Science Advances
AAAS

Supplementary Materials for

Trace metal stoichiometry of dissolved organic matter in the Amazon Plume

Martha Gledhill *et al.*

Corresponding author: Martha Gledhill, mgledhill@geomar.de

Sci. Adv. **8**, eabm2249 (2022)
DOI: 10.1126/sciadv.abm2249

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Supplementary Text

1. Chromatographic characteristics

Size exclusion chromatography of our $DOM_{SPE\text{-}amb}$ fraction showed distinct behaviours for the elements and organic matter properties (Fig. S4). We detected a total of three peaks, indicative of three molecular weight fractions. In all samples we observed a peak for all parameters that eluted with a retention time (RT) between 22 and 28 min (Table S3). We found the average molecular size for this $DOM_{SPE\text{-}amb}$ fraction was lower than for the riverine humic substances Suwannee River humic acid (SRHA, 3S101H, International Humic Substances Society) and Suwannee River fulvic acid (SRFA, 1S101F, International Humic Substances Society), (Table S3) and we designate this the medium molecular weight (MMW) fraction. This MMW peak coincided with maxima in both negative and positive ion counts observed in ESI-MS and maxima in all elements detected by ICP-MS (Table S3). We observed a further, less intense maxima at a lower RT (10-15 min) for $A_{254SPE\text{-}amb}$, $P_{SPE\text{-}amb}$, $Al_{SPE\text{-}amb}$, $Fe_{SPE\text{-}amb}$ and $Mn_{SPE\text{-}amb}$. The RT for this peak was similar to a high molecular weight polystyrenesulfone standard (150 kDa, RT S peak = 12.14 min), and thus we designate this the high molecular weight (HMW) fraction. A third peak was observed for Mn and Zn, with RT of 28-33 min and which eluted at a similar retention time to our lowest MW standard (thiamine, MW = 265 Da, RT S peak = 27.6±0.1) and we designate this a low MW (LMW) fraction. We detected no $A_{254SPE\text{-}amb}$ signature associated with the LMW peak.

In RPLC, we detected one major and one minor fraction for $A_{254SPE\text{-}amb}$, (Fig. S5) which corresponded to $DOM_{SPE\text{-}amb}$ fractions with differing relative hydrophobicities at the pH of our analysis (ca. 6.5) and we refer to these as a hydrophobic fraction (Hphob, $3 < RT < 8$ min) and a hydrophilic fraction (Hphil, $RT < 3$ min), respectively. $P_{SPE\text{-}amb}$ and $S_{SPE\text{-}amb}$ were associated with both the Hphob and Hphil fractions, whilst $Fe_{SPE\text{-}amb}$, $Al_{SPE\text{-}amb}$ and $Cu_{SPE\text{-}amb}$ were only associated with the Hphob fraction. The Hphil and Hphob peaks for $Co_{SPE\text{-}amb}$, $Mn_{SPE\text{-}amb}$, $Ni_{SPE\text{-}amb}$ and $Zn_{SPE\text{-}amb}$ were not well resolved, particularly at low salinities. This is illustrated by an observed decrease in the retention time of the peak apex, which decreased with the order $Cu_{SPE\text{-}amb}$>$Fe_{SPE\text{-}amb}$>$S_{SPE\text{-}amb}$>$Al_{SPE\text{-}amb}$>$P_{SPE\text{-}amb}$>$Zn_{SPE\text{-}amb}$>$Ni_{SPE\text{-}amb}$>$Co_{SPE\text{-}amb}$>$Mn_{SPE\text{-}amb}$ (Table S3).

The different chromatographic fractions showed different trends in the study region. The HMW $P_{SPE\text{-}amb}$ and $A_{254SPE\text{-}amb}$ fractions correlated with chlorophyll *a* ($r=0.54$, $p<0.01$, $n=14$). Although we observed a mid-estuarine peak in the proportion of Fe associated with the HMW fraction (Fig. S6), HMW Fe, Al and Mn were most abundant in the riverine end-member. Elements associated with the HMW $DOM_{SPE\text{-}amb}$ fraction therefore showed a change from Fe, Al and Mn in the riverine end-member, to P and Fe in the inner estuary and mangrove region suggesting a change in character between riverine and inner estuary HMW fractions. The proportion of $Mn_{SPE\text{-}amb}$ associated with LMW peak increased with increasing salinity from 50% in the riverine end-member to 100% in samples from the NBC whilst the relative amount of $Zn_{SPE\text{-}amb}$ in the LMW fraction was not detectable in many samples and overall quite variable, especially in the NBC samples, where $Zn_{SPE\text{-}amb}$ concentrations were low.

The Hphob fraction was consistently the dominant peak for $A_{254SPE\text{-}amb}$, $P_{SPE\text{-}amb}$, $Fe_{SPE\text{-}amb}$, $Al_{SPE\text{-}amb}$ and $Cu_{SPE\text{-}amb}$ throughout the plume (Fig. S6). The proportion of $S_{SPE\text{-}amb}$ associated with the Hphob fraction decreased with increasing salinity in the plume, and $S_{SPE\text{-}amb}$ was mostly associated with the Hphil fraction at higher salinities. In contrast, P was predominantly in the Hphil fraction in our lowest salinity sample, changed rapidly to 80% association with the Hphob fraction at salinities between 3 and 10, but then became increasingly associated with the Hphil fraction with increasing salinity (Fig. S6). About 25% of $Co_{SPE\text{-}amb}$, $Mn_{SPE\text{-}amb}$ and $Ni_{SPE\text{-}amb}$ were associated with the Hphil peak fraction throughout the plume, although $Ni_{SPE\text{-}amb}$ showed some variability.

2. Molecular characteristics

We used our high resolution ESI-MS data to examine the molecular character of features detected chromatograms. Example mass spectra are shown in Fig. S7. We detected a total of 4549 features (i.e. retention time ~ m/z combinations) in positive and negative ion modes in SEC, and 4566 in RPLC (Table S5). Assuming a mass resolution of 3 ppm and protonation/deprotonation was the dominant mechanism of ionisation, approximately 50% of the calculated neutral masses were common to both SEC and RPLC. In contrast, ionisation modes were more distinct, with only 151 and 407 features common to both positive and negative modes in SEC and RPLC, respectively. The majority of the detected features had peak apexes within the retention time ranges observed for Hphob and MMW peaks (>93 % of the detected features), with all but 2 of the remaining features occurring in LMW or Hphil fractions. Calculation of the relationship between the mass defect and the nominal mass showed the MMW and Hphob fractions had slopes slightly higher than those observed for CH_2O (3.5×10^{-4}) (Table S5). The Hphil fraction had slopes lower than the Hphob fraction, but still high enough to suggest the presence of multiple hydrogens per ion, and thus ions of organic rather than inorganic nature. We observed a marked contrast between ionisation modes in SEC, with the mass defect slope of the negative ions pointing to lower H content in comparison to the positive ions. We could successfully assign unambiguous formulas (at 2 ppm resolution) containing only C, H, O and N to at least 83% of the features in positive ion mode and 64% of the features in negative ion mode in the MMW and Hphob retention time ranges. In contrast, less than 30% of the features detected in the Hphil or LMW peak fractions could be assigned a formula. We divided our assigned masses into groups broadly representing polycyclic aromatics (PCAs), mixed aromatic/aliphatic like formula, highly unsaturated, unsaturated aliphatic compounds, saturated compounds with no nitrogen, highly oxygenated compounds and aliphatic compounds containing N according to aromaticity index, hydrogen/carbon, oxygen/carbon and nitrogen content as described in (*29*). The proportions of ions falling into the different groups was similar in MMW and Hphob fractions. The major fractions identified in negative ion mode fell into the highly unsaturated and unsaturated aliphatic compound groups, in-line with observations determined for DOM extracted at pH 2 in the area (*29*). In positive ion mode, we detected many more N containing compounds, and this fraction made up >30% of the assigned formulas in the LMW and Hphob fractions, respectively. The relatively high proportion of N containing compounds likely reflects the greater ease with which N acquires a positive charge in ESI, however we note formation of NH_4^+ adducts could also have contributed to increased detection of N containing compounds.
The limited number of ESI-MS features and low or absent A_{254} makes it difficult to assign a definitive character to these fractions and the elements eluting in these fractions could have been present as either organic or inorganic compounds. For example, metal complexes with ethylenediaminetetraacetic acid (EDTA), which we used as standards for quantification in RPLC all elute at retention times between 1.5 and 1.8 min and metal-DOM complexes with similar hydrophobicities to these complexes (which all have a residual charge) would therefore also be expected to elute in the Hphil fraction. In addition, we cannot rule out the possibility that the HMW fraction was formed via precipitation of colloidal metal oxy-hydroxides during the preparation of the samples, particularly in the lowest salinity samples.

3. Methods

3.1. Determination of elemental concentrations, spectroscopic and molecular characteristics in $DOM_{SPE\text{-}amb}$.

We used biocompatible high performance liquid chromatograph (HPLC, Biocompatible Ultimate 3000 UHPLC, Thermo) coupled in parallel to i) an inductively coupled plasma mass

spectrometer (ICP-MS, Element XR, Thermo) and ii) a diode array detector (DAD, Ultimate 3000, Thermo) followed by an electrospray ionisation mass spectrometer (ESI-MS, Q Exactive, Thermo). The autosampler was temperature controlled (8°C) and a 25 µL injection volume was used for both chromatographic modes of analysis. Reverse phase liquid chromatography (RPLC) was performed using a Hamilton PRP C18 5 µm 150 × 2.1 mm PEEK column using a gradient from 5:95 methanol: 10 mM ammonium acetate (vol:vol) to 100 % methanol over 15.5 minutes at a flow rate of 400 µL min^{-1} followed by 2.5 min at 100 % methanol prior to returning to the start conditions for equilibration (5 mins) and injection of the next sample. Size exclusion chromatography (SEC) was performed using a TOSOH BioAssist G3SWXL, 5 µm, 300 × 7.8 mm PEEK column using an isocratic flow of 25 mmol L^{-1} ammonium acetate with 20 % methanol (*77*) at a flow rate of 500 µL min^{-1}. The column oven was set to 30°C for both modes of chromatography.

