Response to referee comment by Peter Kraal (Referee #1) on “Flocculation of dissolved organic matter controls the distribution of iron in boreal estuarine sediments” by Tom Jilbert et al.

[Referee comments in bold]

[Responses in italics]

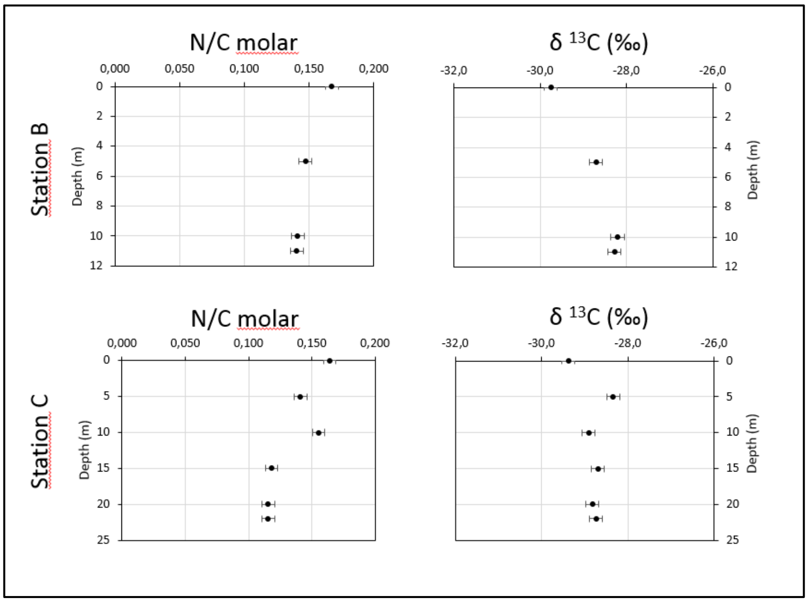
With interest I have read this manuscript, in which the authors explore the role of salinity-driven flocculation of DOM and dFe in controlling the settling and diagenetic fate of riverine Fe along a salinity transect in a Baltic Sea estuary. The authors suggest a key role for flocculation in transferring Fe(III) in the form of Fe(III)-OM complexes and Fe(III) (oxyhydr)oxides from the water column to the sediment. Moreover, the Fe-OM pool is relatively stable and this Fe does not seem to participate in the “normal” reductive diagenetic pathways in organic-rich sediments. The manuscript is well-written, well-structured and an interesting addition to the flourishing research field of Fe-OM interactions in marine and terrestrial systems.

Together with this review, I have uploaded an annotated pdf document with all my questions and comments. Below, I highlight the main questions that arose while reading the manuscript.

We thank the referee for the careful consideration of our manuscript. Below, we respond to each of the major comments followed by the line-by-line comments extracted from the referee’s annotated pdf.

1. On a technical note: If I understand correctly, relative errors for all solid-phase analyses were calculated from replicate analysis of “regular” (powdered) reference materials. I wonder whether this gives relative errors that are also representative for analysis of suspended material on filters. As far as I am aware, such samples are somewhat harder to process and I am curious to know if the authors can comment on how/whether they specifically assessed analytical precision and accuracy associated with filter samples (or why not).

For the analysis of suspended organic matter, complete filters were combusted after packing and compression in tin cups. This avoids issues of heterogeneity within individual filters and this detail will be added to the methods section. We did not take systematic replicates for all GF/F filters due to limitations of sample volume and processing time. However for each site we did take an additional water column sample from <1m above the seafloor. At the stations where this sample is close to (within 2 vertical meters of) the deepest 5m depth interval sample, we suggest it may be used as a replicate to assess precision between filters (with the added value of coming from a separate sampling cast rather than simply being a duplicate from the same Limnos bottle). When this is done we see that the $\delta^{13}C$ *and N/C values for the extra sample are within the analytical error of the sample from the deepest 5m depth interval (see Fig. R1). Hence we conclude that precision for this data is in fact limited by the analysis and not by the sampling procedure. In terms of accuracy, this can only be determined by reference materials analyzed in parallel to the samples. This was done routinely (two standards for every ten samples) and showed accuracy to be <2.5% as stated in the manuscript.*



<u>Figure R1:</u> N/C (mol) and $\delta^{13}C$ of suspended particulate matter at Stations B and C. All data presented here, except for the deepest sample at each station (taken from <1 m above the seafloor), are included in Fig. 3 of the manuscript. Here, the sample from <1 m above the seafloor is treated as a replicate for the deepest 5m sampling interval for each station (depth offset = approx. 1 m for Station B, approx. 2 m for Station C). Horizontal error bars indicate analytical precision, reported as one standard deviation, as determined by 10 repeated measurements of standard materials (N/C =0.005, $\delta^{13}C$ = 0.15 ‰).

When the corresponding exercise is carried out for $Fe_{part.}$ from the polycarbonate filters we observe relative standard deviation (RSD) of up to 15% for "replicate" deep water samples. This is substantially greater than the RSD values for $\delta^{13}C$ and N/C (both less than 1 % in Fig. R1). We interpret this as a consequence of the fact that $\delta^{13}C$ and N/C describe the characteristics of the suspended organic material, but are largely insensitive to variations in its total concentration at a given location. $Fe_{part.}$, on the other hand, describes the absolute concentration of particulate Fe in µmol/L, which (similarly to the absolute concentration of organic matter) is likely to vary on small spatial and temporal scales and may well change in the 10–15 minute interval between Limnos casts at a given location. It should however be noted that the error associated with an RSD of 15%, particularly for low-Fe samples, is several orders of magnitude less than the large-scale changes in $Fe_{part.}$ observed along the salinity gradient which are discussed in the manuscript (Mean value at Station A = 2.54 µmol/L, σ = 1.98 (n = 2); mean value at Station K = 0.30 µmol/L, σ = 0.10 (n = 10)). The concentration of $Fe_{part.}$ at Station A is thus 25 times greater than at Station K.

2. There is no information provided on which standards were analyzed with Mössbauer spectroscopy (or whether reference spectra collected previously were included) and how the selection of (number of) standards for LCF was performed. The LCF fitting routine was only explicitly mentioned in the caption to Fig. 5; it should also feature in the main methods section. The key statement "Quantification of iron-bearing phases and iron oxidation states is based on relative subspectral areas" may be expanded a bit (as it is to some extent in the notes of Table 3). Overall, the procedure of obtaining relative proportions, including that of the "undocumented" Fe phase (in my opinion an awkward term, perhaps "unknown" is more appropriate?), should be more clear, as this phase plays a rather crucial role in the manuscript. Goodness-of-fit is also an important parameter in this respect, as it is basically the (areal) mismatch between the fit and the actual spectrum that is used as a measure of "undocumented" Fe. I think it would be good if all (relevant) reference spectra (perhaps including likely candidates that were not present such as siderite) are clearly presented (they are somewhat hard to discern in the current Figure 5: perhaps a stack plot with offset would work).

We thank the referee for highlighting the need for a more detailed presentation of the Mössbauer approach. To clarify, the LCF fitting was performed using reference spectra rather than freshly prepared and analyzed standards. In the revised version, we will list the reference materials considered, present their spectra, describe how the selection for the LCF was performed, and expand the discussion of the LCF fitting procedure itself including goodness of fit. Furthermore we are in the process of generating Mössbauer spectra of additional sediment samples from the estuarine transect. When presenting these results in the revised manuscript we will consider the referee's terminological suggestions regarding the use of the terms "undocumented" vs. "unknown" Fe phases.

