

**Review to “Technical Note: Uncovering the influence of methodological variations on the extractability of iron bound organic carbon” by Fisher et al.**

Fisher et al. investigated how modifications of the frequently used citrate-dithionite-buffer (CDB) extraction for iron-bound organic carbon influence the respective results. The CDB method is widely applied in soil and marine sciences to extract iron (Fe) and co-precipitated or adsorbed organic matter. Despite its common application, the method has some drawbacks that are, according to the authors, often neglected or at least not properly discussed. In this sense, this study reminds me of the recent publication by Hepburn et al. in Chem. Geol.: “The use of operationally-defined sequential Fe extraction methods for mineralogical applications: A cautionary tale from Mössbauer spectroscopy” and the study by Oonk et al. (2017, Chem. Geol.): “Fraction-specific controls on the trace element distribution in iron formations: Implications for trace metal stable isotope proxies“. As in these previous publications, Fisher et al. try to tackle the problem that wet chemical extractions lead to operationally defined fractions that are not entirely specific to distinct minerals. The authors set up experiments where they varied the strength of the chemical extract as well as the composition of the sample that’s to be leached. They also tested whether a longer duration of the CDB treatment leads to higher Fe and Fe-OC yields. Studies like this are urgently needed to achieve comparability of datasets even though they unfortunately never result in a crystal clear recipe that is to be preferred for all kind of samples. However, this article will make researchers more aware of the shortcomings of the CDB method so that they are put into a better position to judge in which way they should apply it and discuss their data.

The manuscript is for the most part well written and easy to understand. The figures and tables are adequate and the discussion is supported by the presented data. What is missing a bit (probably due to my personal background) is a wider implication of the finding that CDB treatments did not lead to a full recovery of present reactive Fe in any of the tested samples. CDB is not only used for Fe-bound OC, but also for Fe-bound phosphate (see papers by Ruttenberg, Slomp, Kraal) and of course Fe-oxide extractions after Poulton and Canfield (2005), whereby dithionite was recently shown to also extract substantial amounts of magnetite (and clay). This might not be the exact topic of this article, but should at least be mentioned as I feel that it would increase the relevance of this article by a lot. The authors say that Fe<sub>R</sub> extraction was incomplete for their synthesized sediment samples. I would assume that it could potentially also be overestimated in some kind of natural sediments. (At least this is what I often observed.) So my main recommendation would be to revise and complement the discussion accordingly / to expand the “framework” of the discussion a bit further in order to address more readers and demonstrate the real relevance of this nice experimental work. I will give some more recommendations in the following and recommend publication of this study after moderate to major revision.

General comments:

There should not be a period after a title. (You wrote e.g. “Abstract.”) Titles aren’t sentences.

The manuscript should be checked for a consistent use of the expressions “concentration” and “content”. I recommend reading Tolhurst et al. (2005, Estuarine, Coastal and Shelf Science): “Content versus concentration: Effects of units on measuring the biogeochemical properties of soft sediments“. Furthermore, please check the order of references in the text. The publications should be ordered according to their year of publication.

More specific remarks:

[Line 16](#): I suggest to use “synthesized sample” instead of “sediment”.

[Line 36](#): Delete “important” before “for water retention” as it is an unnecessary repetition.

Line 86-88: “Wagai and Mayer (2007) performed a 16 hour extraction (substituting citrate with weak HCL acid rinses to avoid use of organic compounds), and Patzner et al. (2020) extended to 6 hours.” “HCl” instead “HCL”. And just a comment: I'm a bit puzzled by this statement regarding citrate. Citrate is added so that Fe-complexes are formed and Fe is kept in solution. I should probably read the paper by Wagai and Mayer, but acid rinses seem critical to me when it comes to comparability of datasets (which is obviously why you investigated it).

Line 99: “...rapid decomposition of dithionite in aqueous form suggesting, a quick loss of reduction potential...”

Incorrect comma placement.

Line 118-119: “To achieve this, we mixed the precipitate with a marine sediment ‘carrier’ material as described by Fisher et al. (2020), using **the same original carrier sample and similarly treated to liberate OC and inorganic carbon.**”

Weird sentence structure. In order to make it easier for the reader I suggest to include one or two sentences to what the carrier material is. I guess the original Fe<sub>R</sub> contents are known? It's fine to refer to the previous publication, but the reader shouldn't be "forced" to look it up.

Line 126: Replace “A” by “The”.

Lines 164-166: “Initial concentrations of Fe in synthetic samples were obtained by digesting ~2 mg of dried sample in 1 mL 12N HCl at room temperature followed by a 10-fold dilution with 1% HCl solution. Further dilutions were made 165 as necessary, dependent on Fe content, using MilliQ water to produce a subsample within the detectable window (1–10 ppm Fe).”