For ICPMS analysis we gravity fed a solution of 4 µg L^{-1} chromium and 1 µg L^{-1} cerium in 0.32 mol L^{-1} nitric acid (sub boiled redistilled) into the flow from the HPLC via a 1 L infusion bag (Flexboy, Sartorius), 50 µm ID PEEK tubing and a T connector. The Ce/Cr solution allowed for ICP-MS tuning and drift correction during the ICP-MS analysis. The flow rate for the Ce/Cr solution was set to approximately 75 µL min^{-1} via adjustment of the PEEK tubing length. Solvents in the flow were removed via a 200 µL min^{-1} desolvating nebuliser (Aridus, CETAC II) and oxygen was added as a plasma gas in order to prevent build up carbon during the analysis. The flow rates of oxygen, nitrogen and sweep gas were adjusted to optimise ion counts and reduce the presence of cerium oxide to <10 % of the Ce abundance prior to the analysis. Results reported here were obtained with sweep gas and nitrogen gas flow rates of 8.54 L min^{-1} and 7 L min^{-1}, respectively for SEC analysis and 8.46 L min^{-1} and 7 L min^{-1} for RPLC analysis. An oxygen flow rate of 0.021 L min^{-1} was used for SEC and 0.035 L min^{-1} for RPLC.

We assessed the potential contribution from the absorbance of inorganic TMs onto our SPE cartridges by preconcentrating 1 L artificial seawater (n=4) comprising 0.4 mol L^{-1} NaCl, 0.02 mol L^{-1} $MgCl_2$, 0.01 mol L^{-1} $CaCl_2$, 2.5 mmol L^{-1} $NaHCO_3$ and containing 100 nmol L^{-1} of each trace metal (Al, Co, Cu, Fe, Mn, Ni and Zn). Cartridges were rinsed and eluted following the same procedure as applied for samples. The percentage inorganic TMs observed within the retention time range of the Hphob and MMW peaks are provided in Table S6. The values are at least one order of magnitude lower than recoveries estimated for samples (Table S6).

The DAD detector collected absorption spectra between 210 and 800 nm at a rate of 1 Hz with a bandwidth of 10 nm.

The ESI-MS detector was mass calibrated prior to each analysis using recommended calibration solutions (Pearson, Thermo) according to the manufacturer's instructions. We used charge switching to allow for collection of both negative and positive ions in one run. The mass resolution was set to 140 000 ppm. Maximum injection time was 100 msec and the AGC target was 1×10^6. The m/z range was 150-2000, capillary temperature was 350 °C, nitrogen gas flow rate set to 40 arbitrary units and auxiliary gas flow rate to 10 arbitrary units and the S-lens RF level was set to 50 Hz. The source temperature was set to 350°C and 400°C in SEC and RPLC modes respectively. Spray voltage was 2.5 and 3.5 V for positive ion mode for SEC and RPLC respectively and 1.5 and 2.5 V in negative ion mode for SEC and RPLC respectively. We used the following lock masses which are present as ubiquitous background ions in our system: I^- (M^-; m/z = 126.9050) in negative ionisation mode and polysiloxane ($[M+H]^+$; m/z = 371.10124) and diisooctyl phthaltate ($[M+H]^+$; m/z = 391.284) in positive ionisation mode.

The autosampler triggered data collection for all three detectors, however differences in dead volume meant that there was a small time offset (0.7 min) between the ICP-MS and UV detector and 0.01 min between UV detector and ESI-MS. We used the elution time of thiamine to unify retention times.

3.2 Data analysis pipeline.

Data collected from ICP-MS and DAD detectors were converted into comma delimited text files and imported into R (*78*) for determination of peak areas using the trapz function in the package pracma (version 1.9.9).

For ESI-MS data Xcalibur RAW files were converted to either negative or positive .mzML files using MSconvert in Proteowizard (*79*) prior to import into MZmine (*80*). Masses with an abundance over 5×10^5 were detected and then mass chromatograms extracted for these masses over the time ranges of 1.2-10 mins for RPLC and 9-30 mins for SEC. The m/z tolerance was set at 3 ppm. Features (i.e peaks) were smoothed with a value of 5, isotopes grouped and an aligned list created using the join aligner algorithm, an m/z tolerance of 3 ppm and a retention time tolerance of 2 absolute minutes. Duplicates were subsequently removed, and then features further filtered so that any detected feature was observed at least twice in the data set, none of the detected features had peaks that also occurred in the procedural SPE blanks, and all features had a duration of at least 0.5 minutes. The feature list was then manually checked to ensure successful removal of background ions. Finally, the aligned feature list was gap-filled requiring the same retention time and m/z with a m/z tolerance of 3 ppm. The average mass to charge ratio for each feature in MZmine was used to predict molecular formula using the R package MFAssignR (*81*).

Contamination issues

We tried to minimise contamination from both organic compounds and trace metals. Typically, analysts concerned with organic matter use glass or metal and tend to avoid plastic. On the other hand, trace element analysts cannot use metals and glass for sample handling and prefer plastic laboratory materials (*82*). Analysis of TMs bound to DOM therefore requires compromises. Where possible, we used fluorinated plastics (e.g. polytetrafluoroethylene for solutions used in sample preparation and SPE elution and perfluoroalkoxy alkane for aqueous HPLC eluants). We used high density polyethylene for sample bottles and polypropylene vials for sample extract storage (Eppendorf, 5 mL) and for autosampler vials (Fisher).

To remove peaks arising from plastic contaminants from our positive ionisation data we derived a custom list of polyethylene glycols (PEGs), polypropylene glycols (PPGs) and nylon derivatives from both RPLC and SEC data. We identified these compounds by comparison with known m/z values (*83*) and from the characteristic way in which the features associated with these compounds increase in retention time and m/z space. We also observed significant formation of acetate clusters in both RPLC and SEC since we used ammonium acetate to buffer our eluants to pH 6.5. The features identified as PEGs, PPG, Nylons or acetate clusters consisted of discrete peaks with peak durations of < 1 min, and thus contrasted with the broad peaks associated with marine DOM. We identified multiple putative PEG series with different headgroups and/or adducts and extended any PEG or acetate cluster series to cover a mass range from m/z 150 to 1000. Only one PPG series was identified, while we identified two peaks for Nylon with several adducts and many clusters of Ca, Mg, Na and K acetates. We did not detect the presence of any TM acetates. We provide the list of identified polymers and acetate clusters as a data file that can be downloaded from https://doi.pangaea.de/10.1594/PANGAEA.941455. The m/z values were used to identify plasticizers and acetate features in MZmine and exclude these compounds from our feature lists.

3.3 NICA-Donnan modelling.

We used either the generic fulvic acid NICA parameters or NICA parameters derived from the equations described in ref (*36*) for fulvic acids to predict TM:C ratios from DOC and dissolved TMs at ambient pH, ionic strength (calculated from ion pairing) and temperature in our study area, assuming binding sites were composed of a heterogeneous mixture with properties similar to fulvic acid.

The generic binding site density for fulvic acid was used together with generic fulvic acid NICA parameters for H^+, Mg^{2+}, Ca^{2+}, Sr^{2+} and Mn^{2+} (*36*, *84*). We used the Donnan model as described for fulvic acid (*85*), however it should be noted that in our study the maximum percentage of any TM partitioned into the Donnan volume was 1.2% for Mn at our lowest salinity (0.03).

NICA parameters for Fe^{3+} and Cu^{2+} were taken from ref (*86*), and were derived for the estuarine waters of the Solent in the United Kingdom. Parameters for Co^{2+}, Ni^{2+} and Zn^{2+} were derived from their hydrolysis parameters K_{OH}, (Table S7) using the following equations provided in ref (*36*). We obtained log K_{OH} from the MINTEQ4 data base in ORCHESTRA.

$n_1 = 0.14 - 0.055 \times \log K_{OH}$

$n_2 = 0.76 \times n_1$

$n_1 \times \log K_1 = 0.35 \times \log K_{OH} + 2.43$

$n_2 \times \log K_2 = 0.78 \times \log K_{OH} + 8.21$

The generic fulvic acid NICA parameters for binding of Fe^{3+} and Al^{3+} binding to phenolic groups ($\log K_2$) reported in ref (*36*) are high (36 and 12.1 respectively) and Fe^{3+} values derived specifically for marine DOM have proved to be much lower (8-11.2) (*19*, *86*, *87*). Since C:Al ratios in the MMW and Hphob fractions, and the chromatographic behaviour of Al was similar to Fe and Cu, we hypothesized that Al would have similar NICA binding parameters to Fe or Cu. We found parameters describing Cu binding produced C:Al ratios close to observed values and therefore used the Cu NICA parameters for Al^{3+} as a first estimate to describe Al binding to DOM in our calculations. The complete list of NICA parameters is given in Table S7.

Speciation calculations were undertaken with the software program ORCHESTRA (*73*) following approaches described in (*19*, *87*). We used pH on the NBS scale for calculations since this is the scale appropriate for the thermodynamic constants used by the ORCHESTRA data base. Carbonate concentrations and pH_{NBS} used in ORCHESTRA were calculated using CO2SYS (*74*), all other major ion concentrations were calculated from salinity assuming conservative mixing of the Amazon River freshwater end member concentrations (*88*, *89*) with those of seawater (*90*).