3. The authors assign the "undocumented" Fe phase fully to complexes of non-sulfidized Fe(II) with organic matter (p 12, L 17-19). They base this on the study of Yu et al. (2015), who found Fe(II)-OM phase "to be a major component of sedimentary Fe in a nearby boreal estuary." In the Yu et al. study, OM-complexed Fe(II) was identified by Fe X-ray absorption spectroscopy using standards of Fe(II) and Fe(III) complexed with organic matter. As far as I could see, no empirical data in support of the assumption of OM-Fe(II) is provided in this manuscript. Because the Fe(II)-OM phase plays such an important role in the discussion, I wonder whether the authors can further substantiate their assumption that all the "undocumented" Fe, that could not be assigned to their (to the reader unknown) library of standards, was present as Fe(II)-OM? It should be clear exactly which Fe phases could be ruled out based on LCF with reference spectra. Some more focus on which spectral features could not be explained by the available reference standards, and how these may point to OM-associated Fe, would also be welcome. In the absence of a "smoking" gun, perhaps some more consideration should be given to the fact that OM-Fe(II) may not necessarily account for all "undocumented" Fe (to what extent do fitting uncertainties play a role?). In particular, the authors may want to address the validity of extrapolating Mössbauer data (key for assigning Stage 3 and 4 Fe to OM-Fe(II)) for the uppermost sample (0-1 cm) to the whole sediment record (up to 60 cm) at all sites (see also Comment 7).

Our detailed response to this comment is partly dependent on the ongoing Mössbauer analysis of an additional five sediment samples from the estuarine transect. In total, the new dataset will comprise seven samples (riverbed sediment, shallow + deep sediment from Station A, shallow + deep sediment from Station D, shallow + deep sediment from Station J). At the time of writing, four of the additional 5 samples have been measured and the fifth is in progress. Analysis of these spectra will be performed by LCF using references, and as stated in the response to the previous comment, the references used in the LCF (and the rationale for their selection) will be explicitly stated in the manuscript, complete with presentation of spectra.

We have also performed additional sediment extractions on 15 samples from Stations A, D and J to investigate the robustness of our initial conclusions concerning the dominance of OM-Fe(II) in Stages 3 and 4 of the Poulton and Canfield (2005) method. These results will be included in the revised manuscript and considered alongside the new Mössbauer data. In brief, we carried out parallel extractions of subsamples dried under N_2 after frozen storage, and treated under anoxic conditions as follows:

Subsample 1: 1M HCl, 1 hr (including trap for H_2S evolved from Acid Volatile Sulfur)
Subsample 2: 1 M HCl + 1 M hydroxylamine-HCl, 1 hr
Subsample 3: 0.2 M Sodium citrate at pH 4.8, 2 hr

Fe (II) in each extract was then determined by spectrophotometry after complexation with 1,10 phenanthroline, while AVS was determined by iodometric titration. This approach is similar to that *used by Yu et al. (2015) to determine sulfidized vs. unsulfidized Fe (II), and Fe(II) /Fe(III), in the cold 1M HCl-soluble fraction of boreal estuarine sediments. However we also included Subsample 3 to mimick Stage 3 of the Poulton and Canfield extraction – but excluding dithionite – to test the hypothesis that the citrate ligand alone may be able to extract Fe (II) direct from OM-complexes. This specifically addresses a comment of Referee #2 concerning our interpretations of citrate-dithionite-soluble (i.e. nominally reducible) Fe as Fe (II).*

The results of the additional extractions suggest that our initial assumptions concerning the solubility of OM-Fe (II) in Stages 3 and 4 may indeed require modification. We do detect unsulfidized Fe(II) in the HCl extracts, and its concentration in most samples is higher than both sulfidized Fe (II) (AVS) and HCl-soluble Fe (III) (derived from the difference between Subsamples 1 and 2, see Figs. R2 and R3). Furthermore the parallel citrate-only extraction does appear to dissolve approximately 60% of the unsulfidized Fe (II) pool (see Fig. R4). While these observations support our claim that OM-Fe(II) complexes are present in the sediments and suggest that at least a fraction of this material is citrate-soluble, the total amount of Fe dissolved in the additional extracts is equivalent to (only) the combined total from Stages 1 and 2 of the Poulton and Canfield (2005) procedure (Fig. R5). Hence, the Fe phases dissolved by 1M HCl are likely identical to those dissolved by sodium acetate and hydroxylamine-HCl, while the Stage 3- (citrate-dithionite) soluble Fe fractions remain largely intact during the 1M HCl extraction.

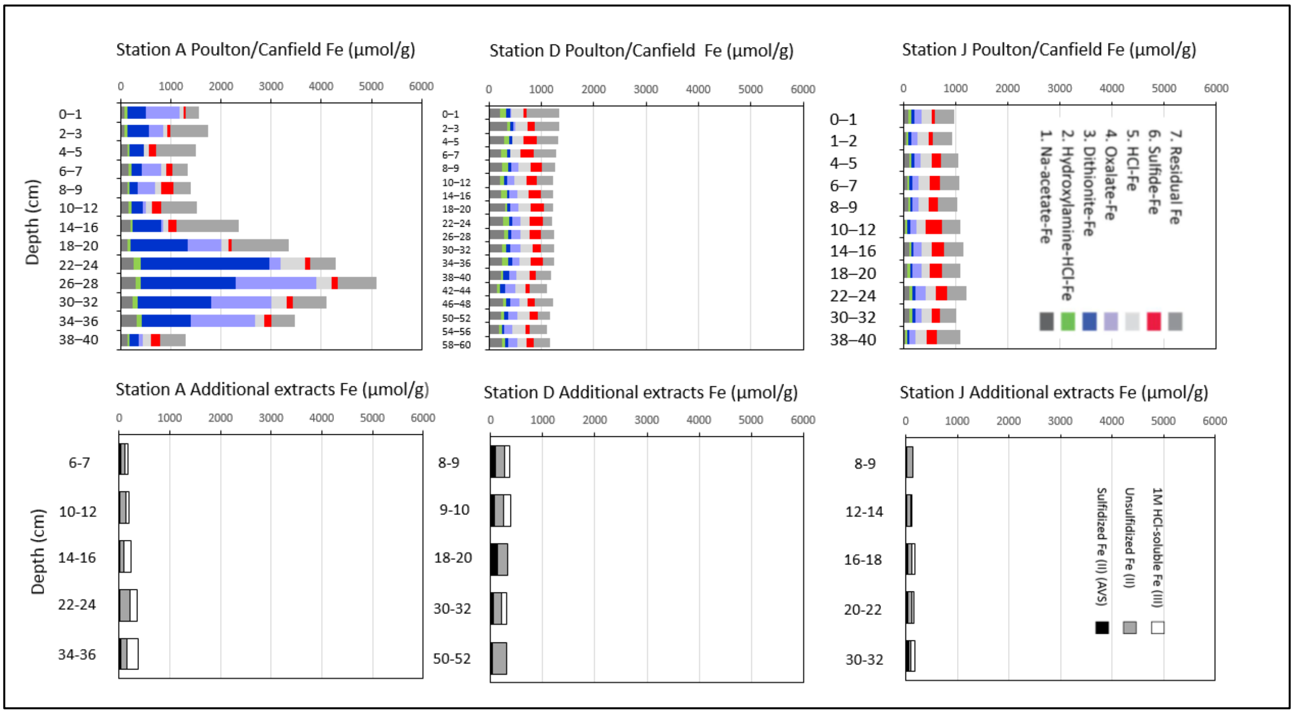
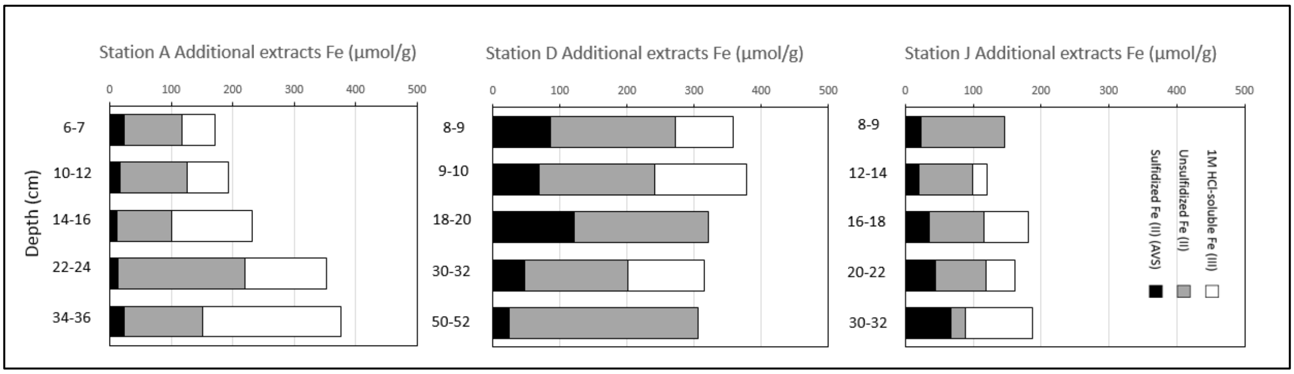
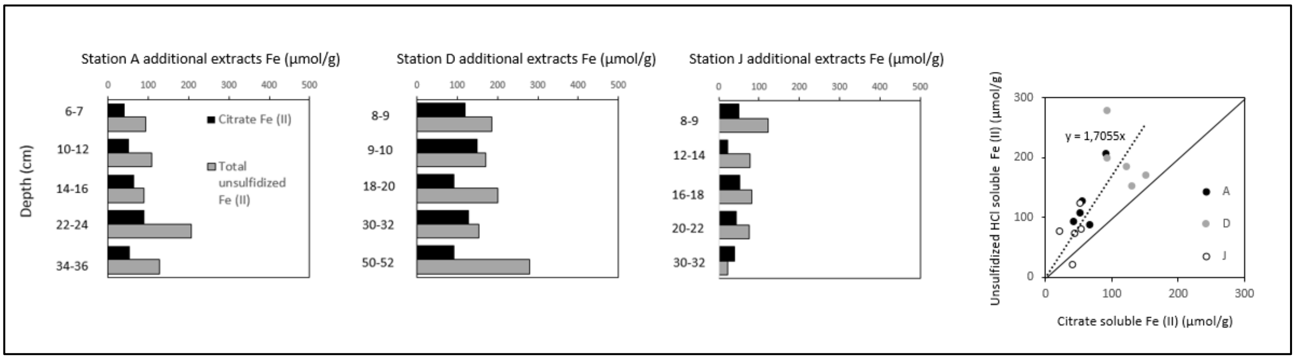


