Craven, D.B., Henrichs, S.M., 1985. Surface-film microheterotrophs: amino acid metabolism and solar radiation effects on their activities.

Cen-Lin, H., Tzung-May, F., 2013. Air-Sea Exchange of Volatile Organic Compounds: A New Model with Microlayer Effects. Atmospheric Ocean. Sci. Lett. 6, 97–102. https://doi.org/10.1080/16742834.2013.11447063

Chen, Y., Yang, G.-P., Xia, Q.-Y., Wu, G.-W., 2016. Enrichment and characterization of dissolved organic matter in the surface microlayer and subsurface water of the South Yellow Sea. Mar. Chem. 182, 1–13. https://doi.org/10.1016/j.marchem.2016.04.001

Chen, Y.-C., 2017. A tutorial on kernel density estimation and recent advances. Biostat. Epidemiol. 1, 161–187.

Crocetti, E., 2016. Systematic Reviews With Meta-Analysis: Why, When, and How? Emerg. Adulthood 4, 3–18. https://doi.org/10.1177/2167696815617076

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Cunliffe, M., Engel, A., Frka, S., Gašparović, B., Guitart, C., Murrell, J.C., Salter, M., Stolle, C., Upstill-Goddard, R., Wurl, O., 2013. Sea surface microlayers: A unified physicochemical and biological perspective of the air–ocean interface. Prog. Oceanogr. 109, 104–116. https://doi.org/10.1016/j.pocean.2012.08.004

Cunliffe, M., Murrell, J.C., 2009. The sea-surface microlayer is a gelatinous biofilm. ISME J. 3, 1001–1003. https://doi.org/10.1038/ismej.2009.69

Cunliffe, M., Salter, M., Mann, P.J., Whiteley, A.S., Upstill-Goddard, R.C., Murrell, J.C., 2009. Dissolved organic carbon and bacterial populations in the gelatinous surface microlayer of a Norwegian fjord mesocosm. FEMS Microbiol. Lett. 299, 248–254. https://doi.org/10.1111/j.1574-6968.2009.01751.x

Dietz, A.S., Albright, L.J., Tuominen, T., 1976. Heterotrophic activities of bacterioneuston and bacterioplankton. Can. J. Microbiol. 22, 1699–1709. https://doi.org/10.1139/m76-251

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., Zäncker, B., 2017. The Ocean's Vital Skin: Toward an Integrated Understanding of the Sea Surface Microlayer. Front. Mar. Sci. 4, 165. https://doi.org/10.3389/fmars.2017.00165

Engel, A., Galgani, L., 2016. The organic sea-surface microlayer in the upwelling region off the coast of Peru and potential implications for air–sea exchange processes. Biogeosciences 13, 989–1007. https://doi.org/10.5194/bg-13-989-2016

Engel, A., Thoms, S., Riebesell, U., Rochelle-Newall, E., Zondervan, I., 2004. Polysaccharide aggregation as a potential sink of marine dissolved organic carbon. Nature 428, 929–932. https://doi.org/10.1038/nature02453

Feenstra, S., 2006. Use of logarithmic-scale correlation plots to represent contaminant ratios for evaluation of subsurface environmental data. Environ. Forensics 7, 175–185.

Frew, N.M., 1997. The role of organic films in air–sea gas exchange, in: Liss, P.S., Duce, R.A. (Eds.), The Sea Surface and Global Change. Cambridge University Press, Cambridge, pp. 121–172. https://doi.org/10.1017/CBO9780511525025.006

Frew, N.M., Bock, E.J., Schimpf, U., Hara, T., Haußecker, H., Edson, J.B., McGillis, W.R., Nelson, R.K., McKenna, S.P., Uz, B.M., Jähne, B., 2004. Air-sea gas transfer: Its dependence on wind stress, small-scale roughness, and surface films. J. Geophys. Res. Oceans 109, 2003JC002131. https://doi.org/10.1029/2003JC002131

Gao, Q., Leck, C., Rauschenberg, C., Matrai, P.A., 2012. On the chemical dynamics of extracellular polysaccharides in the high Arctic surface microlayer. Ocean Sci. 8, 401–418. https://doi.org/10.5194/os-8-401-2012