I'm not quite sure about the fraction that is intended to be dissolved here. Bulk Fe? I guess it's okay in case that the synthetic sample does not contain Fe-bearing silicates. As mentioned above it would be good to add what was used as "carrier" for the Fe oxide-OC spikes.

Add “of the extract” after „10-fold dilution”.

Section 2.6 about ICP-OES analyses: Generally (for future), I would recommend using an internal standard for correction of different ionic strengths.

Line 171-173: “Extraction of Fe was calculated by subtracting the amount of Fe lost in the control experiment from Fe lost following extraction, then subtracting this from the initial Fe of each sample.”

The formulation of this sentence seems more complicated than necessary and I don't fully get it. How about: "The recoveries of the extractions were determined as extracted Fe compared to the initial Fe content." (I suppose you made sure that the carrier sediment that you spiked did not contain any Fe?)

Line 188: “...requires a 0.25 g addition relative to 0.25 g of dried sediment sample”

Recommend to use “per” instead of “relative to”.

Lines 191-194: “All samples show incomplete reduction of Fe regardless of Na dithionite addition, with those samples containing the least Fe proving extractable for the greatest proportion of Fe.”

Unnecessarily complicated formulation. How about: "highest recovery of Fe in samples with low OC-Fe contents".

Line 199-201: “From this, we can deduce the maximal %Fe in sediment extractable by 0.25 g Na dithionite lies between a 20 and 30% OC-Fe<sub>R</sub> mix, equivalent to 7-10 wt% Fe content in the sediment.”

I would slightly reformulate the last part of the sentence to not imply that this is total Fe you're talking about. And wouldn't it make sense (for practical reasons) to translate your "20 and 30% OC-Fe<sub>R</sub> mix" into an absolute amount of Fe (e.g. in mmol or mg) that can be liberated?

Lines 203-205: This is about LECO data, right? I wonder whether you could avoid confusion by just calling it "extracted OC" or OC<sub>Fe<sub>R</sub></sub>. Calling this fraction OC-Fe<sub>R</sub> is a bit confusing as I would intuitively translate it as "OC-bound reactive Fe". But you mean "reactive Fe-bound OC". As for the Fe I assume that your carrier did not contain any further OC?

Line 255-256: "For the four synthetic samples we subjected to dithionite reduction, these differed in composition (7-24 wt% Fe, 20-50% initial OC-Fe<sub>R</sub> content)."  
Didn't you also have a batch with 100% OC-Fe<sub>R</sub>??? (See Table 2.)

Lines 256-257: "The concentration of Fe in these samples results in an effective dithionite to (wt) Fe reduction reaction ratio of 1:0.07-0.24."  
Replace "concentration of Fe" by "Fe contents", "results" by "resulted" and replace "effective dithionite to (wt) Fe reduction reaction ratio" by "effective dithionite to Fe mass ratio".

Line 261-262: "This has the potential to drive wt% Fe higher in small samples of sediment such as those treated by the method (0.25 g)."  
Recommend to replace "treated by the method (0.25 g)" by "typically used for the CDB extraction".

Lines 266-269: "Maximal extraction here is defined as the point from which further addition of Na dithionite does not increase the extraction of Fe beyond the amount of Fe extracted under the previous dithionite addition mass ± error. For example, the 20% OC-Fe<sub>R</sub> sample subject to 0.25 g dithionite is removable for 88.79% ± 3.55 of Fe<sub>Total</sub> while 0.375 g addition extracts 90.94% ± 3.64; ..."  
What is meant by "is removable for"??? Unnecessarily complicated formulation. Use "yields" or "liberates". Delete "beyond the amount of Fe extracted under the previous dithionite addition mass ± error" and add a "further" before "increase". How this is meant is getting clear through your example.

Line 276: Missing space before 2.69.

Line 276-278: "This finding demonstrates that the OC-Fe<sub>R</sub> composition would not be correctly determined following the method of Lalonde et al. (2012) for these OC-Fe<sub>R</sub> rich sediments, and the overall extent of OC-Fe<sub>R</sub> in the marine sediment pool would be underestimated."  
You can delete the "pool". It kind of implies that you're talking of a specific fraction of the marine sediment, but here you mean the sediment itself (bulk). Why would you limit this to marine sediments? Couldn't you say this is a general outcome of your study no matter which sediment (fluvial or marine or soil) is used? (Now, again, it would be nice to know the composition of your carrier material.) I would write "amount" instead of "extent".