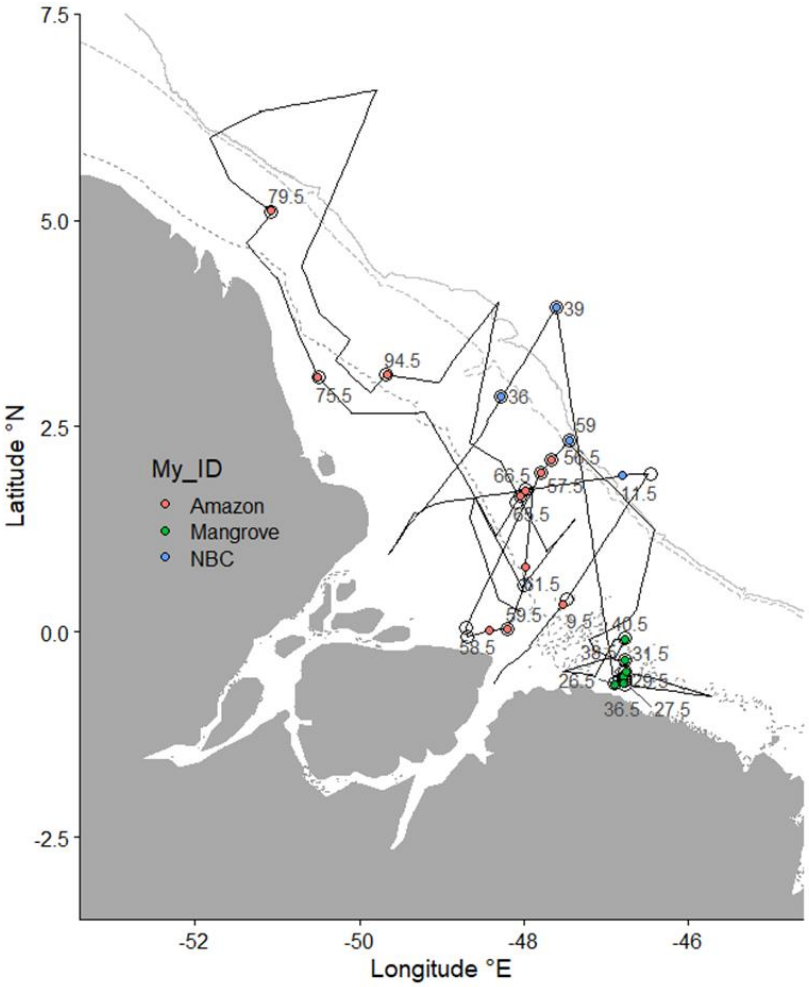

Fig. S1. Map of the study area showing the cruise track (black line) used for interpolation of salinity in Fig 1.

In addition, this map shows the position of samples take for the determination of trace metal concentrations (filled circles). Sample positions shown with .5 denote surface samples collected underway with the Towfish, whilst integer stations (36, 39 and 59) denote surface samples (ca. 20 m depth) collected with the TM clean CTD rosette (stations 36, 39, 59). With three exceptions, the positions of the underway samples corresponded closely to the position of the $DOM_{SPE\text{-}amb}$ samples (shown in Fig. 1), which are also plotted for reference as the larger, open circles. For three $DOM_{SPE\text{-}amb}$ stations (19 and 66, 68, see Fig 1) either failure of the Towfish (68, 66, shown on Fig 1) or the TM-CTD (19, shown on Fig 1) meant that the closest TM data sample (11.5 and 58.5) was offset from the position of the $DOM_{SPE\text{-}amb}$ samples. This is most critical for the stations 66 & 68 since these stations were the lowest salinity samples.

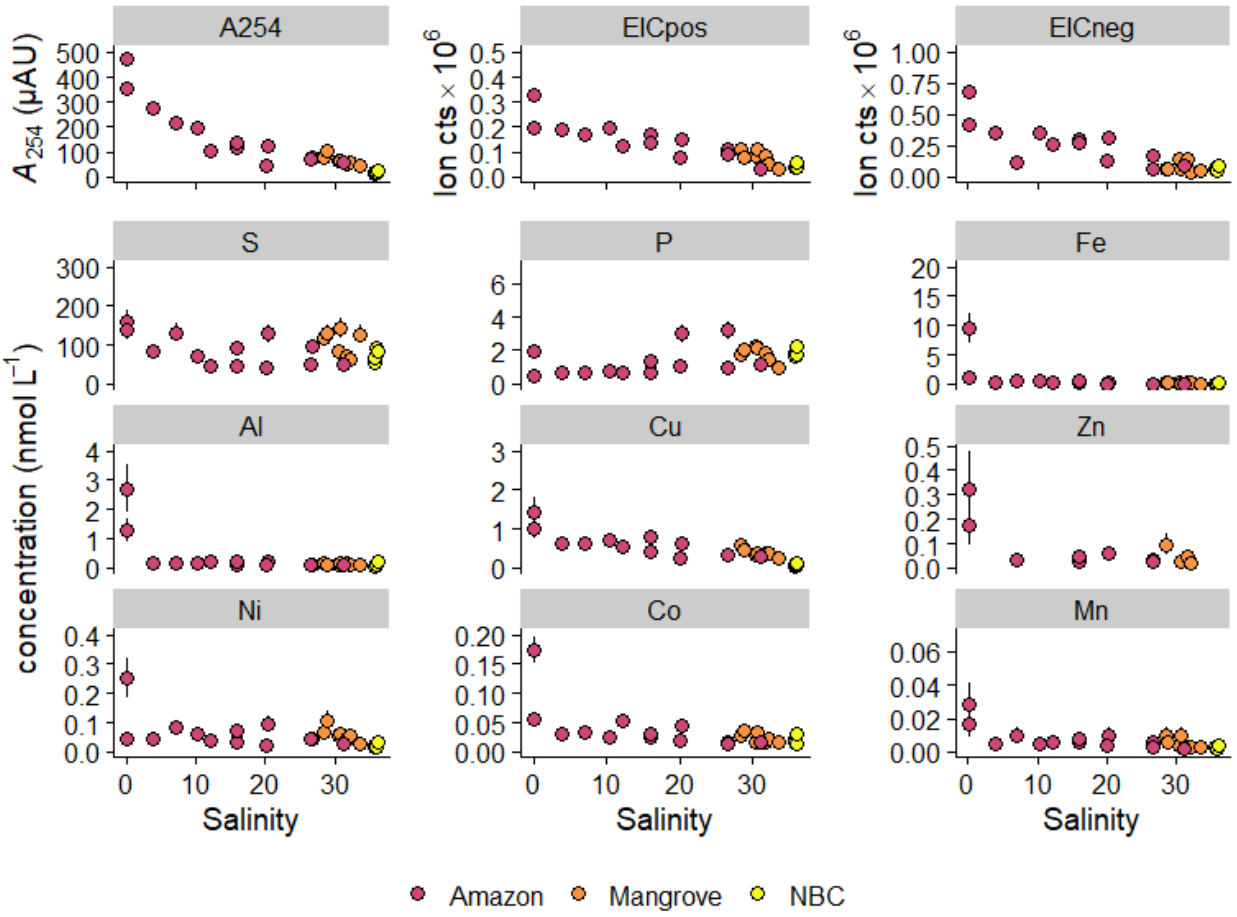

Fig. S2. Total concentrations of elements and ion abundances determined in DOM$_{SPE\text{-}amb}$ samples collected in our study area.

Results shown here were calculated from the sum of peak areas observed in reversed phase chromatography data. Vertical bars represent the analytical uncertainty associated with element concentrations determined in this study. Colours show geographic grouping with red indicative of Amazon estuary samples, orange indicative of samples collected of an area of mangrove forest and yellow indicative of samples collected in the NBC.