Garabetian, F., Romano, J.-C., Paul, R., Sigoillot, J.-C., 1993. Organic matter composition and pollutant enrichment of sea surface microlayer inside and outside slicks. Mar. Environ. Res. 35, 323–339. https://doi.org/10.1016/0141-1136(93)90100-E

Gašparović, B., Plavšić, M., Ćosović, B., Saliot, A., 2007. Organic matter characterization in the sea surface microlayers in the subarctic Norwegian fjords region. Mar. Chem. 105, 1–14. https://doi.org/10.1016/j.marchem.2006.12.010

Goldman, J.C., Dennett, M.R., Frew, N.M., 1988. Surfactant effects on air-sea gas exchange under turbulent conditions. Deep Sea Res. Part Oceanogr. Res. Pap. 35, 1953–1970. https://doi.org/10.1016/0198-0149(88)90119-7

Härdle, W., Müller, M., Sperlich, S., Werwatz, A., 2004. Nonparametric and semiparametric models. Springer Science & Business Media.

Hardy, J., 1997. Biological effects of chemicals in the sea-surface microlayer. Camb. Univ. PRESS CAMBRIDGEUK 339–370.

Hedges, J.I., Cowie, G.L., Richey, J.E., Quay, P.D., Benner, R., Strom, M., Forsberg, B.R., 1994. Origins and processing of organic matter in the Amazon River as indicated by carbohydrates and amino acids. Limnol. Oceanogr. 39, 743–761. https://doi.org/10.4319/lo.1994.39.4.0743

Henrichs, S.M., Williams, P.M., 1985. Dissolved and particulate amino acids and carbohydrates in the sea surface microlayer. Mar. Chem. 17, 141–163. https://doi.org/10.1016/0304-4203(85)90070-2

Hillbricht-Ilkowska, A., Kostrzewska-Szlakowska, I., 2004. Surface microlayer in lakes of different trophic status: nutrients concentration and accumulation. Pol. J. Ecol. 52, 461–478.

Hoaglin, D.C., Iglewicz ,Boris, and Tukey, J.W., 1986. Performance of Some Resistant Rules for Outlier Labeling. J. Am. Stat. Assoc. 81, 991–999. https://doi.org/10.1080/01621459.1986.10478363

Hunter, K.A., 1980. Processes affecting particulate trace metals in the sea surface microlayer. Mar. Chem. 9, 49–70. https://doi.org/10.1016/0304-4203(80)90006-7

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Hunter, K.A., Liss, P.S., 1981. Chapter 9 Organic Sea Surface Films, in: Duursma, E.K., Dawson, R. (Eds.), Elsevier Oceanography Series. Elsevier, pp. 259–298. https://doi.org/10.1016/S0422-9894(08)70331-3

Hunter, K.A., Liss, P.S., 1977. The input of organic material to the oceans: air—sea interactions and the organic chemical composition of the sea surface. Mar. Chem. 5, 361–379. https://doi.org/10.1016/0304-4203(77)90029-9

Joux, F., Agogué, H., Obernosterer, I., Dupuy, C., Reinthaler, T., Herndl, G., Lebaron, P., 2006. Microbial community structure in the sea surface microlayer at two contrasting coastal sites in the northwestern Mediterranean Sea. Aquat. Microb. Ecol. 42, 91–104. https://doi.org/10.3354/ame042091

Keene, O.N., 1995. The log transformation is special. Stat. Med. 14, 811–819. https://doi.org/10.1002/sim.4780140810

Knipping, E.M., Lakin, M.J., Foster, K.L., Jungwirth, P., Tobias, D.J., Gerber, R.B., Dabdub, D., Finlayson-Pitts, B.J., 2000. Experiments and Simulations of Ion-Enhanced Interfacial Chemistry on Aqueous NaCl Aerosols. Science 288, 301–306. https://doi.org/10.1126/science.288.5464.301

Knulst, J.C., Backlund, P., Hessen, D.O., Riise, G., Södergren, A., 1997. Response of surface microlayers to artificial acid precipitation in a meso-humic lake in Norway. Water Res. 31, 2177–2186. https://doi.org/10.1016/S0043-1354(97)00061-4

Kock, A., Schafstall, J., Dengler, M., Brandt, P., Bange, H.W., 2012. 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

Kuznetsova, M., Lee, C., 2002. Dissolved free and combined amino acids in nearshore seawater, sea surface microlayers and foams: Influence of extracellular hydrolysis. Aquat Sci 64, 252–268.