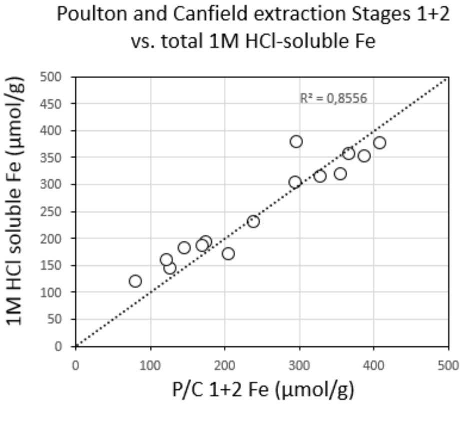
Fig. R2: Comparison of Fe extraction data for the Poulton and Canfield (2005) method reported in the original manuscript (upper panels), and additional extractions with 1M HCl performed in response to the interactive discussion comments (lower panels). All panels are plotted on the same scales for comparability. Where possible, samples for the additional extractions were selected from the set originally extracted by the Poulton and Canfield method. When no sample material was remaining, adjacent samples were taken (n=4).



<u>Fig. R3:</u> Zoom of lower panels of Fig. R2 for clarity.



<u>Fig. R4:</u> Comparison of citrate-soluble Fe (II) (Subsample 3) and total unsulfidized Fe (II) (derived from Subsample 1) for the additional extraction samples. Note that approximately 60% of total 1M HCl-soluble unsulfidized Fe (II) is extracted by the citrate solution, as given by the relationship in the final panel (solid line in this panel indicates 1:1).



<u>Fig. R5.</u> Relationship between Fe extracted in Stages 1+2 of the Poulton and Canfield method, and total 1M HCl-extractable Fe in the additional extractions. Where no equivalent sample was available, adjacent samples have been compared (n =4). Dashed line represents 1:1 and the least-squares regression is performed against this line.

Therefore, assuming the 1M HCl extraction to be a reliable determinant for OM-Fe(II) as suggested by Yu et al., 2015, we should modify our interpretation of the composition of the phases dissolved in Stages 3 and 4 of the Poulton and Canfield method. We hope that the forthcoming Mössbauer data will clarify this issue, especially the deep sample from Station A, whose combined Stage 3 and Stage 4-soluble Fe pool is in excess of 3000 µmol/g sediment (i.e. 16.8% of the sediment by weight).

4. p 14, L 13-19. The authors attribute the excess removal of Fediss relative to DOC to "preferred association of Fe with higher molecular weight compounds, which are more sensitive to flocculation (Asmala et al., 2014) or a mechanistic enhancement of flocculation by the presence of Fe (Forsgren et al., 1996)." Firstly, the second mechanism deserves some explanation (it now implies enhancement of flocculation of Fe by Fe?). Secondly, as the authors also observe a ferrihydrite signal in their Fe flux to the sediment, could flocculation of Fe independent of DOM -> POM dynamics play a role in the removal of Fe from the water column (driven by salinity, pH, perhaps DO)? The authors mention that "Flocculated material in the oxic estuarine water column is likely present as Fe (III) partitioned between organic-Fe (III) complexes and ferrihydrite (Neubauer et al., 2013)." (p 16, L 20-21). Changes in surface charge of Fe(III) particles due to adsorption of ions at higher salinity may also affect the solubility/flocculation of Fe(III) particles?

We will clarify these statements in the revised manuscript. Several authors have noted that the respective behavior of Fe and bulk OM along salinity gradients in Baltic Sea estuaries differ significantly, with non-conservative behavior being much more apparent for Fe. This may indeed imply additional mechanisms for flocculation that are specific to Fe, rather than simply to the fraction of OM to which Fe is associated (the context within which this statement was written). These include, as the referee suggests, direct flocculation of Fe oxides in response to pH changes. Typically, pH of the the Mustionjoki river drainage is in the order 6.0–6.5 (Lahermo et al., Geochemical Atlas of Finland, 1996). Along the salinity gradient in the surface waters of the estuary, pH indeed rises, towards the offshore value of 8.0–8.4 in the open Gulf of Finland (exact value depends on season, e.g., Omstedt et al., Tellus B, 62B, 2010), due to mixing between fresh and brackish water masses. As outlined in Neubauer et al., ES&T 47, 2013, pH – independent of salinity – may determine the partitioning of Fe between NOM-Fe complexes and ferrihydrite, and between different size classes of ferrihydrite in natural waters (their Figure 1e). In their model, freshwater at pH 6 should already contain a substantial fraction of Fe oxides that would not pass through a 0.2µm filter. We used 0.4µm filters, but nevertheless measured a majority of Fe in the particulate phase in the freshwater endmember sample (surface water of Station A, Fig. 2) which supports the Neubauer et al. model for a pH 6.0–6.5 river. However we see an immediate loss of particulate Fe between Stations A and B, despite the onset of the salinity gradient and hence the rise in pH. This suggests that also current strength influences the concentration of this material in the water column, due to its susceptibility to sedimentation. Further offshore, we observe a second maximum in particulate Fe (Fig. 2) which we attribute to flocculated formerly dissolved material. We cannot discount that pH, increasing in tandem with salinity, influences this process. The density of minerogenic matter such as clays may also influence the likelihood of flocculation in boreal estuaries (as discussed in Forsgren et al., 1996).

5. The reasoning behind the conclusion "that flocculation of DOM to POM in the estuarine environment may provide the second fraction of POM detected in the N/CPOM and 13CPOM data." is somewhat unclear to this non-expert in that field (Fig. 3 and 4 and section 5.3). The authors state that the POM signal in surface waters is dominated by phytoplankton and therefore DOM-POM transitions cannot be discerned in N/CPOM vs 13CPOM signature of surface water-DOM from a previous study plot in the field of C3 plant material, suggesting that phytoplankton is not an important component of surface water DOM? I found this in itself

surprising, as I would expect the DOM to be impacted by decomposition (POM -> DOM) of fresh algal material.

