General comments – response to associate editor and reviewers

We appreciate the thoughtful feedback from the associate editor and reviewers. We changed the text of the Introduction and Abstract to better reflect the novelty of our research and state our research goals more clearly. We also reorganized parts of our results and discussion to connect more clearly and logically with these research goals.

Response to Reviewer 1

Overall comments:

> In "Reproducible determination of dissolved organic matter photosensitivity", Armstrong et al. present a new methodology for obtaining reproducible results for photodegradation experiments of DOM SPE extracts. In the end, I believe that this paper is likely publishable, but I am recommending further revision. At its heart, this paper is a methods paper, and it is unclear to me whether the method is superior to less arduous methodologies. Comparisons of this technique to alternative approaches that clearly show why this approach should/could be adapted or clearer demonstrations of how inferences cannot be obtained without this technique are lacking or not (yet) convincingly explained. Below, I elaborate in more detail.

We agree that the approach presented here is more arduous than other methods used to study DOM photochemistry, but hope our revisions illuminate the fruits of and need for this effort. As noted the introduction has been revised to clarify our goals and the relevance of our results. We presented evidence of several previously unreported sensitivities that may affect experimental photodegradation results, which has relevance for the broader research community even if they do not adopt our approach to address these issues. Our research goals included inferences about compositionally similar DOM (the comparison of two wetlands) – we needed a method that was sensitive enough to reliably discern small differences in photosensitivity that were potentially important to the biogeochemistry and ecology of the larger ecological system. In other cases this rigor may not be required. However, both the findings from our specific study system and the broader methodological insights we present are relevant to researchers interested in DOM photochemistry, even if they do not wish to replicate our particular approach.

> In Fig. 6, the authors show that the SPE DOM and whole water samples differed in apparent composition. I would have liked to have seen more discussion for how this would affect ecological inferences. Does the apparent enhanced reproducibility warrant removal of the sample matrix (and loss of environmental conditions/relevance)?

The reviewers asked whether our focus on solid phase extracts to isolate DOM from its matrix would weaken ecological inferences. As the reviewer noted, we employed extracts to improve reproducibility and comparisons between samples (especially if samples required storage before experiments were possible) but we also used extracts because we think more work needs to be done to disentangle the roles of matrix composition and DOM composition in DOM

photochemistry. Our results here in fact emphasize the need for complementary research on potential matrix influences and their relative contributions to DOM photochemistry. Our revised manuscript should make this distinction more clearly.

> Also, did the authors compare the various kinetic parameterizations for the 2 user comparison in Fig. 5… and is the affect of the different users on the modeling greater than that of the sample prep?

The biexponential model parameters for the 2 user comparison shown in Fig. 5 were compared with 2-tailed t-tests that did not indicate differences in mean parameter values. While not conclusive given n=3 for each operator, the same test was able to distinguish between other comparisons using n=3 on either side such as the RO vs PPL test of SRNOM, or wetland DOM PPL vs. SRNOM PPL. We interpreted this as evidence that with proper control of experimental procedures and conditions, variability due to user differences was less than variability from real compositional differences, evidence of method reproducibility not always reported explicitly. We recommend researchers collecting photodegradation kinetics data undertake similar tests of their procedures.

> Is the SRNOM PPL data in Fig. 6 the same as presented in Fig. 5?

They are separate comparisons with different data sets. Initial tests to compare operator effects (as in Figure 5) produced discrepancies. A component fault arising during these initial comparisons may have been responsible, but we also rewrote operating procedures to ensure they left no room for seemingly trivial or arbitrary differences in execution. Then the two researchers took turns collecting the data shown in this figure. The data in Figure 6 was collected earlier by one of the researchers alone, following the same procedures as those later codified in the more exacting protocol. We grouped the data to reflect the comparisons we intended to make before executing the experiments.

> It's not clear to me whether these conclusions require the approach used here. I would have liked to have seen a comparison to the alternative (some combination of fewer datapoints, no manipulation of pH, whole water, etc. that is currently being used in the literature) to demonstrate the superiority of this method or how using an alternate method gives misleading information. (Alternatively, I missed the explanation, and the manuscript would benefit from a clearer statement of methodological benefits.)

We agree that the specific utility of our approach could be brought into greater relief with comparisons to other methods. This is most clear in (revised) lines 557-575 and Fig. 12, where we compared results from our approach to use of absorbance data alone to highlight what is possible with the fluorescence kinetics. There is also more discussion of our focus on isolating influence of DOM composition on photodegradation potential in lines 426-449.