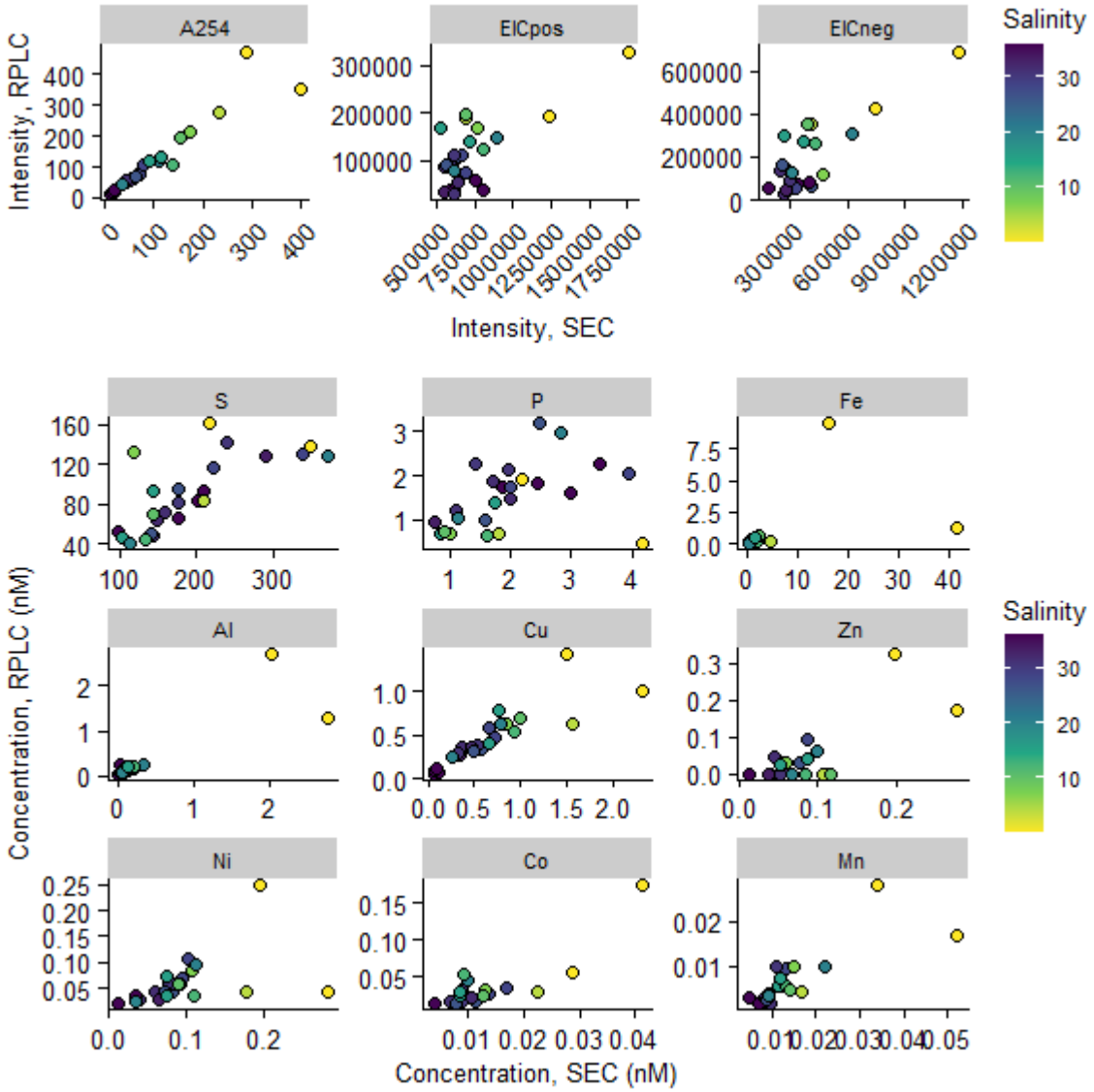

Fig. S3. Comparison of the sum of peak areas determined by size exclusion chromatography (SEC) and reversed phase liquid chromatography (RPLC)

Values show absorbance at 254 nm (A254, µAU), total extracted positive ion current (EICpos, ion counts), total extracted negative ion current (EICneg, ion counts) and concentrations determined for elements quantified by ICP-MS in SPE extracts obtained in this study. Colour bar shows salinity. Regression slopes indicated that elemental concentrations determined by SEC were between 0.2 (Co) and 2 (Fe) times those determined from RPLC analysis. Since values for absorbance at 254 nm observed in SEC were within 12% of those observed in RPLC, we suggest the differences observed for elemental abundancies arose because of a combination of matrix effects arising from the different solvents used in SEC versus RPLC and analytical uncertainties. One of our riverine end-member samples showed a large reduction in intensity of concentration in RPLC compared to SEC and we suspect this was due to loss of the high molecular weight (HMW) fraction on RPLC analysis of this sample.

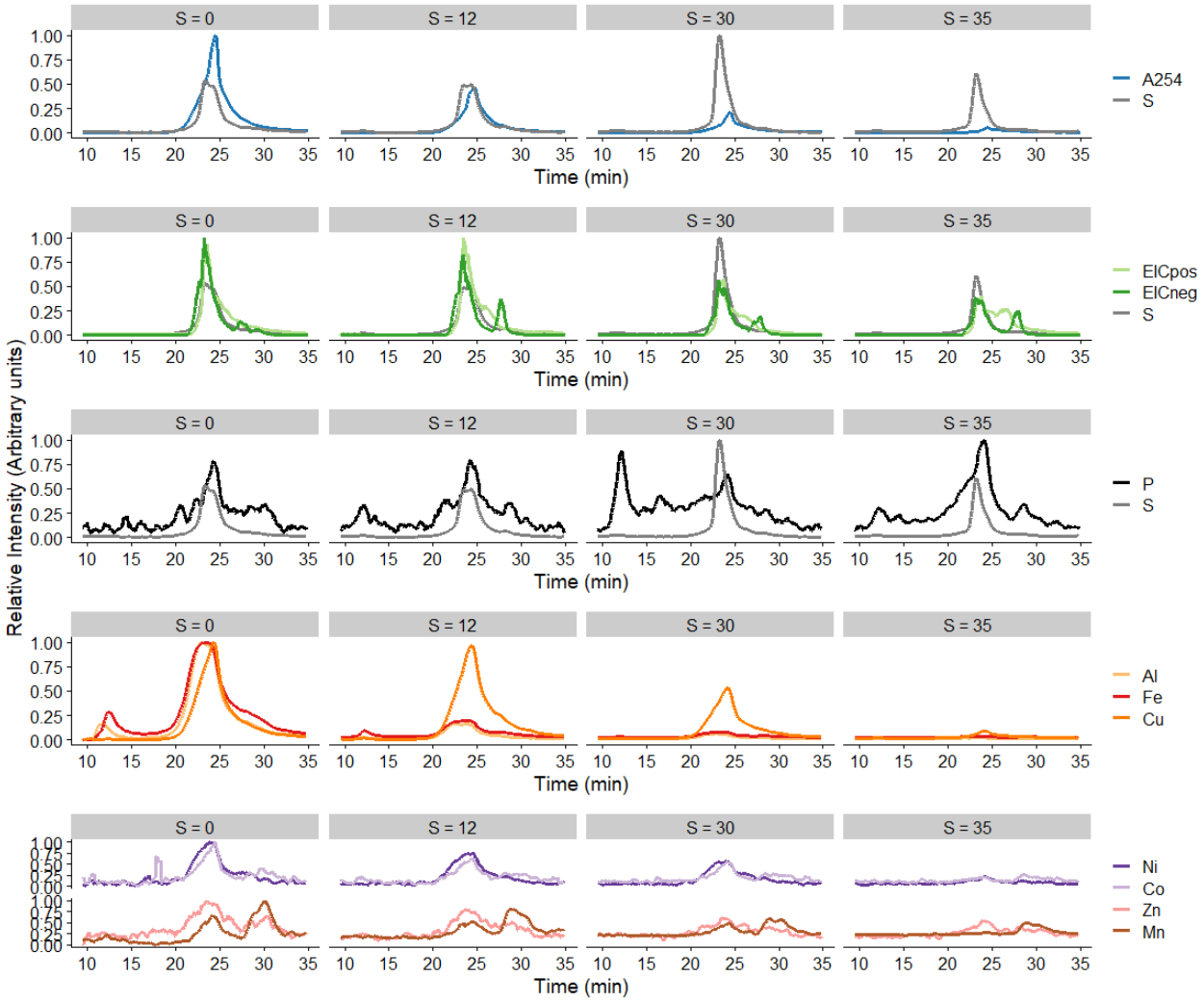

Fig. S4. Example size exclusion chromatograms for four representative samples collected at the riverine end-member (Station 66, S=0), in the estuary (Station 72, S=12), in the Mangrove area (Station 45, S=30) and in the North Brazil current (Station 19, S=35).

Chromatograms show UV absorbance at 254 nm (A254), extracted ion abundance for mass:charge ratios for ESI-MS in positive and negative ionisation modes (EICpos and EICneg) and element intensity determined by ICP-MS. Intensities were normalised to the maximum intensity observed for each parameter. Response for S is shown in the top three rows for reference.