To answer this point first: It is true that degradation of phytoplankton material may theoretically impact on DOM content (and therefore net composition). However in this system the mass balance argues against such a hypothesis. Concentrations of river-derived DOM in the estuary are an order of magnitude greater than the potential supply from degrading phytoplankton. In the study of Asmala et al., Biogeosciences 10, 2013, DOM concentrations along the salinity transect of the Mustionjoki estuary declined from 585 µmol/L in river water to 363 µmol/L in the offshore endmember (average of three seasons). As shown in the same study (and in our Fig. 3), despite transformations in the estuarine environment including flocculation and utilization of DOM in foodwebs, bulk DOM retains a strongly "terrestrial" C/N (or (N/C) ratio even in the offshore region. In this study, we measured POC in the range of 20–60 µmol/L for most stations (see Fig. R7 and convert to molar units). Hence, complete degradation of this material to DOM would yield only minor changes in the C/N (or (N/C) ratio. We will rephrase the text to emphasize this.

The POM data from deeper waters from this study plot towards the C3 plant signature. The authors use this trend as an indication that the second OM source besides phytoplankton, i.e. C3 plant material, is transferred from DOM to POM through flocculation. I wonder whether the POM N/CPOM vs 13CPOM plots cannot also be determined/impacted by magnification of the C3 signal in the POM reservoir below the phytoplankton-dominated surface waters? I guess this strongly depends on the "rapid remineralization of fresh phytoplankton material during settling" (p 15, L8-9) which would remove the phytoplankton signal from POM and DOM and the persistence of the phytoplankton POM signal.

Our conceptual model for interpreting the plots in Fig. 3 is that there are essentially two end-members of POM in the estuarine water column: phytoplankton and (C3) terrestrial plant material. At any given location (water depth, distance from river mouth), we then observe a mixture of phytoplankton and terrestrial material whose proportions determine the bulk N/C_{POM} and $\delta^{13}C_{POM}$. In deeper waters, these bulk signals are closer to those of the plant material, partly because remineralization of sinking phytoplankton detritus depletes the contribution of this end-member. If we understand the referee's question correctly, it concerns the possibility that plant material in the bulk POM may derive from a direct POM input to the estuary (rather than from flocculated DOM). This is indeed a possibility which cannot be ruled out with the data presented. However as demonstrated by Mattsson et al., Biogeochemistry 76, 2005 in a study of 86 Finnish river catchments, 94% of TOC in river water is present as DOC. This implies that transformations of DOC in the estuarine water column are a more likely source of plant matter-derived POC than a direct POC input from the catchment. Furthermore DOM in the Mustionjoki estuary is known to undergo non-conservative mixing at low salinities (considered mainly to be due to flocculation of DOC to POC as outlined in Asmala et al., JGR Biogeosciences 10.1002, 2014). Hence we stand by the initial interpretation that the terrestrial plant material end-member of POM is likely derived from flocculated DOM.

6. p16, L18-27. The authors interpret dithionite- and oxalate-extractable Fe as OMFe(II), and state that this is formed in the water column rather than the sediment. As far as I can see, this is again based on the findings of Yu et al. (2015) for a nearby boreal estuary. It may be good if the authors highlight data from this study or further literature besides Yu et al. (2015) in support of the conclusion that OM-Fe(II) is exclusively formed in the water column and is not of diagenetic origin.

As outlined in the response to comment 3, our interpretations concerning the role of OM-Fe(II) will require modification. The additional 1M HCl extraction data confirm that unsulfidized Fe (II) is present in all measured samples (n=15) but the comparison in Fig. R5 shows that it likely contributes to Stages 1+2-soluble Fe (rather than Stages 3+4-soluble Fe) in the Poulton and Canfield method. In fact we did not claim that the OM-Fe(II) complexes were formed in the water column, rather that the association between Fe and OM may initially occur in the water column (where Fe is present as Fe(III)) and persist in the sediments (where Fe is then reduced to Fe(II)). In any case we will rewrite this section based on the new information from the additional extractions and Mössbauer analyses.

7. p 17, L 4-7. The maximum accumulation of ferrihydrite in a (seasonally oxygen depleted) "pit" of Station D is interesting, in that redox shuttling apparently causes maximum accumulation Fh (the most labile and easily reduced Fe(III) phase) in the surface sediment (Fig. 6), where H2S already seems to accumulate (Fig. 7). Striking is also the persistence of Stage 1 and 2 Fe ("ferrihydrite") with depth through sulfidic depth intervals at Station D and to a lesser extent Station J. Could ferrihydrite perhaps be only part of the answer? The authors assign all Fe extracted by Na acetate at pH 4.5 to ferrihydrite, while this mineral is very slow to dissolve at that pH. Iron monosulfide was not specifically quantified in this study, could the presence of FeS help explain the Na acetate-extractable Fe pool and the persistence of Stage 1 and 2 Fe with depth? Egger et al. 2014, ES&T, mention that FeS in Baltic surface sediments is extracted in Stage 1 and especially Stage 2 of the same Fe extraction scheme as used in this study.

To answer this point first: As outlined in the response to comment 3 we now have a more detailed picture of the likely composition of the Na-acetate- and hydroxylamine-HCl-soluble fractions, which helps to answer these questions. First we recall that total 1M HCl-soluble Fe is equivalent to total Stage 1+2-soluble Fe from the Poulton and Canfield method (Fig. R5), implying that the same phases are dissolved in both approaches. The results of the 1M HCl extractions therefore confirm the referee's suggestion that FeS accounts for a portion of total Na-acetate- and/or hydroxylamine-HCl-soluble Fe (Fig. R3). The remainder is made up of unsulfidized Fe (II) and HCl-soluble Fe (III). It is notable, as the referee also states, that Station D is characterized by the highest total proportions of these "labile" phases (Figs. R2 and R3) which is very likely a consequence of lateral Fe shuttling into the bathymetric depression, followed by diagenetic cycling of Fe in the sediment. This diagenetic cycling may include reduction of ferrihydrite and precipitation as FeS.

Although the additional extractions were performed on only 5 samples per station – and therefore trends with depth in the relative proportions of FeS, unsulfidized Fe(II) and HCl-soluble Fe(III) are difficult to discern – the data suggests that HCl-soluble Fe (III) is indeed present in most samples. This implies a persistent occurrence of ferrihydrite at depth in the sediments throughout the transect, as originally suggested in the manuscript. An alternative theory would be that the HCl-soluble Fe (III) is contributed by labile OM-Fe(III) complexes, which may be less susceptible than ferrihydrite to reaction with H_2S. These possibilities will be discussed in more detail in the revised version. (We note that three of the 15 samples appear not to yield any HCl-soluble Fe(III) (Fig. R3), but that this may be an artefact caused by heterogeneity between Subsamples 1 and 2, which were used to estimate HCl-soluble Fe (III) from the difference between HCl-soluble Fe (II) and HCl-soluble $Fe_{tot.}$).

The Mössbauer data indicate abundant superparamagnetic Fe(III) at Stations A and to a lesser extent Station D, but these are data from the uppermost sample (0-1 cm) that was in contact with bottom water O2 that was still 75 % (Station A) to 30 % (Station D) of saturation, eyeballing Fig. 1. FeS persistence far below the zone of H2S accumulation may be unlikely (unless the system is very dynamic and far from any steady state). This zone, characterized by abundant Fe2+ but no sulfide, is often associated with formation of reduced Fe minerals such

as vivianite and siderite. Could these also be present here and extracted in the low pH Stage 1 and 2 of the extraction scheme? The fact that these phases were not observed in the Mössbauer data from the surface sample does not rule out their presence at depth. Overall, some more words (beyond "seasonally oxygen depleted") can be spent regarding the steadiness of a geochemical state with apparently co-occurring "Fh" and H2S accumulation, and the nature of Fe extracted from anoxic non-sulfidic and sulfidic sediments during Stages 1 and 2 of the Fe extraction scheme.