Kuznetsova, M., Lee, C., Aller, J., Frew, N., 2004. Enrichment of amino acids in the sea surface microlayer at coastal and open ocean sites in the North Atlantic Ocean. Limnol. Oceanogr. 49, 1605–1619. https://doi.org/10.4319/lo.2004.49.5.1605

Li, L., Wu, P., Zhang, P., Huang, S., Zhang, Y., 2024. An improved model for air–sea exchange of elemental mercury in MITgcm-ECCOv4-Hg: the role of surfactants and waves. Geosci. Model Dev. 17, 8683–8695. https://doi.org/10.5194/gmd-17-8683-2024

Liss, P.S., Duce, R.A., 1997. The sea surface and global change. Cambridge University Press.

Liu, K., Dickhut, R.M., 1998. Effects of wind speed and particulate matter source on surface microlayer characteristics and enrichment of organic matter in southern Chesapeake Bay. J. Geophys. Res. Atmospheres 103, 10571–10577. https://doi.org/10.1029/97JD03736

López-Puertas, A., Wurl, O., Frka, S., Ribas-Ribas, M., 2025. Diel Variability Affects the Inorganic Marine Carbon System in the Sea-Surface Microlayer of a Mediterranean coastal area (Šibenik, Croatia). https://doi.org/10.5194/egusphere-2025-2090

MacIntyre, F., 1974. Chemical fractionation and sea-surface microlayer processes. The sea 5, 245–299.

Maki, J.S., Hermansson, M., 2020. The dynamics of surface microlayers in aquatic environments. Biol. Part. Aquat. Syst. Second Ed. 161–182.

Mari, X., Burd, A., 1998. Seasonal size spectra of transparent exopolymeric particles (TEP) in a coastal sea and comparison with those predicted using coagulation theory. Mar. Ecol. Prog. Ser. 163, 63–76. https://doi.org/10.3354/meps163063

Marty, J.C., Saliot, A., 1976. Hydrocarbons (normal alkanes) in the surface microlayer of seawater. Deep-Sea Res. 23, 863–873.

McKenna, S.P., McGillis, W.R., 2004. The role of free-surface turbulence and surfactants in air–water gas transfer. Int. J. Heat Mass Transf. 47, 539–553. https://doi.org/10.1016/j.ijheatmasstransfer.2003.06.001

Menge, D.N.L., MacPherson, A.C., Bytnerowicz, T.A., Quebbeman, A.W., Schwartz, N.B., Taylor, B.N., Wolf, A.A., 2018. Logarithmic scales in ecological data presentation may cause misinterpretation. Nat. Ecol. Evol. 2, 1393–1402. https://doi.org/10.1038/s41559-018-0610-7

Mengist, W., Soromessa, T., Legese, G., 2020. Method for conducting systematic literature review and meta-analysis for environmental science research. MethodsX 7, 100777. https://doi.org/10.1016/j.mex.2019.100777

Milinković, A., Penezić, A., Kušan, A.C., Gluščić, V., Žužul, S., Skejić, S., Šantić, D., Godec, R., Pehnec, G., Omanović, D., Engel, A., Frka, S., 2022. Variabilities of biochemical properties of the sea surface microlayer: Insights to the atmospheric deposition impacts. Sci. Total Environ. 838, 156440. https://doi.org/10.1016/j.scitotenv.2022.156440

Münster, U., Heikkinen, E., Knulst, J., 1998. Nutrient composition, microbial biomass and activity at the air-water interface of small boreal forest lakes, in: Tamminen, T., Kuosa, H. (Eds.), Eutrophication in Planktonic Ecosystems: Food Web