Lines 278-281: "While 30-40% OC-Fe<sub>R</sub> content is above the average for marine sediments, many samples exist in the 20-30% range. Indeed, the average value for marine sediment OC-Fe<sub>R</sub> composition given by Lalonde et al. (2012) is greater than 20% with individual marine sediments recorded as exceeding 30% OC-Fe<sub>R</sub> (e.g. Equatorial Pacific 0°N, 34.79% (Barber et al., 2017))."  
I am, to be honest, a bit confused by these numbers. 20-30% of Fe-OC really seems high to me. I never had such high amounts of reactive Fe. Fe plus bound OC is surely higher than reactive Fe alone, but with dithionite you typically reduce all kinds of Fe oxides including ferrihydrite, lepidocrocite, hematite and goethite as well as (unfortunately) some magnetite. (At least with the Poulton and Canfield method published in 2005.) Nevertheless, by applying this method I never ended up with more than 3 wt% extracted Fe in the sediment out of usually around 6 wt% total Fe including all silicate Fe and sulfides. Please double-check your numbers!

As mentioned above, the usual CDB extraction includes crystalline phases like goethite and hematite that might not be so relevant for OC. I am missing a statement concerning how the (maybe in your case unintended?) leaching of more crystalline phases potentially skews the  $Fe_R$  : OC relationship. The typical amount of highly reactive Fe (amorphous phases) in shelf sediments is, I would say, less than 1 wt% (so by far lower than what you were testing for). So I would therefore be a bit hesitant to transfer the results of your experimental data to real marine sediments and it's good that you included tests with Antarctic sediments in this study.

Line 290-291: "If the increased strength dithionite treatment increases dissolved Fe beyond the complexing capacity of citrate, then excess Fe likely precipitates out of solution before measurement."

This can be avoided when performing the extraction under anoxic conditions (e.g. Henkel et al. 2016).

Line 293-295: "Measurement of OC- $Fe_R$  extracted for the concentration of Na dithionite at which maximum Fe is extracted showed incomplete OC- $Fe_R$  loss (Fig. 1)."

I would replace "loss" by "liberation" or "recovery".

Line 294-295: "The similarity of OC- $Fe_R$  and raw Fe extraction values indicates that OC and Fe are reductively released from the sediment in comparable proportions, as is expected due to the low molar OC: $Fe_R$  ratio of the coprecipitate (~0.7:1) ."

What is meant by "raw" values? Raw data is typically used in another sense.

Lines 297-298: "...could benefit from using increased strength Na dithionite compared to the 0.1 M treatment currently used."

Or shorter: "compared to the conventional 0.1 M treatment".

Line 308: Replace "have been" by "are".

Line 309: Replace "defined" by "assessed so far".

Lines 324-328: "We postulate that freeze drying-induced aggregation of sediment particles could result in reduced Fe extractability compared to non-dried samples since grain size is a known key factor in limiting determination of bioavailable Fe (Raiswell et al., 1994). Aggregation could reduce surface contact with dithionite, preventing reduction of 'shielded' sediment particles, while this could be overcome (e.g. by crushing), and this in itself would introduce further variability in grain size (Raiswell et al., 1994)."

I'd actually argue the other way around. I am wondering about the potential differences between grinded and non-grinded natural samples. You typically freeze-dry samples to be able to grind them and make them more homogenous. I would assume that the freeze-drying itself might result in a transfer of Fe from a more reactive into a less reactive pool. But at the same time I would guess you reduce effects of grain size differences or clogging/shielding of grains (coatings) by grinding the samples. Ok. I see that you mention this in the following sentence. (Add a space before "The influence...") As grinding is what's typically done, I'm not convinced that the aggregation plays the dominant role. I'd rather think that the amorphous Fe compounds aren't stable during the processing (freeze drying). Would be worth checking whether there is a transformation of ferrihydrite during and after drying...

Line 335-336: "The alternate tested method of using wet samples has largely been avoided, with only a few studies (e.g. van Bodegom et al., 2003; Chen et al., 2020) reporting the use of a wet slurry sample in soils and none for sediments."

Suggestion: The alternative method of using wet samples has largely been avoided, with only a few exceptions in soil studies (e.g. van Bodegom et al., 2003; Chen et al., 2020).

That's not true when you don't limit your view to the Fe-OC extraction by dithionite but also consider the many studies focussing on the Fe or P. The Poulton and Canfield (2005) method that includes a similar dithionite step is often applied to wet sediments. Check papers by Natascha Riedinger, Laura Wehrmann and Katja Laufer (2019, Reactivity of Iron Minerals in the Seabed Toward Microbial Reduction – A Comparison of Different Extraction Techniques). The same is true for Fe-P extractions with CDB (Kraal, e.g. 2017 GCA paper). One reason for people sticking to the freeze-drying and grinding is that with lots of samples, that's the only option. I'm thinking of IODP material (usually pretty hard mud rock) or black shales.

Line 346-347: However, the use of wet sediments is likely to be inappropriate for some analyses or sample sites.