We are currently analyzing five additional samples by Mössbauer, including one sample from the deeper part of the core from Station D (26–28 cm), below the H_2S *zone. As indicated above, these data will be included in the revised version of the manuscript along with the relevant reference spectra and expanded details of the LCF procedure. The referee is correct that precipitation of Fe(II) phosphates and carbonates is theoretically possible in this depth interval, as shown for example for Bothnian Sea sediments in Egger et al., GCA 169, 2015. We will acknowledge this in the text as well considering these phases in the interpretation of the Mössbauer data. It is also indeed noteworthy that Station D appears to display co-occurrence of "ferrihydrite" and* H_2S *in the depth interval 5–20 cm, whereas ferrihydrite is expected to react with* H_2S *with a half-life measured in hours (Table 4.1 in Raiswell and Canfield, Geochemical Perspectives 1, 2012). Although we have only sampled Station D pore waters on one occasion, we expect from nearby – repeatedly sampled – stations that the* H_2S *peak will be persistent throughout the year, which supports the need for an explanation why "ferrihydrite" should survive in the sediments. As outlined in the previous response, one possibility is that the HCl-soluble Fe(III) is in fact labile OM-Fe(III), which may be less susceptible than ferrihydrite to reaction with* H_2S*. Another is that ferrihydrite is indeed present, but somehow protected from reaction with* H_2S *by surface sorption processes. We note the referee's own recent conference abstract (Kraal et al., Goldschmidt 2017) which suggests a greater chemical stability for ferrihydrite particles to which P and Si is sorbed.*

I look forward to enjoying the revised version of this manuscript.

Kind regards, Peter Kraal Utrecht University, Department of Earth Sciences-Geochemistry Please also note the supplement to this comment: https://www.biogeosciences-discuss.net/bg-2017-181/bg-2017-181-RC1-supplement.pdf

Responses to Line by Line supplementary comments from P. Kraal (Referee #1)

P3 Line 6: And, conversely, OM has been shown to stabilize ferrous Fe (Toner et al, Nat Geosci 2, 2009)

We will add the reference to the revised manuscript.

P3 Line 13: Michel et al, Science 316, 2007; Hiemstra et al, GCA 105, 2013: $Fe_{10}O_{14}OH_2.nH_2O$

We will clarify that there are several formulas for ferrihydrite (and indeed several models for its structure) in the literature. The simplified formula we use is taken from Raiswell, Elements 2011.

P3 Line 25: Fh binding sites are overwhelmingly Fe-O groups, not Fe-OH

We will clarify this statement with references to developments in understanding of the structure and sorption characteristics of ferrihydrite. Relevant to the referee's comment, the original OH-rich model for ferrihydrite presented by Drits et al. (Clay Minerals, 28, 1993) has indeed been superseded by more recent works implying a dominance of Fe-O bonds in the structure and thus in sorption sites on the mineral surface, as outlined in Hiemstra, GCA 105, 2013.

P3 Line 29: net

The text will be modified accordingly.

P5 Line 28: I can imagine that the precision and accuracy of standards may differ from data obtained from solids on filters. Were complete filters combusted, or parts? Were there any replicates for filter analyses?

This issue is addressed in detail in our response to the referee's main comments.

P6 Line 9: Here I have the same question as above, about the relationship between results from standards and results from (pieces of) filters.

This issue is addressed in detail in our response to the referee's main comments.

P6 line 10: I do not follow: if total Fe on filter is measured, and the filtered volume is known, why is TSS needed to convert from umol Fe on filter to umol Fe/L water?

The referee is correct that TSS is not needed to perform this calculation. In our spreadsheets we had used TSS first to estimate the concentration of $Fe_{part.}$ *as µmol/g suspended solid material, and subsequently converted this value back to* $Fe_{part.}$ *in µmol/L. We now checked the direct conversion as described by the referee and the results are consistent with those initially calculated. We will reword this section.*

P8 Line 15: And, in light of the possible role of FeS, maybe spell out that it's a low estimate for sulfide-associated Fe.

This issue is addressed in detail in our response to the referee's main comments.

P8 Line 21: No info on standards.

This issue is addressed in detail in our response to the referee's main comments.

P10 Line 4: Based on measurements of gravimetric water loss?

The assumed constant value of 90% is of course a simplification of the true porosity profile, but sufficient for the qualitative discussion of diagenetic processes presented in the original paper. For completeness, we will adjust the CH_4 *concentration data using a best-fit line through the porosity data derived from gravimetric water loss, which are indeed available.*

P12 Line 8: Are the stacked 100% plots necessary?

It is difficult to convey clearly all the information in this large dataset (multiple operational fractions + multiple stations, Fe + organic matter dynamics) in a single figure. We decided to include the *stacked 100% plots to improve our chances of succeeding in this. For example in the case of the plot of the organic matter fractions, the absolute concentrations show clearly that there are lower values in the stations close to the sill, while the stacked 100% plot shows that the relative concentrations of terrestrial and phytoplankton material are unaffected by the presence of the sill (i.e. the samples plot approximately where expected in the offshore trend).*

P12 Line 8: Well, station B stands out but for the rest I do not see much of a (general) trend.

We will adjust the text accordingly.

P12 Line 14: The "marked" difference does not become apparent, because the phases normally associated with the stages are not mentioned here. For clarity, I would refer to the operational fractions above (L12, 13) as it is the first mention of the "stages" in the results section (instead of later on in L25 and 27), mention the corresponding stages and then keep referring to the stages.

This issue is addressed in detail in our response to the referee's main comments.

P12 Line 19: Any way to substantiate this? Yu et al used XAS, here it's Mossbauer without any Fe-OM reference phase?

This issue is addressed in detail in our response to the referee's main comments.

P12 Line 25: Consider the order of section 4.5 and the place of this paragraph; would it not fit better before the Mossbauer results?

We will consider this advice when restructuring the manuscript.

P13 Line 15: Why H2S in the text, when HS- (the major species at circumneutral pH) is used in the formula?

We will correct this inconsistency.

P13 Line 18: And why H2S here, while HS- in Eq 5?

We will correct this inconsistency.

P13 Line 20: What does "efficiently" mean in this context? Rapid reaction?

Yes, we will clarify this.

P13 Line 21: vertical extent

We will adjust the text.

P14 Line 6: Why is there no direct evidence for OM flocculation for this study? Fig 3b presents data, but does not show DOM/POM trends along the gradient.

This is because the original experimental design was focused on Fe and POM did not include measurements of DOM in 2 dimensions June 2015. The DOM data were added later to help explain the distribution of N/C_{POM} *and* $\delta^{13}C_{POM}$.

P14 Line 10: I find it hard to understand this remark and its context, some more detail would be beneficial. What kind of variations in end-member values?

We will expand this section briefly. The statement refers to the fact that the freshwater DOM endmember characteristics (and magnitude) are temporally variable, for example in response to the seasonal cycle or discharge events. In estuaries with a long freshwater residence time this variability may be transmitted downstream slowly, meaning that an instantaneous sampling of an entire transect for DOM characteristics may reveal not only the steady-state mixing dynamics, but also the signal of the changing freshwater endmember. Hence interpretations of conservative vs. non-conservative mixing need to be made with care. These concepts are discussed at length in Asmala et al., Frontiers in Marine Science 12, 2016.

P14 Line 15: I wonder about effects of changes in pH and DO, that may affect the kinetics of Fe oxidation and precipitation. Are there jumps in these values when going from the river into the estuary?

This issue is addressed in detail in our response to the referee's main comments.

P15 Line 5: Crucial. Unsure about the reasoning; Can it simply be the "natural" distribution between DOM and POM for plant material?

This issue is addressed in detail in our response to the referee's main comments.

P16 Line 13: This is the smoking gun; but is it syngenetic or diagenetic? And, there is no direct evidence for Fe-OM from the data, there is just a pool of "undocumented Fe" that is assumed to be organic-bound Fe based on XAS data from Yu et al. I assume that siderite was in the library? Dithionite-citrate may have potential for siderite dissolution? More info on the standards explored is necessary to validate the claim that the Fe was undocumented and thus likely associated with OM.