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Dynamics and Elemental Cycling. Springer Netherlands, Dordrecht, pp. 261–270. https://doi.org/10.1007/978-94-017-1493-8_21

Mustaffa, N.I.H., Ribas-Ribas, M., Banko-Kubis, H.M., Wurl, O., 2020. Global reduction of *in situ* CO_2 transfer velocity by natural surfactants in the sea-surface microlayer. Proc. R. Soc. Math. Phys. Eng. Sci. 476, 20190763. https://doi.org/10.1098/rspa.2019.0763

Obernosterer, I., Catala, P., Lami, R., Caparros, J., Ras, J., Bricaud, A., Dupuy, C., van Wambeke, F., Lebaron, P., 2008. Biochemical characteristics and bacterial community structure of the sea surface microlayer in the South Pacific Ocean.

Obernosterer, I., Catala, P., Reinthaler, T., Herndl, G., Lebaron, P., 2005. Enhanced heterotrophic activity in the surface microlayer of the Mediterranean Sea. Aquat. Microb. Ecol. 39, 293–302. https://doi.org/10.3354/ame039293

Parzen, E., 1962. On Estimation of a Probability Density Function and Mode. Ann. Math. Stat. 33, 1065–1076.

Passow, U., 2000. Formation of transparent exopolymer particles, TEP, from dissolved precursor material. Mar. Ecol. Prog. Ser. 192, 1–11. https://doi.org/10.3354/meps192001

Passow, U., Alldredge, A.L., 1999. Do transparent exopolymer particles (TEP) inhibit grazing by the euphausiid Euphausia pacifica? J. Plankton Res. 21, 2203–2217. https://doi.org/10.1093/plankt/21.11.2203

Pereira, R., Ashton, I., Sabbaghzadeh, B., Shutler, J.D., Upstill-Goddard, R.C., 2018. Reduced air–sea CO2 exchange in the Atlantic Ocean due to biological surfactants. Nat. Geosci. 11, 492–496. https://doi.org/10.1038/s41561-018-0136-2

Petersen, M.K., Iyengar, S.S., Day, T.J.F., Voth, G.A., 2004. The Hydrated Proton at the Water Liquid/Vapor Interface. J. Phys. Chem. B 108, 14804–14806. https://doi.org/10.1021/jp046716o

Reinthaler, T., Sintes, E., Herndl, G.J., 2008. Dissolved organic matter and bacterial production and respiration in the sea-surface microlayer of the open Atlantic and the western Mediterranean Sea. Limnol. Oceanogr. 53, 122–136. https://doi.org/10.4319/lo.2008.53.1.0122

Sabbaghzadeh, B., Upstill-Goddard, R.C., Beale, R., Pereira, R., Nightingale, P.D., 2017. 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}. Geophys. Res. Lett. 44, 2852–2858. https://doi.org/10.1002/2017GL072988

Salter, M.E., Upstill-Goddard, R.C., Nightingale, P.D., Archer, S.D., Blomquist, B., Ho, D.T., Huebert, B., Schlosser, P., Yang, M., 2011. Impact of an artificial surfactant release on air-sea gas fluxes during Deep Ocean Gas Exchange Experiment II. J. Geophys. Res. Oceans 116, 2011JC007023. https://doi.org/10.1029/2011JC007023

Schartau, M., Engel, A., Schröter, J., Thoms, S., Völker, C., Wolf-Gladrow, D., 2007. Modelling carbon overconsumption and the formation of extracellular particulate organic carbon. Biogeosciences 4, 433–454. https://doi.org/10.5194/bg-4-433-2007

Schartau, M., Landry, M.R., Armstrong, R.A., 2010. Density estimation of plankton size spectra: a reanalysis of IronEx II data. J. Plankton Res. 32, 1167–1184. https://doi.org/10.1093/plankt/fbq072

Shine, J.P., Wallace, G.T., 1996. Flux of surface-active organic complexes of copper to the air-sea interface in coastal marine waters. J. Geophys. Res. Oceans 101, 12017–12026. https://doi.org/10.1029/96JC00616

Sieburth, J.McN., 1983. Microbiological and Organic-Chemical Processes in the Surface and Mixed Layers, in: Liss, P.S., Slinn, W.G.N. (Eds.), Air-Sea Exchange of Gases and Particles. Springer Netherlands, Dordrecht, pp. 121–172. https://doi.org/10.1007/978-94-009-7169-1_3

Silverman, B.W., 1986. Density estimation for statistics and data analysis. Chapman & Hall.