Yes! You should add one or two sentences to that. I believe it's for most cases not as if people using these methods are not aware of its shortcomings.

Line 354: You can delete the "method" after "storage". "Any storage" is enough.

Line 366: "...slurry form..."

Delete "form".

Lines 370-373: "As we observed incomplete Fe extraction (Fig. 1) for all our samples, a range of CBD extraction times were trialled to understand whether increasing the length of a reaction would increase Fe liberated, as seen for other chemical Fe extractions; oxalate, for example, is known to continue to extract Fe beyond a standard 1 hour treatment (McKeague and Day, 1966)."

Okay, but it does not make too much sense to compare te CDB method to the oxalate method, because the oxalate extraction works differently. The extraction is actually catalyzed by dissolved  $Fe^{2+}$ . So the longer the extraction continues, the more  $Fe^{2+}$  is in solution and the stronger gets the extraction (well described in Oonk et al., 2017, Chem Geol. and references therein).

Lines 373-375: "Additionally, as previously mentioned, some iterations of the CBD method have been repeated multiple times in succession to extract the full  $Fe_{CBD}$  pool, but it is unclear whether time or reagent concentration limit full extraction of this pool on the first treatment."

By you or others? It's not getting clear here.

"Iteration repeated multiple times in succession..." Here you say the same thing twice (or actually three times).

Line 379: "... concluding that an increase in chemical exposure time has no difference on Fe extractability."

Replace "has no difference on" by "has not enhanced" or "has no effect on".

Line 381: "We would perhaps not expect any benefit from increasing the length of CBD treatment as dithionite, ..."

You don't seem to be very convinced by your data. Replace "would perhaps not" by "do not".

Line 382-383: "... with a rapid second order rate constant ( $K_2$ ) of  $3.0 \text{ (g-molecule/L)}^{-1} \text{ min}^{-1}$  at  $79.4^\circ \text{C}$ , ..."

The unit is written in an unnecessarily complicated way. I guess it should be  $\text{L}/(\text{mol} \cdot \text{min})$ ? Please check!

Line 415-417: “We suggest that if future studies were to increase Na dithionite addition in the CBD method this should be followed by a similar increase in trisodium citrate to ensure the entire reduced Fe pool is complexed, preventing precipitation of Fe before quantification.”

I ran some tests with citrate myself with Fe contents that are comparable to natural occurrences and found that it's usually not limiting. The citrate concentration can in fact be reduced compared to original protocols (I tested the Poulton and Canfield method) as long as you work under strictly anoxic conditions (Henkel et al. 2016).

Lines 421-422: “Freeze drying induced aggregation appears to reduce Fe liberation in synthetic coprecipitates that were freeze dried relative to slurried, however, we were unable to replicate this increased extraction for natural samples.”

Suggestion: Freeze drying induced aggregation appears to reduce Fe liberation from synthetic coprecipitates. However, we were unable to confirm this reduced Fe extraction for a set of natural samples.

Lines 422-424: “While we speculate this may be due to the use of freeze thawed samples, which can introduce aggregation in itself, it is hard to see a practical implementation of this adjustment for marine sediments due to the difficulty in transport of pristine samples.”

Replace “which” by “where freeze thawing”, otherwise your reference isn't fully correct. (You'd refer to the samples and not the process of thawing.) And I believe you can delete the “in” before “itself”.

Line 425: Add “the” before “dry weight”.

Line 427: Period missing after “extraction”.

Figure 1:

I have difficulties understanding your black and blue symbols. Shouldn't the percentage of extracted Fe be equivalent to the extracted OC-Fe? Or is the data behind the blue symbols the LECO-data? Do you really need the separate axis with the different scale??? It's just (at first glance) confusing that e.g. the blue diamond is so much further up the fitted curve. And the offset in “dithionite added” between blue and black symbols (equivalent to maybe 0.1 g) is odd, too.

Figure 2: You don't need 3 different patterns if you distinguish between the different OC contents by different colors (gray scales). So reduce the complexity of this graph by just using 3 colors for the three differing OC batches and filled vs. hatched bars for dry and wet. I would also (for clarity) change the figure a bit so that it doesn't appear as if the OC-Fe to total sediment ratio was 5% for the lowermost wet batch 3 COOH mix and close to 30% for the lowermost dry batch 1 COOH mix. You know what I mean? Those extractions all belong to the 20% test, right? Figure caption: Colon after “Figure 2”.

Figure 3: Colon after “Figure 3”.

Table 1: Use format “left-aligned” in the first column.

Table 2: I find the expression “%OC-Fe:sediment” a bit confusing. I guess you mean % of OC-Fe coprecipitate to total sample”. It's inconsistent because when you write “Sediment (mg)” you mean the carrier only.

I trust this review is fair and constructive.

Susann Henkel