This issue is addressed in detail in our response to the referee's main comments.

P16 Line 18: This is pretty speculative (based on Yu); any evidence for the fact that it occurs in the water column? Following sentences infer the role of diagenesis, the Fe(III)-OM is purely hypothetical?

This issue is addressed in detail in our response to the referee's main comments.

P16 Line 33: OK, so siderite was a standard (which is a Fe(II) mineral)

Yes. Siderite was one of the reference spectra. This issue is addressed in detail in our response to the referee's main comments.

P16 Line 33: Even fresh ferrihydrite dissolved very slowly at the pH of Na acetate (4.5), are there lit examples of Fh dissolving under those conditions?

This issue is addressed in detail in our response to the referee's main comments.

P17 Line 5: This is interesting, in that redox suttling causes maximum Fh (the most labile and easily reduced Fe(III) phase) accumulation in the surface sediment, where H2S already seems to accumulate (Fig. 7). Perhaps some more words (beyond "seasonally oxygen depleted) can be spent on the equlibrium of a state with apparently co-occurring abundant Fh and H2S accumulation?

This issue is addressed in detail in our response to the referee's main comments.

P18 Line 13: Is there not potential for a role of sediment DOC/POC as Fe(II) sink, i.e. diagenetic OM-Fe(II) formation?

Diagenetic formation of OM-Fe complexes is indeed suggested in the papers of Lalonde et al. (2012) and Shields et al. (2016). i.e. according to their model, the association between Fe and OM occurs after sedimentation. Due to the close association of Fe and OM in the water column of boreal estuaries, we have focused on the idea that Fe and OM are transferred together to the sediments. However we cannot discount the possibility that diagenesis affects the nature of the association and we will acknowledge this in the revised text.

P19 Line 3: The mechanism could be explained more explicitly: how does the recovery drive increased Fe and DOM transport?

We will expand this section briefly. The mechanism is related to the ionic strength of freshwaters and consequent residence time of DOM in drainage systems.

P20 line 25: (

We will correct the typo.

Response to referee comment by Anonymous Referee #2 on “Flocculation of dissolved organic matter controls the distribution of iron in boreal estuarine sediments” by Tom Jilbert et al.

[Referee comments in bold]

[Responses in italics]

This is a paper on the removal mechanisms of Fe to sediments in the riverine-estuarine transition of a long riverine estuary on the Eastern Baltic Sea and the effect of this removal for the corresponding iron diagenesis in the underlying sediment. The key message of this paper, if I understand this right, is that riverine Fe is removed by flocculation in the riverine-estuarine transition of the Eastern Baltic Sea. These results are similar to the Fe story presented in Yu et al (2015) Chemical Geology in another Finnish estuary. The flocculation occurs as an organic Fe complex.

These results are based on the interpretation of Mössbauer spectra, which revealed an unknown Fe phase that was neither well-crystallized silicate, magnetite, nor ferrihydrite. The inference is therefore not direct, but indirect, and that is a major shortcoming.

In the underlying sediments of the outer estuary and the Baltic Sea, Fe accumulation occurs as different phases than in the riverine part, where more crystalline Fe hydroxides occurred than in the distal part where organic Fe dominates.

In principal I believe some of the story, e.g., the organic-Fe transport and rapid removal in the estuary. But I failed to see how the results translate into different diagenetic Fe processes in the sediment.

In particular, I felt that the story on the anaerobic oxidation of methane by iron was underdeveloped. While this is an impressively large data set from many stations, much of the potential novelty hinges upon the interpretation of the Mössbauer data. The combination of Mössbauer/extraction data call for a major reinterpretation of the operational Fe extractions, in particular of a re-assessment of the dithionite extraction as an organic Fe phase. This has large ramifications for many published papers and the authors need to be very careful in their assessment and interpretations.

We thank the referee for the careful consideration of our manuscript. Below, we respond to each of the question posed by the referee.

I would like the authors to address and comment on a number of questions I have:

Why do you not present Mössbauer spectra of riverine material, dissolved and particulate? This would be most interesting to see.

We agree with the referee that such data would be extremely interesting. However the present study was not designed to include Mössbauer analysis of riverine dissolved and particulate material. This would require far greater volumes of water to be filtered (and with various grades of filters) than that required for the determination of total Fe_{diss}*. and* $Fe_{part.}$*. Indeed, this would constitute a separate study in itself. However in response to the comments of both referees we are currently analyzing additional sediment samples by Mössbauer. These include a sample from the river bed of the*

Mustionjoki, upstream from Station A, which may shed some light on the composition to material transported in the river.

Why do you claim that hydroxylamine-extracted Fe occurs as Fe2+, when hydroxylamine is a strong reducing agent? No information on prior oxidation state is possible using this extraction method.

It is true that hydroxylamine-HCl is a reducing agent and in fact we did not claim in the original text that hydroxylamine-HCl extracts Fe^{2+}*. We assume the referee is referring to our claim that dithionite (another reducing agent) extracts* Fe^{2+} *from OM-Fe(II) complexes. In response to the comments of both referees we have addressed the issue of solubility of various Fe phases in our original extraction scheme extensively. Please refer to Figures R2–R5 and associated text in this file.*

Why do you not even consider or discuss the extraction of FexSy phases with dithionite? This is well known. Not all Fe may be organically associated to the same extent throughout the transect and not necessarily as Fe2+, because dithionite is also a reducing agent.

The exact nature of the dithionite-soluble phase(s) remain difficult to determine conclusively, and we require the additional Mössbauer data to advance this discussion. The additional extractions suggest that unsulfidized Fe (II), including the potential OM-Fe (II) complexes, is in fact dissolved in Stages 1+2 of the Poulton and Canfield method, rather than Stages 3+4 (dithionite + oxalate) as suggested in the original manuscript. With respect to the possible dissolution of FexSy phases in dithionite, pyrite is not considered to be dithionite-soluble (see for example Berner, Amer. J. Sci. 268, 1970; Canfield, GCA 53, 1989; Raiswell et al., Chem. Geol., 111, 1994) while greigite is not expected to be a major phase in the sediments at this location. The role of FeS has been established by the additional extractions: this is expected to be dissolved in Stages 1+2 (Fig. R3).

In your table or on the Mössbauer spectra you should show the patterns for FexSy or FeS2 phases to convince the reader that the composite spectrum is not influenced by these phases.

This is a fair criticism and similar to comments from Referee #1. We will present all relevant reference spectra in the revised version.

The Mössbauer spectroscopy standardization and reference spectra are not explained. It is not clear why your conclusion is that the phase must be an organic Fe2+ phase. If it is ferrihydrite associated with organic matter, what kind of association is this?

This is a fair criticism and similar to comments from Referee #1. We will present all relevant reference spectra in the revised version. The allocation of Fe(II) vs. Fe (III) to unknown phases in Mössbauer spectra is made according to the position of the spectrum in x-y space of quadrupole splitting vs. isomer shift, as outlined in Murad and Cashion, Springer 2004. This detail will be included in the revised version.

Is there a possibility that the unknown spectrum is an amorphous Fe-silica phase?

We have no basis to suspect this at the present time.

Why is it that there are hardly any changes in Fe speciation at Station D, although the S content changes so significantly and therefore likely the concentrations of FeS and FeS2?