Södergren, A., 1987. Origin and composition of surface slicks in lakes of differing trophic status1. Limnol. Oceanogr. 32, 1307–1316. https://doi.org/10.4319/lo.1987.32.6.1307

Springer, T.G., Pigford, R.L., 1970. Influence of Surface Turbulence and Surfactants on Gas Transport through Liquid Interfaces. Ind. Eng. Chem. Fundam. 9, 458–465. https://doi.org/10.1021/i160035a025

Thornton, D.C.O., Brooks, S.D., Chen, J., 2016. Protein and Carbohydrate Exopolymer Particles in the Sea Surface Microlayer (SML). Front. Mar. Sci. 3. https://doi.org/10.3389/fmars.2016.00135

Tsai, W., Liu, K., 2003. An assessment of the effect of sea surface surfactant on global atmosphere-ocean CO_2 flux. J. Geophys. Res. Oceans 108, 2000JC000740. https://doi.org/10.1029/2000JC000740

Tukey, J.W., 1977. Exploratory data analysis. Springer.

Upstill-Goddard, R.C., 2006. Air–sea gas exchange in the coastal zone. Estuar. Coast. Shelf Sci. 70, 388–404. https://doi.org/10.1016/j.ecss.2006.05.043

https://doi.org/10.5194/egusphere-2025-4050
Preprint. Discussion started: 21 August 2025

Verdugo, P., Alldredge, A.L., Azam, F., Kirchman, D.L., Passow, U., Santschi, P.H., 2004. The oceanic gel phase: a bridge in the DOM–POM continuum. Mar. Chem. 92, 67–85. https://doi.org/10.1016/j.marchem.2004.06.017

Vitousek, P.M., 2004. Nutrient cycling and limitation: Hawai'i as a model system. Princeton University Press.

Wegman, E.J., 1972. Nonparametric probability density estimation. J. Stat. Comput. Simul. 1, 225–245. https://doi.org/10.1080/00949657208810017

Williams, P.M., Carlucci, A.F., Henrichs, S.M., Van Vleet, E.S., Horrigan, S.G., Reid, F.M.H., Robertson, K.J., 1986. Chemical and microbiological studies of sea-surface films in the Southern Gulf of California and off the West Coast of Baja California. Mar. Chem. 19, 17–98. https://doi.org/10.1016/0304-4203(86)90033-2

Wurl, O., Holmes, M., 2008. The gelatinous nature of the sea-surface microlayer. Mar. Chem. 110, 89–97. https://doi.org/10.1016/j.marchem.2008.02.009

Wurl, O., Miller, L., Röttgers, R., Vagle, S., 2009. The distribution and fate of surface-active substances in the sea-surface microlayer and water column. Mar. Chem. 115, 1–9. https://doi.org/10.1016/j.marchem.2009.04.007

Wurl, O., Stolle, C., Van Thuoc, C., The Thu, P., Mari, X., 2016. Biofilm-like properties of the sea surface and predicted effects on air–sea CO2 exchange. Prog. Oceanogr. 144, 15–24. https://doi.org/10.1016/j.pocean.2016.03.002

Yang, G.-P., 1999. Dimethylsulfide enrichment in the surface microlayer of the South China Sea. Mar. Chem. 66, 215–224. https://doi.org/10.1016/S0304-4203(99)00042-0

Zhang, Z., Liu, C., Liu, L., 2006. Physicochemical Studies of the Sea-Surface Microlayer. Front. Chem. China 1. https://doi.org/10.1007/s11458-005-0003-8

Zuur, A.F., Ieno, E.N., Smith, G.M. (Eds.), 2007. Linear regression, in: Analysing Ecological Data. Springer New York, New York, NY, pp. 49–77. https://doi.org/10.1007/978-0-387-45972-1_5