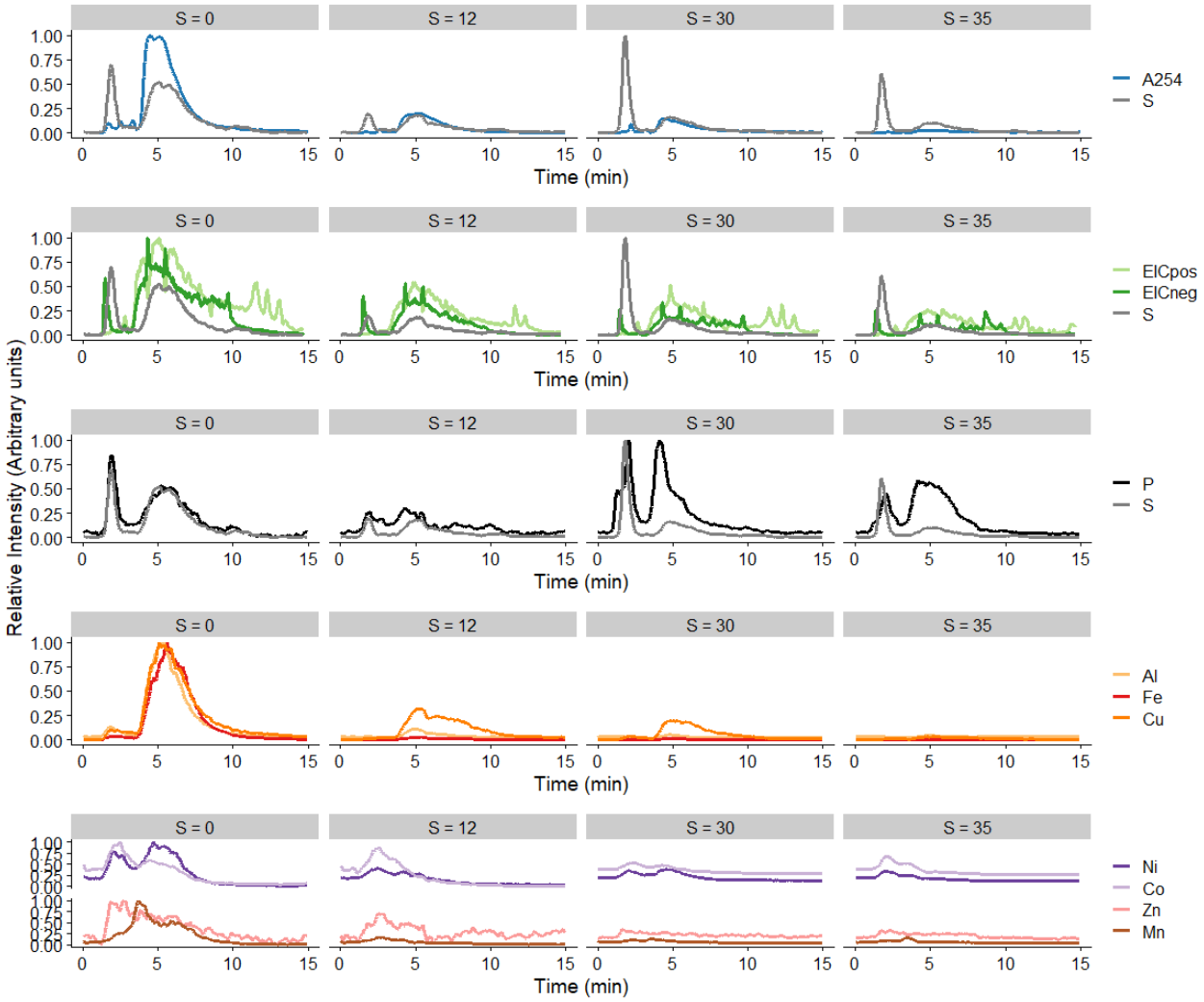

Fig. S5 Example reverse phase chromatograms for four representative samples collected at the riverine end-member (Station 66, S=0), in the estuary (Station 72, S=12), in the Mangrove area (Station 45, S=30) and in the North Brazil current (Station 19, S=35).

Chromatograms show A254, extracted ion abundance for mass:charge ratios for peaks with apexes between 1 and 8 minutes observed on ESI-MS in positive and negative ionisation modes (EICpos and EICneg) and element abundance determined by ICP-MS. Intensities were normalised to the maximum intensity observed for each parameter. Response for S is shown in the top three rows for reference.

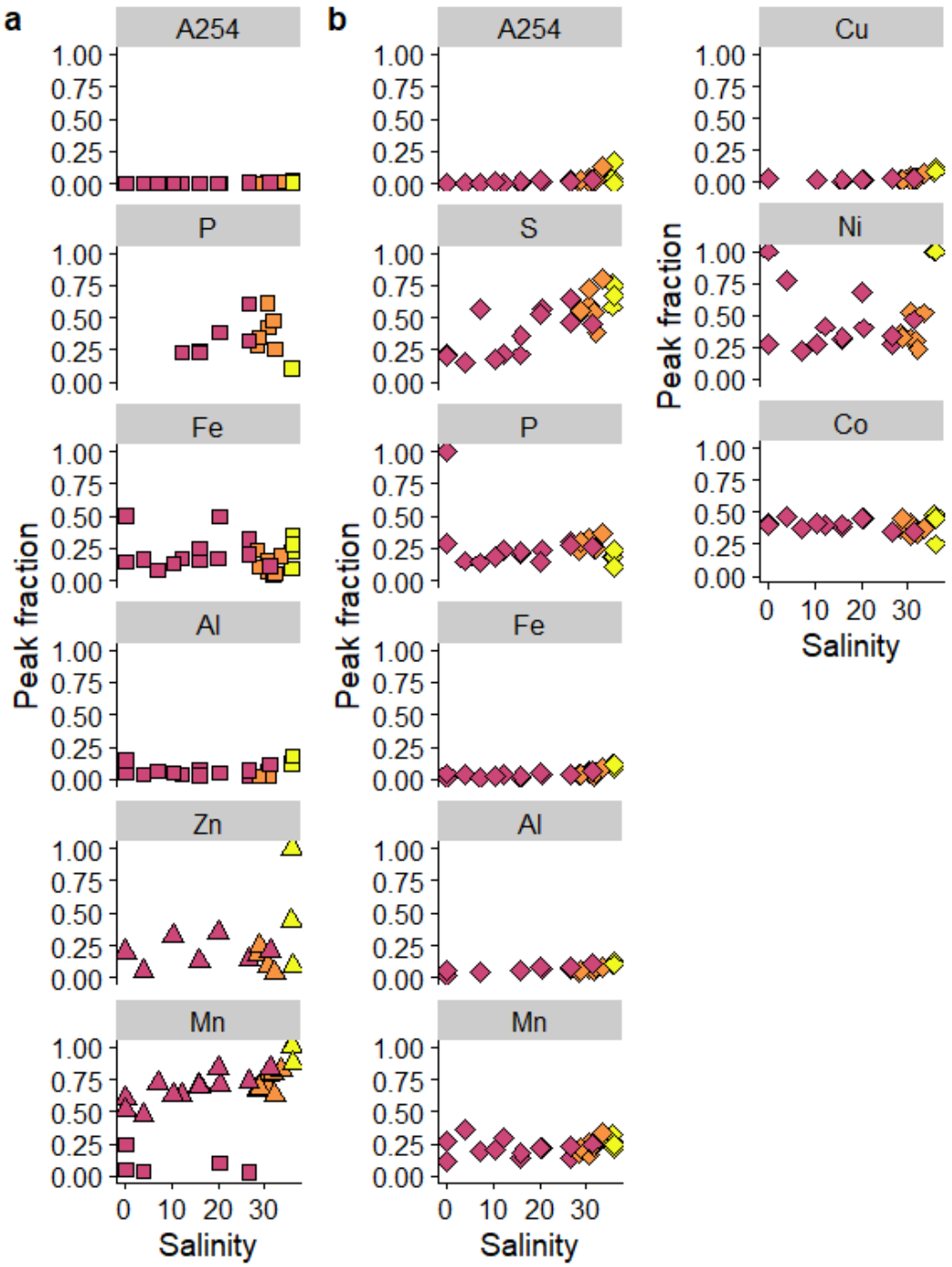

Fig. S6 Variation in abundance (relative to the total concentration) of different peak fractions with salinity.

(a) Relative abundances of the HMW (square) and LMW (triangle) peaks for A254, P, Fe, Al, Zn and Mn observed in SEC chromatograms. (b) Variation in abundance relative to the total concentrations for the hydrophilic fraction (diamonds) for all parameters except Zn and EIC observed in RPLC chromatograms. Hydrophobic and hydrophilic Zn peaks could not be separated in RPLC because of the low sensitivity of Zn ICP-MS analysis and thus Zn associated with the different fractions could not be calculated. Colours show geographic grouping with red indicative of Amazon estuary samples, orange indicative of samples collected off an area of mangrove forest and yellow indicative of samples collected in the NBC.

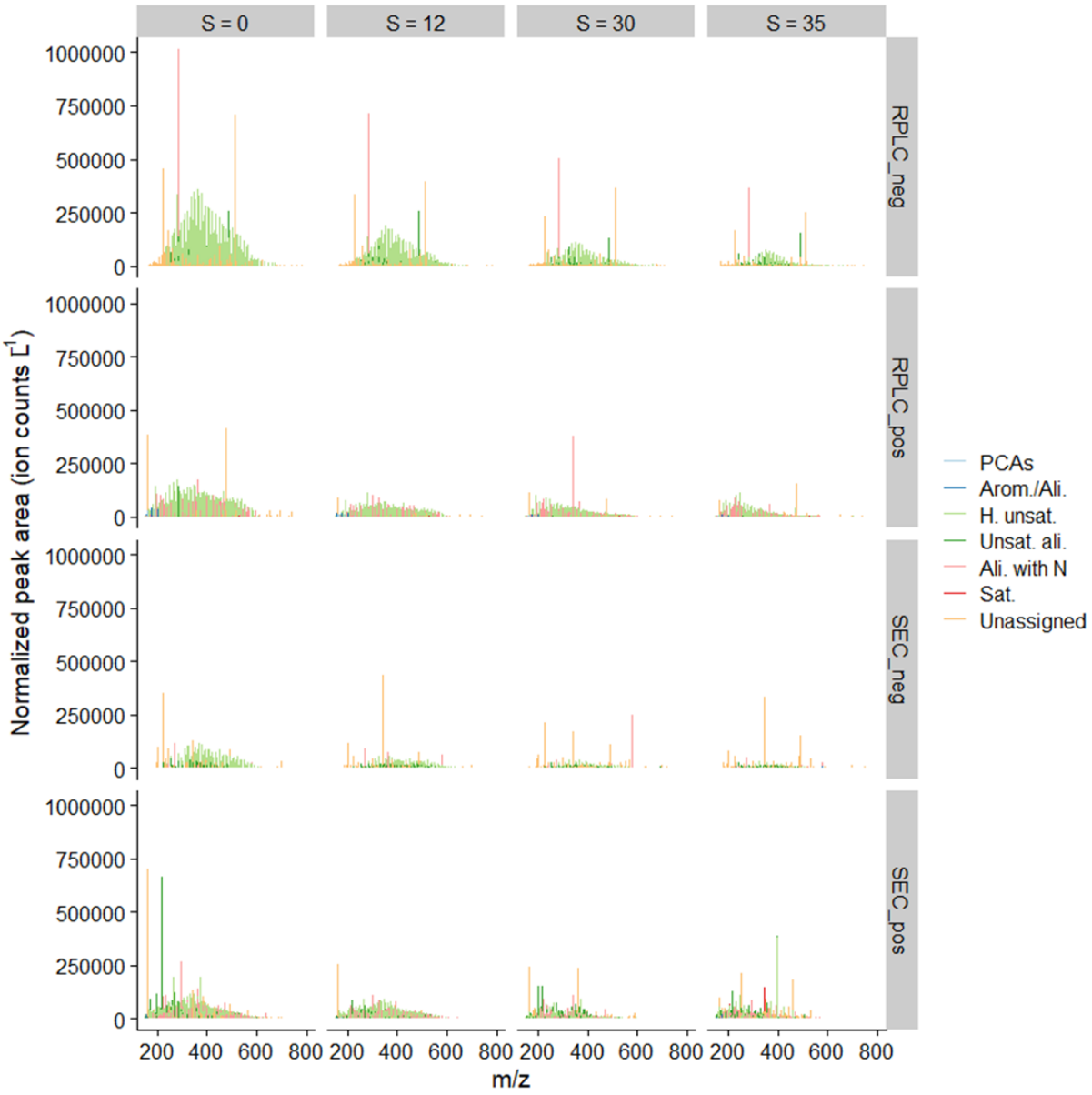

Fig. S7 Example mass spectra obtained for samples from stations 66 (S=0), 72 (S=12), 45 (S=30) and 19 (S=35) for the Hphob (RPLC) and MMW (SEC) peaks after ionisation in positive (pos) or negative (neg) ion mode.