We apologize to the referee, but we made a significant mistake in the plotting of Figure 6 (and associated text) which is relevant to this question and likely influenced the referee's understanding of the data. The scale of the sedimentary Fe content reads 0–4 %, while it should read 0–40 %. Hence the true figure should look like this:

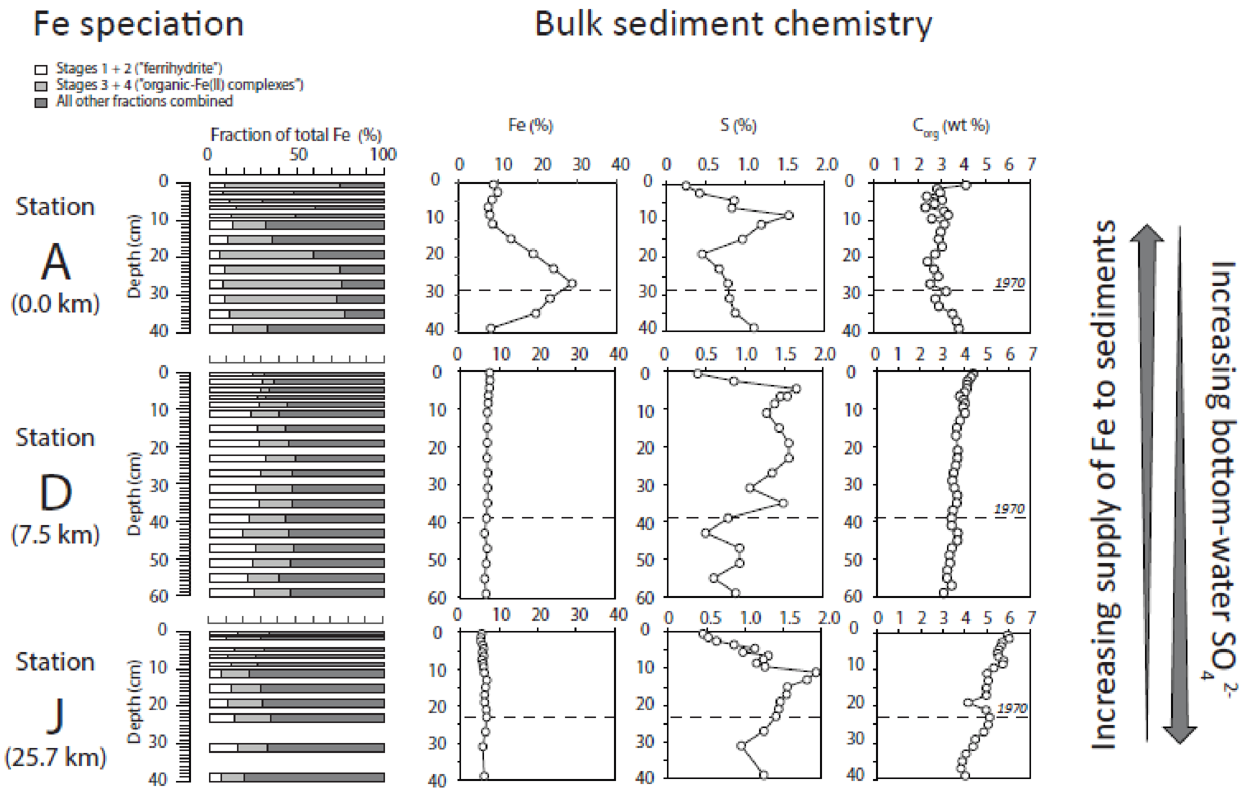


Fig. R6. Corrigendum to Figure 6 in the original manuscript. Note the scales on the axes of weight % Fe in the sediments (0–40%).

Note that all other plots of sedimentary Fe content in µmol/g units are correct, as are the Fe/Al weight ratios given in Fig. 8.

Clearly, the Fe content of the sediments at all stations is far higher than the S content, also on a molar basis, and this is the reason that the downcore changes in S content are not reflected in changing Fe speciation. This is best illustrated in Fig. R2 of this document, where "sulfide-Fe" (calculated as described in the manuscript assuming all S as pyrite) is shown to be only a minor component. This conclusion does not change when the calculation is performed taking into account the AVS (FeS) component now calculated in the additional extractions. For clarity we will add both molar and weight % scales to the figures in the revised version.

What is the major Fe carrier down the river, i.e., what is the speciation of Fe in river and estuarine water? How much is associated with the organic fraction – how much is present as ferrihydrite – what is the exact molecular association?

For the reasons outlined earlier, we did not carry out speciation work on the suspended particulate (or operationally "dissolved") Fe. We emphasize again the relevance of the study of Neubauer et al., ES&T 47, 2013, which discusses the speciation of Fe in boreal rivers.

Figure 5 figure caption and left figure panel don't agree. The caption reads Fe2+/FeT, the other Fe3+/FeT ratio.

We thank the referee for this observation and will correct the text and figure.

Figure 6: The down-core operational Fe profiles don't exactly make sense in light of the pronounced changes in S content with depth at Station A. At least, the 100% scaling makes it difficult to associate the species changes with the S concentration changes. I recommend to show the Fe species as concentrations, e.g., as summed bars totaling to the actual Fe concentration. That would help at least for comparing the data of Station A. The Fe species do not correspond at all to the sulfur concentrations. How is this possible, if FeS/FeS2 forms? My conclusion would be that the dithionite-extracted species seem to be associated at least partly with some FeS/FeS2. This needs to be accounted for.

Again, we apologize to the referee for the mistake in Figure 6 which is relevant to this comment (see earlier response). We will take onboard the suggestion to plot the Fe speciation data as concentrations. A version of Fig. R2 will be included in the revised manuscript.

p.14 l.25: I think the authors mean 'isotopically depleted'

No. The deeper water samples are indeed more isotopically enriched (less negative values than surface water samples).

p.16, l.16-17 Fe-Si amorphous phases; FeS mackinawite-like material?: Why Fe2+, could also have been Fe3+?

We will present all relevant reference spectra in the revised version. The allocation of Fe(II) vs. Fe (III) to unknown phases in Mössbauer spectra is made according to the position of the spectrum in x-y space of quadrupole splitting vs. isomer shift, as outlined in Murad and Cashion, Springer 2004. This detail will be included in the revised version.

p.17, l.1-5 hydroxylamine hydrochloride is a strong reducing agent suggesting that the Fe could have been reduced by the extraction.

It is true that hydroxylamine-HCl is a reducing agent and is specifically used to target poorly crystalline Fe oxides in this scheme. We do not claim in this passage of text that hydroxylamine-HCl extracts Fe^{2+}.

I would also like to see a transect plot of DOM and POM.

We do have a transect plot of POC from the same samples used to generate the corresponding plots for N/C and $\delta^{13}C$. *(Fig. R7). This can be included as a supplement if needed but does not contribute* *significant extra information. The distribution strongly resembles N/C, suggesting that phytoplankton dominates the POC pool at the time of sampling in June 2015.*

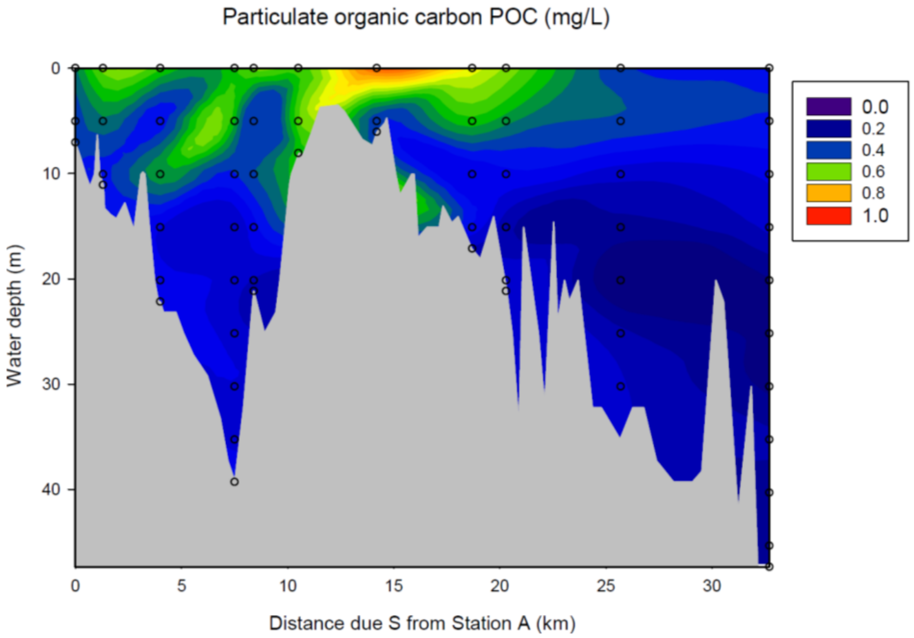


Fig. R7. POC along the transect in June 2015 (mg/L)

Unfortunately we do not have a corresponding plot for DOC (or DOM) as this sampling was not included in the original experimental setup (which was focused on Fe and POM).

p.18. Are the salinity differences significant enough to impact the Fe-S system? I don't think so.