Height of the line represents the normalised peak area (i.e. peak area after accounting for the preconcentration factor for each sample). Peaks are coloured according to different groups of assigned formulas (see Table S5 for definitions). Averaged mass spectra for Hphob and MMW had features typical of marine DOM isolates (*29*, *91*).

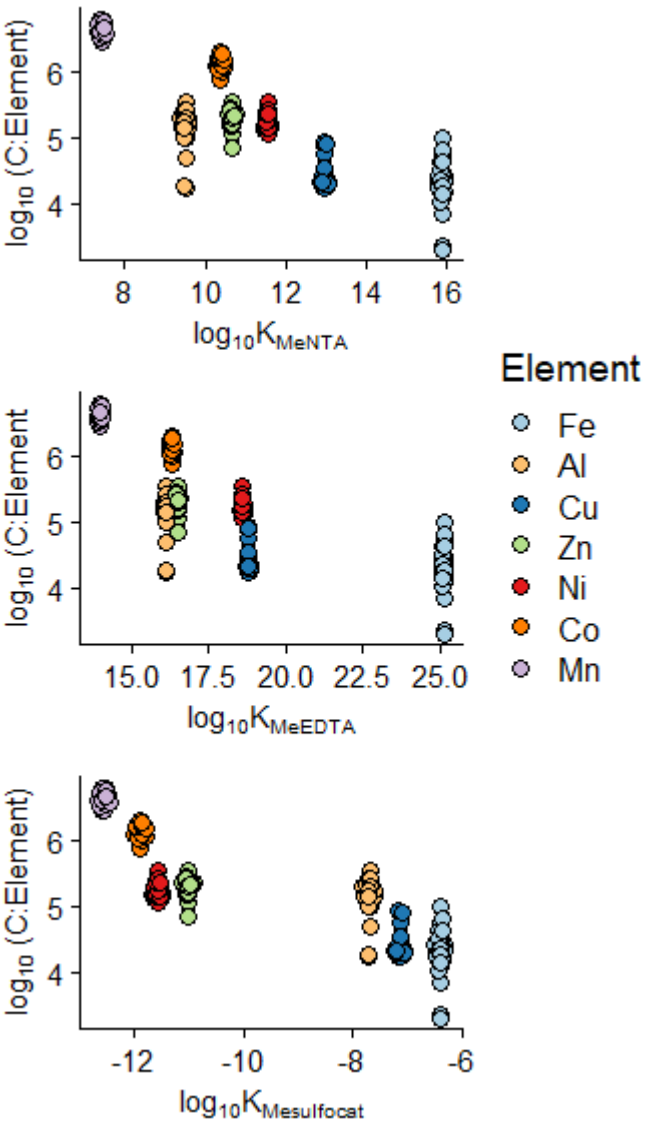

Fig. S8 Relationships between TM thermodynamic stability constants for different ligands and C:TM ratios observed for the MMW peak fraction.

NTA: nitrilotriacetic acid, EDTA: ethylenediaminetetraacetic acid, sulfocat: 4-sulfocatechol. In each case the constants for the 1:1 reaction are presented and charges are omitted from the x-axis title for simplicity. Constants were obtained from the NIST critical stability constant data base (*92*).

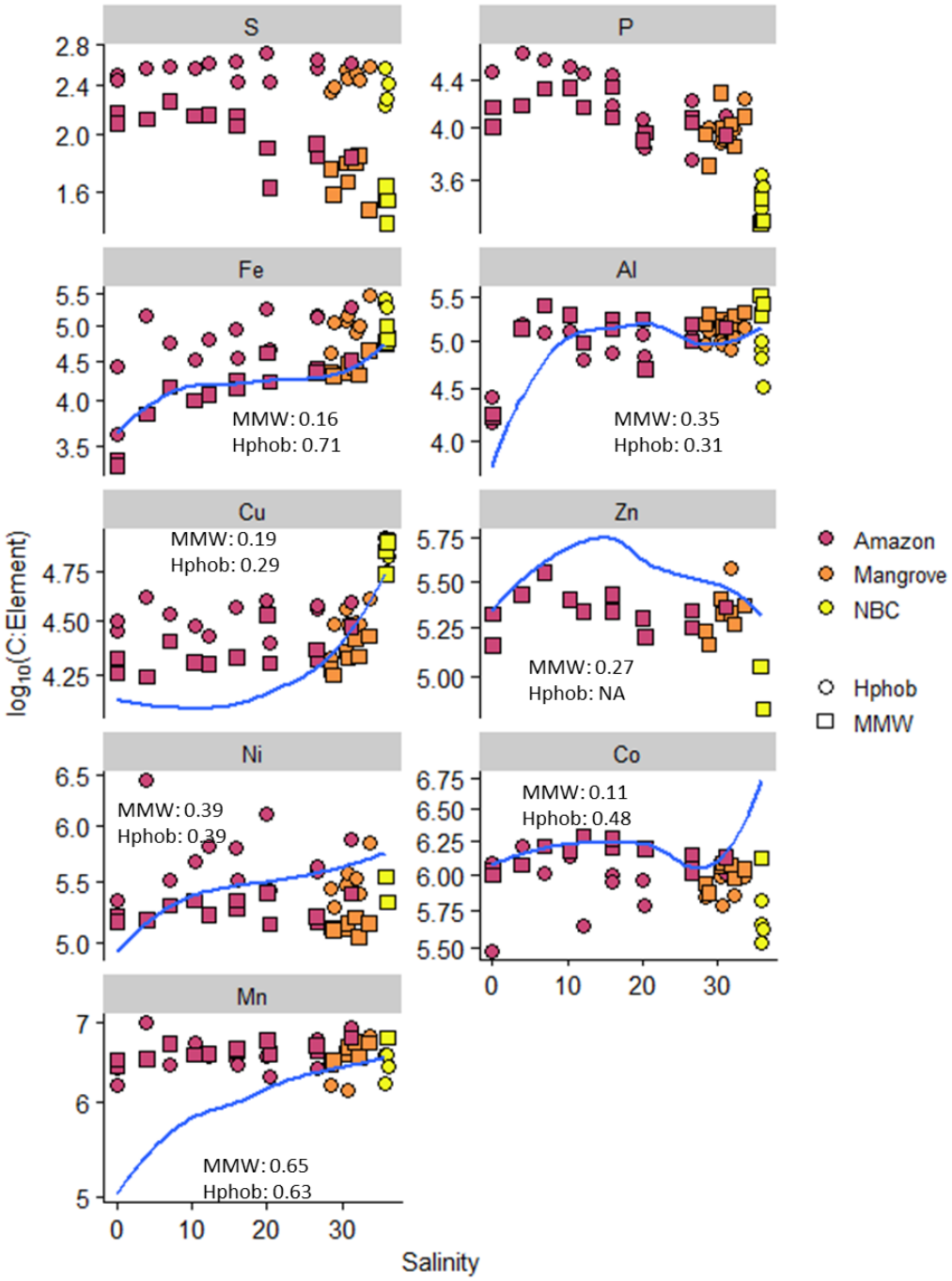

Fig. S9 Variation in observed elemental stoichiometry as a function of salinity in the study area for the MMW (squares) and the Hphob (circles) DOM fractions.

The blue line shows the trend in predicted C:TM ratios calculated from dissolved trace metal, dissolved organic carbon, ambient pH and temperature using the NICA-Donnan model assuming a fulvic acid like organic matter pool (*19*, *86*). The text shows the root mean squared error for the $\log_{10}$ transformed predicted versus observed stoichiometries in the MMW and Hphob DOM fractions. RMSE for Hphob Zn values were not calculated due to the low number of values obtained for this fraction (n=2). Colours show geographic grouping with red indicative of Amazon estuary samples, orange indicative of samples collected of an area of mangrove forest and yellow indicative of samples collected in the NBC.

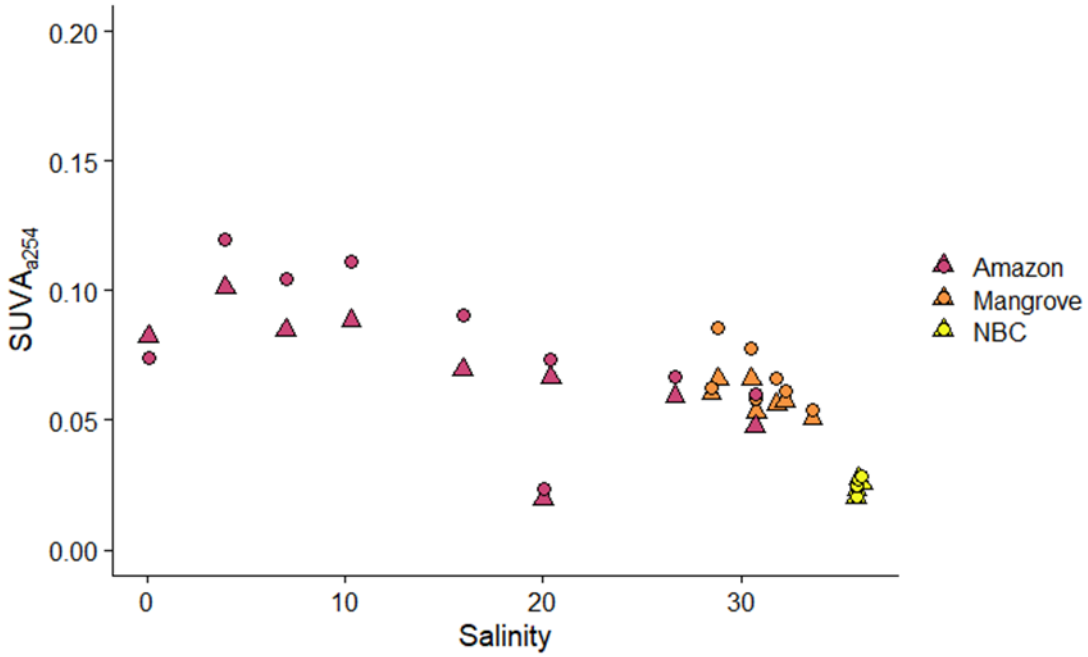

Fig. S10 Variation in specific UV absorbance at 254 nm ($SUVA_{254}$, L mg^{-1} m^{-1}) with salinity in the study area.

$SUVA_{254}$ has been shown to relate to aromaticity of dissolved organic matter (*52*) but the absolute values shown here will be different from those determined directly in aqueous samples (*52*) because these values are derived from chromatographic peak areas rather than being directly determined in seawater. Triangles show values obtained for the MMW peak and points for the Hphob peak. Differences between RPLC and SEC could relate to sample matrix as well as aromaticity of the peak. For reference, we determined $SUVA_{254}$ values for MMW peaks of 0.24±0.03 for SRHA and 0.13±0.01 for SRFA. Colours show geographic grouping with red indicative of Amazon estuary samples, orange indicative of samples collected off an area of mangrove forest and yellow indicative of samples collected in the NBC.

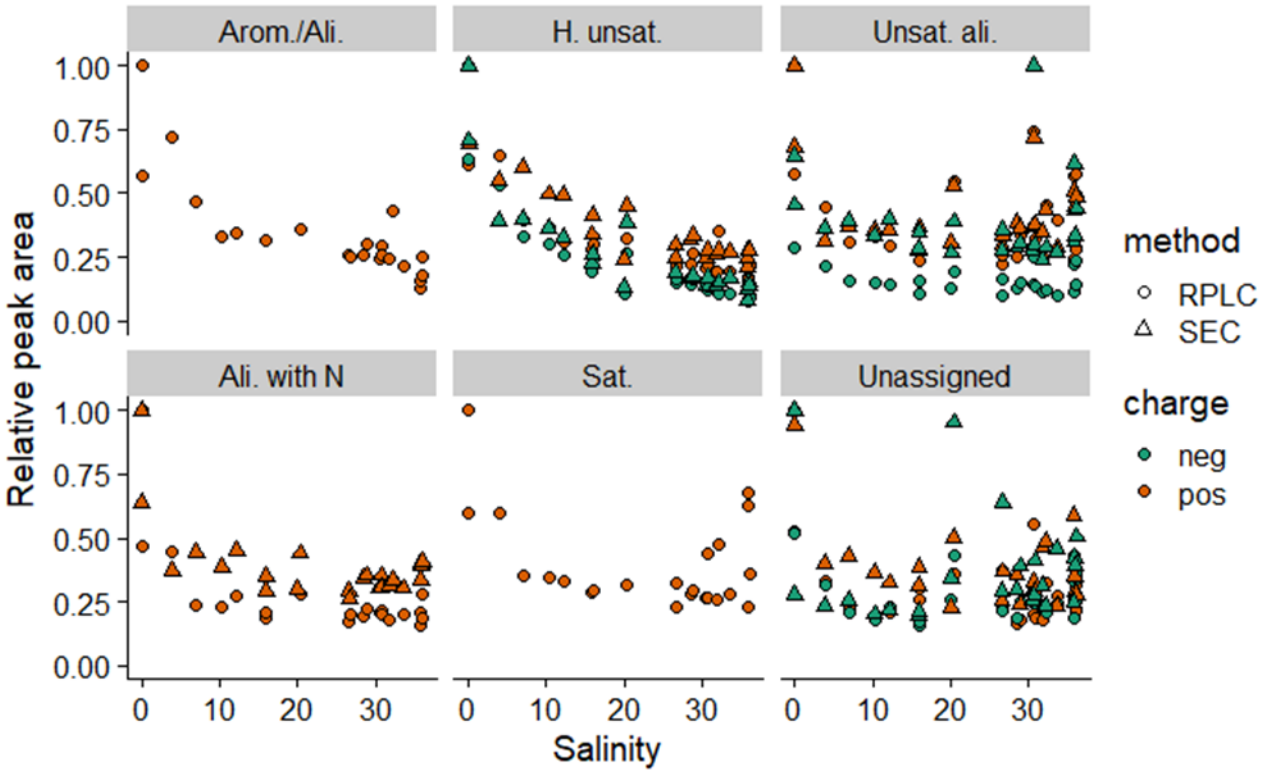

Fig. S11. Variability in the sum of areas of features with assigned or unassigned formulas grouped according to their sum formula properties (Supplementary text).

Peak areas were normalised to the maximum value. Only groups with more than 40 features per group are shown. The number of individual retention time ~ m/z features detected in all chromatographic and ionisation modes did not change markedly with salinity even though the sum of all peak areas (normalised to preconcentration volume) with peak apexes within the retention time range of the MMW and Hphob peaks decreased with salinity (Fig. 3, main manuscript). Aromatic/Ali.: 0.5<Aromaticity Index (AI)<0.67 (Mixed aromatic/aliphatic), H. unsat. : AI ≤0.5, H/C < 1.5 (Highly unsaturated), Unsat. ali.: AI ≤0.5, 1.5 ≤ H/C < 2, N < 1 (Unsaturated aliphatic), Ali. With N: AI ≤0.5, 1.5 ≤ H/C < 2, N>0 (Aliphatic compounds containing N), Sat.: AI ≤0.5, 2 ≤ H/C, O/C < 0.9 (Saturated), Unassigned: peak areas for masses where a formula could not be assigned.

Table S1. Comparison between SEC and RPLC values obtained from summing all peaks observed in either chromatographic mode.

Equations for linear regression of data in Fig. S3 according to the expression SEC (A_{254}, ion counts or nmol L^{-1}) = slope× RPLC (A_{254}, ion counts or nmol L^{-1}) + intercept. Number of observations =24.

Parameter	Intercept	slope	Adjusted r^2	p value
A_{254}	8.8±9.7	0.78±0.06	0.88	<0.01
EICpos	4.3E05±0.8E5	2.7±0.6	0.49	<0.01
EICneg	2.0E5±0.4E5	1.1±0.2	0.67	<0.01
Al	-0.002±0.088	1.0±0.1	0.72	<0.01
Co	0.006±0.002	0.21±0.03	0.69	<0.01
Cu	0.009±0.111	1.4±0.2	0.70	<0.01
Fe	2.1±1.8	1.9±0.9	0.14	0.04
Mn	0.005±0.002	1.36±0.23	0.61	<0.01
Ni	0.064±0.020	0.52±0.25	0.08	0.11
P	1.29±0.43	0.48±0.25	0.11	0.06
S	54±32	1.5±0.3	0.50	<0.01
Zn	0.023±0.011	1.0±0.17	0.46	<0.01

Table S2. Percentage of elements removed at low salinity.

Values were determined from the percentage difference between the concentrations observed at salinities between 5 and 15 and those expected assuming conservative mixing between the riverine end-member and NBC waters.

Element	Total $DOM_{SPE\text{-}amb}$ concentrations		MMW peak fraction	Hphob peak fraction
C	49±3			
	SEC	RPLC		
A_{254}	47±1	49±14	47±1	41±19
Al	93±4	90±2	92±4	90±3
Co	62±5	70±14	62±5	74±14
Cu	46±8	41±8	46±8	39±7
Fe	94±1	94±3	89±2	94±2
Mn	65±1	68±10	50±9	72±9
Ni	52±6	68±10	52±6	67±13
P	68±11	63±3	71±5	
S	56±6	42±30	56±5	50±11
Zn	43±20	-	54±10	
EICpos	48±6	35±13	48±5	35±14
EICneg	53±3	52±24	54±3	53±24

Table S3. Average retention times for peaks observed on size exclusion and reversed phase chromatography of $DOM_{SPE\text{-}amb}$ obtained from the study area.