We disagree. The changes in sulfate penetration depth from Station A–D–J (Fig. 7) indicate a significant impact of bottom water salinity on the diagenetic zonation of the sediments. As highlighted in the discussion of this Figure, the depth of the SMTZ, and the intensity of the associated H_2S peak, contrast strongly between the sites. As mentioned also by Referee #1, the distribution of H_2S in the pore waters (between stations and the downcore profile at each station) should then have a strong impact on the stability of Fe phases. We will devote more lines to this in the revised version.

p.18 l.24 what do you mean by background rates: How are these background rates? Aren't these the major reactions? What about sulfate reduction rates and coupled oxidation of sulfide by ferrihydrite?

We will rephrase the sentence to remove the term "background".

Undoubtedly, both sulfate reduction coupled to organic matter oxidation, and oxidation of sulfide by ferrihydrite – along with many other diagenetic reactions – are also occurring at these sites. We will expand the discussion to give a broader overview of the various diagenetic processes, including first order estimates of process rates derived from pore water profiles. However it was not (and is still not) our intention to perform a detailed diagenetic modeling study, rather to highlight the diagenetic zones that are clearly visible in the pore water profiles, in order to qualitatively discuss the differences that are observed along the transect (and the link to Fe inputs from flocculation).

The anaerobic oxidation of methane by iron is often invoked these days, but to argue for this process there has to be good direct evidence. I am sure the authors are aware that AOM also can be coupled indirectly to reduction of iron oxyhydroxides through sulfide oxidation, but cannot be distinguished easily without performing specific experiments. Concentration profiles alone are not enough. The authors should refrain from inferring that AOM by iron is a major process controlling deep iron diagenesis when they have not addressed sulfide oxidation processes. They do not even present DIC data to support their assertion. In addition, should this process occur, it is easy to assess the quantitative significance by assessing the methane flux and the required removal of Fe to account for methane oxidation.

We are indeed aware of the alternative indirect pathways by which reduction of Fe oxyhydroxides may be coupled to methane oxidation. The most relevant is of course the so-called cryptic sulfur cycle as described in Holmkvist et al., GCA 75 (2011), in which downward-diffusing H_2S from the SMTZ is oxidized by Fe oxyhydroxides, leading to the formation of native sulfur and subsequent disproportionation to SO_4^{2-} and H_2S (the SO_4^{2-} then going on to oxidize CH_4). We do not dispute the validity of the mechanism presented by those authors and indeed coupled sulfur cycling and methane oxidation have been confirmed by further studies (e.g. Milucka et al., Nature 491, 2012).

However, as discussed at length in the review process for Slomp et al. Plos ONE 8, 2013 and Egger et al., ES&T 49 (2015) for sediments from the Bothnian Sea, in the low-salinity systems of the northern Baltic the SMTZ is sufficiently shallow that H_2S diffusing downwards from the SMTZ is exhausted well above the base of a typical GEMAX core. This is confirmed in the profiles in Figure 6, where H_2S is undetectable below approx. 20 cm at Stations D and J. It is therefore problematic to invoke the cryptic sulfur cycle as the cause of high pore water Fe^{2+} at depth, because the downward diffusion of H_2S is the ultimate driver of this process. Although H_2S is regenerated during sulfur disproportionation within the cryptic sulfur cycle, only three moles are produced for every four moles of H_2S that initially react with Fe oxyhydroxides. So cryptic sulfur cycling is in fact a net consumer of H_2S and cannot sustain H_2S-driven Fe oxyhydroxide reduction well below the downward-penetrating H_2S front. This concept is nicely illustrated in the study of Egger et al., Biogeosciences 13, 2016 for Black Sea sediments (see their Figure 8, where pore water Fe^{2+} production due to cryptic sulfur cycling is shown to be restricted to a narrow depth interval just below to the downward-penetrating H_2S front). Hence we are confident that the large increase in pore water Fe^{2+} observed in the deeper parts of the cores in our study are not driven by cryptic sulfur cycling.

To clarify, in the original manuscript we also give two possible mechanisms for the production of pore water Fe^{2+} (Eq. 7 and 8), thereby acknowledging that both dissimilatory reduction of Fe oxyhydroxides, and Fe-AOM, may be active in the deep sediments. Hence coupled to the above discussion re. cryptic sulfur cycling, we disagree with the assertion that too much weight is given to the likelihood of Fe-AOM.

In related work we are performing experiments to determine the rates of AOM in these sediments, but these results are beyond the scope of the current paper.

Along these lines, generally there is also too little discussion on sulfide oxidation coupled to iron reduction in the surface sediments. These organic-rich sediments likely have very low oxygen penetration depths of a few mm. Based on many other studies in estuarine systems, it is likely that anaerobic degradation processes such as sulfate reduction commence in the first centimeter. This makes it possible that FeS phases already occur in the topmost cm, and that not only iron reduction, but co-existing iron and sulfate reduction take place in the topmost cm.

The referee is correct that oxygen penetration is in the order of mm, and that both sulfate reduction and Fe reduction likely co-occur in the uppermost cm. This is implied by the immediate decline in pore water sulfate below the sediment-water interface at all sites, and the basic principle that Fe oxide reduction is more than twice as energy efficient as sulfate reduction per mole carbon oxidized (Stumm and Morgan, Wiley, 1981) and therefore should activate at a shallower depth horizon. In reality the diagenetic zones in these sediments overlap extensively, which in the surface sediments is also related to bioturbation and bioirrigation processes including by the invasive polychaete Marenzellaria (e.g. Norkko et al., Glob. Ch. Biol. 18, 2011). The referee is also correct that FeS may be present in the surface cm, either formed in situ by the co-occurrence of sulfate reduction and Fe reduction, or transported vertically from deeper horizons by the "smoothing" action of bioturbation. All of these concepts will be included in an expanded discussion of the diagenetic processes in the revised version of the manuscript.

Finally, although the authors do very well in describing the bathymetry of this estuarine system, they fail to associate the bathymetric features with the current transport/ hydrography and the resulting particle transport and accumulation. For example, Station D likely must be influenced by saline water transport upstream, which is the only explanation to explain the higher bottom salinities. Therefore, inshore/upstream transport of organic material and of Fe has to be considered for the deep depressions.

It is true that saline inflows across the sill are responsible for the relatively high salinity deep water of the inner estuary. Inflows typically occur in winter as established in early literature on this system (e.g. Virta, Nordic. Hydrol. 8, 1977) and Niemi (Meri, 4, 1977). We will include an expanded introduction to the hydrodynamics in Section 2 (Study location). We are also aware of the likely lateral transport of Fe and OM (indeed Page 17, Lines 4–14 of the original manuscript specifically address this issue). We will rephrase this section to emphasize that lateral transport into the bathymetric depression of Station D may occur in both and upstream and downstream direction. Figure 4 will also be modified to highlight the position of the bathymetric depression and the sill in the data series of OM and Fe.

In addition, Station D, being the deepest station of the inner estuary, is clearly a particle trap of fine-grained material, also of organic material. As such, the focusing and accumulation of material here may override the estuarine mixing signal the authors have as their overriding study target. Stations C or E may be more informative in this context. Do the authors note the same signals at stations C, D, and E?

As outlined in the previous response we are aware of the potential effects of lateral transport and focusing at Station D and will emphasize these more clearly in the revised version. In terms of the relative effects of the focusing and estuarine mixing on the observed signals, and the comparison of Stations C, D, and E: It could indeed be argued that Stations C and D display slightly elevated concentrations of "phytoplankton-derived" OM relative to the offshore trend (see Fig. 4b, lower left panel). We will give more weight to this in the revised version. The missing Fe data from Station C makes the equivalent assessment more difficult in the case of Fe.