	Size exclusion chromatography		Reversed phase chromatography	
Parameter/ Element	Peak identity	Retention time	Peak Identity	Retention time
Absorbance at 254 nm	HMW	10.9±0.9	Hydrophilic	2.2±0.1
Al		11.9±0.6		1.6±0.1
Co				2.22±0.2
Cu		11.8±0.6		2.24±0.3
Fe		11.7±1.0		1.9±0.1
Mn		11.7±0.9		2.3±0.2
Ni				2.1±0.2
P		12.1±0.1		1.9±0.1
S				1.8±0.1
Zn				2.1±0.3
A_{254}	MMW	24.7±0.5	Hydrophobic	5.0±0.8
EICpos		23.8±0.4		4.7±0.2
EICneg		23.9±0.8		5.25±1.1
Al		23.4±0.9		5.0±0.2
Co		24.3±0.2		3.5±0.4
Cu		24.3±0.3		5.7±0.7
Fe		23.7±1.0		5.3±0.3
Mn		24.4±0.2		3.3±0.3
Ni		24.2±0.4		4.1±0.5
P		24.2±0.1		4.7±0.6
S		23.3±0.3		5.2±0.3
Zn		23.9±1.0		4.3±1.3
Mn	LMW	28.9±0.2		
Zn		28.5±0.5		

Table S4. Retention times and elemental concentrations for Suwannee River Humic and Fulvic acid reference materials determined in our study.

	Fulvic Acid (1S101F)			Humic acid (3S101H)		
Element	Retention time	Concentration (μmol g^{-1})	C:Element (mol mol^{-1})	Retention time	Concentration (μmol g^{-1})	C:Element (mol mol^{-1})
A_{254}	22.2±0.1			22.2±0.1		
S	22.0±0.1	460 ±180	80±31	22.9±0.0	697±20	55±2
P	25.2±2.1	7.9±3.9	4700±2300	23.7±0.1	31±15	1250±600
Al	21.9±0.1	2.4±0.1	$15±1×10^3$	21.7±0.2	4.9±0.2	$7.9±0.3×10^3$
Fe	21.9±0.2	0.45±0.22	$82±40×10^3$	21.9±0.3	0.8±0.1	$48±6×10^3$
Cu	22.2±0.1	7.2±2.9	$5±2×10^3$	22.1±0.2	35±2	$1.1±0.1×10^3$
Zn	21.9±0.2	2.6±1.3	$14±7×10^3$	22.0±0.3	4.6±0.4	$8.4±0.7×10^3$
Ni	22.2±0.3	1.2±0.2	$31±5×10^3$	22.1±0.5	3.6±0.3	$11±1×10^3$
Co	22.1±0.1	0.14±0.04	$2.6±0.8×10^5$	22.1±0.3	0.4±0.1	$1.0±0.2×10^5$
Mn	22.9±0.1	0.03±0.02	$12±8×10^5$	22.1±0.1	0.05±0.01	$7.8±1.6×10^5$

Table S5. Molecular characteristics of ESI-MS peaks identified after non-targeted feature detection.

Features common to SEC and RPLC were identified using mzmine (*80*) as neutral masses with < 3 ppm difference assuming ionisation occurs via loss or gain of a proton. Masses were assigned based on the average m/z ratio calculated from all samples for features identified in MZmine using the R package MFAssignR (*81*).

	Size exclusion chromatography			Reversed phase chromatography		
	Peak identity	Positive	Negative	Peak Identity	Positive	Negative
Total number of features	Overall value	3572	1274	Overall value	1695	950
Mean number of C per feature ±1σ		19.5±5.8	20.4±5.2		18.4±5.3	20±5.3
Mean number of H per feature ±1σ		28.9±9.3	28.0±7.9		26.6±8.6	25.3±8.3
Mean number of O per feature ±1σ		6.1±2.9	8.9±2.2		5.8±2.9	8.1±2.2
Mean number of N per feature ±1σ		0.94±0.93	0.2±0.6		1.1±1.0	0.2±0.5
Features common to both SEC and RPLC	2417					
Features per peak	LMW (RT>28 min)	144	49	Hydrophilic (RT<3 min)	28	79
Unassigned features		50	46		15	74
AI ≥ 0.67 (PCAs)		6	1		0	1
0.5<AI<0.67 (Mixed aromatic/aliphatic)		2	0		0	0
AI ≤0.5, H/C < 1.5 (Highly unsaturated)		12	2		2	2
AI ≤0.5, 1.5 ≤ H/C < 2, N < 1 (Unsaturated aliphatic)		2	0		1	0
AI ≤0.5, 2 ≤ H/C, O/C < 0.9 (Saturated)		10	0		0	0
AI ≤0.5, 2 ≤ H/C, O/C ≥ 0.9 (Oxygen rich)		0	0		0	0
AI ≤0.5, 1.5 ≤ H/C < 2, N>0 (Aliphatic compounds containing N)		18	0		10	2
Mean m/z±1σ		371±157	310±102		307±189	378±148
Slope × 10^{-4} (Mass defect ~ nominal mass)		5.9±0.5	Not significant		3.1±0.6	3.5±0.8
Features per peak	MMW (15<RT<28 min)	3426	1225	Hydrophobic (3<RT<8 min)	1667	871
Unassigned features		595	478		244	217
AI ≥ 0.67 (PCAs)		5	6		19	1
0.5<AI<0.67 (Mixed aromatic/aliphatic)		32	6		41	3
AI ≤0.5, H/C < 1.5 (Highly unsaturated)		1428	495		806	528

AI ≤0.5, 1.5 ≤ H/C < 2, N < 1 (Unsaturated aliphatic)		379	198		85	101
AI ≤0.5, 2 ≤ H/C, O/C < 0.9, N = 0 (Saturated)		54	3		42	2
AI ≤0.5, 2 ≤ H/C, O/C ≥ 0.9 (Oxygen rich)		0	0		0	0
AI ≤0.5, H/C ≥ 1.5, N>0 (Aliphatic compounds containing N)		933	39		447	19
Features common to both +ve and -ve modes		168			162	
Mean m/z±1σ		393±128	424±132		374±122	378±110
Slope × 10^{-4} (Mass defect ~ nominal mass)		5.2±0.1	3.2±0.2		5.8±0.1	5.1±0.2
Features per peak	HMW (8<RT<15 min)	2	0			

Table S6. Absorption of inorganic TMs onto SPE cartridges.

Values were obtained after preconcentration of artificial seawater (organic free) containing 100 nmol L^{-1} of each TM. Values represent concentrations of TM observed at the same retention times as that observed for peaks in samples and are expressed as the percentage of the total TM concentration added to the media. Estimated recoveries for TMs in samples are provided for reference.

Trace metal	Percentage inorganic TM recovered RPLC	Percentage inorganic TM recovered SEC	Percentage recovery of dissolved TM estimated for samples RPLC	Percentage recovery of dissolved TM estimated for samples SEC
Fe	0.32±0.16	0.54±0.22	2.3±1.6	10±5
Cu	0.06±0.01	0.08±0.02	7.9±5.6	10±6
Al	0.07±0.01	0.02±0.02	0.3±0.2	0.4±0.8
Ni	0.023±0.004	n.d.	1.3±0.6	2±0.7
Co	0.012±0.004	n.d.	20±25	5±2
Zn	0.006±0.001	n.d.	8.5±5.1	14±11
Mn	0.002±0.001	n.d.	0.13±0.09	0.34±0.34

Table S7. First hydrolysis and NICA parameters (log*k*, *n*) used for calculations in this study.
FA1 represents the carboxylic like binding site group and FA2 the phenolic like binding site group.

Binding site type	Trace metal ion	Non ideality constant (*n*)	First hydrolysis constant ($logK_{OH}$) or NICA affinity constant ($logK$) used for TMs.
OH^-	Al^{+3}		-5
OH^-	Co^{+2}		-9.7
OH^-	Cu^{+2}		-7.5
OH^-	Fe^{+3}		-2
OH^-	Mn^{+2}		-10.6
OH^-	Ni^{+2}		-9.9
OH^-	Zn^{+2}		-9
FA1	Al^{+3}	0.42	1.64
FA1	Co^{+2}	0.67	-1.4
FA1	Cu^{+2}	0.63	0.26
FA1	Fe^{+3}	0.3	3.6
FA1	Mn^{+2}	0.72	-1.55
FA1	Ni^{+2}	0.68	-1.51
FA1	Zn^{+2}	0.64	-1.13
FA2	Al^{+3}	0.36	7.26
FA2	Co^{+2}	0.5	1.26
FA2	Cu^{+2}	0.36	7.26
FA2	Fe^{+3}	0.15	11.2
FA2	Mn^{+2}	0.56	-1.10
FA2	Ni^{+2}	0.53	0.94
FA2	Zn^{+2}	0.48	2.47

Description of data associated with manuscript.

Files are available separately as supplementary data files associated with the manuscript or can be downloaded from
https://doi.pangaea.de/10.1594/PANGAEA.941455**:**

Data set components:

1) Water column parameters:
Table containing water column data from stations sampled for this study.
2) SEC data:
Table containing data obtained from size exclusion chromatography
a) Concentrations of elements S, P, Fe, Al, Cu, Zn, Ni, Mn and Co observed in samples obtained as detected by ICP-MS.
b) UV absorbance (254 nm) per litre of seawater observed with a diode array detector
c) Total ion counts per litre of seawater obtained by summing peak area of all ions with retention times coeluting with UV 254 and element peaks obtained on ESI-MS analysis in positive and negative ion modes.
3) RPLC data:
Table containing data obtained from reversed phase chromatography
a) Concentrations of elements S, P, Fe, Al, Cu, Zn, Ni, Mn and Co observed in samples obtained as detected by ICP-MS.
b) UV absorbance (254 nm) per litre of seawater observed with a diode array detector
c) Total ion counts per litre of seawater obtained by summing peak area of all ions with retention times coeluting with UV 254 and element peaks obtained on ESI-MS analysis in positive and negative ion modes.
4-7) Feature lists obtained for ESI-MS analysis for size exclusion and reversed phase chromatography in positive and negative ion modes
8-9) List of ions used to exclude acetate and plasticizer ions from the feature lists.
10) Input data for speciation calculations. This data can be copied and pasted into the input field in the speciation program ORCHESTRA.
11) Output of speciation calculations obtained by ORCHESTRA.