> They then suggest a series of speculative hypotheticals for how photosensitivity differences might have implications for other ecosystem processes. These remain hypotheticals at this stage, but if the method is demonstrated to be superior/necessary,

answers to many of the questions that are posed would be quite valuable.

We thought additional tests to explore the hypotheses generated by our results would be overwhelming in a manuscript that already explores several experiments, though we hope to continue these lines of inquiry in the future.

In the Introduction, line 70, the authors state that the goal is "to compare the photosensitivity of DOM sources…." But this is misleading. That may be the ultimate goal, but it is not a major focus in this paper which seems to be a proof of concept for the methodology – as noted by the title of the manuscript. I recommend revising the stated goal to reflect the actual work.

We agree and have revised the manuscript to state the scope of our goals more clearly.

I found the storage comparisons to be useful but distracting from the main body of the manuscript. I recommend they be moved to supplemental information.

We mostly agree. The storage experiments are mentioned in the main text but the figure with results and description of its relevance are now in Appendix B to improve the flow of the paper.

The term DOM photosensitivity adds to the already confusing dictionary of photochemical jargon. -why not use DOM potential photodegradation which is the definition provided by the authors for photosensitivity?

We disagree. Avoiding unnecessary jargon improves our science but the meaning of "photosensitivity" seems intuitive in the context of audiences who are interested in photochemical properties of DOM and the mechanisms of its chemical transformation when exposed to sunlight. It seems easier to read "photosensitivity" in a sentence than "DOM potential photodegradation" in this context of this manuscript and related works, as long as the meaning of the term is clearly communicated.

In the paper and in the responses to one of the reviewers, the authors note that their approach may not be suitable for every situation and that researchers will have to make that decision for themselves. But ultimately, the authors are writing this paper because they hope the method will be adopted in the future. I suggest the authors make specific recommendations for when using this approach makes sense (and maybe suggest when it does not).

We have expanded our thoughts on this point in lines 557-589, though did not offer any ironclad rules or guidance. Ultimately it is up to the researcher to decide which methods are best suited to their questions, and those criteria will vary between individuals.

"optically thin" is mentioned in line 135 but not defined until line 315. Move its definition to first mention.

We changed the wording of the first occurrence to focus on the purpose of using optically thin samples to save the term and its definition for a more substantial discussion of the concept relevant to our results.

> line 172, are there any effects of adding Cl with HCl? Removing halides was mentioned as a benefit of the PPL step.

This is an excellent question. For some applications this could create problems – for example, investigations into disinfection byproducts. However, pH control was an overriding concern and alternative acids all have their own drawbacks – for example nitric acid absorbs in UV ranges (potentially affecting both photodegradation itself and optical measurements), formic acid would add organic C and absorbs in the UV range. We decided HCl additions would be the least disruptive, most feasible approach.

> -line 290 – it would be beneficial to state what the two types of reactions are.

The revised manuscript includes more information in lines 495-499 and 519-523.

> In Fig. 4, the results of the stir/no stir comparison are not discussed. Stirring does not appear to have much effect on the PARAFAC results (less effect than is seen for Fig. 5), is that correct?

This is correct. We compared data from trials with and without a stir bar present in the equilibration flask and the results did not differ. This did not seem as important as other methodological insights but we have added our observations to the text (lines 314-315).

> In Fig. 8 caption, what are DF, FN, and QB?

These are site codes identifying different wetland locations. We agree these designations are not meaningful in this context, distracting the reader, and have modified the caption.

> nitrite actinometry is an important recommendation from this study but is not described in detail in the methods. Please add these details so researchers can replicate your work.

We agree and have expanded our description in the text (lines 114-124).

> line 439, a sentence was started but never finished.

We were not able to find the fragment described here.

Response to Reviewer 2

> Line 117: Can the authors be more specific what "longer storage" means? In the cited Murphy et al. (2018) study I could not find anything on the effect storage has on bulk

> DOM vs PPL extracted DOM. Can they provide an estimate how long bulk DOM can be stored (I assume frozen) for fluorescence analysis as opposed to PPL extracts?

We appreciate the reviewer's careful attention to our clumsy word choice in reference to this citation. This question deserves more thorough investigation than can be presented here. In our samples, optical properties of DOC-rich wetland water (20-40 mg DOC L^{-1}) were not stable after only 2-3 weeks while stored at 4°C after filtration to 0.2 µm. However, one of the authors of the current manuscript has observed apparent stability in DOC-poor seawater sample optical properties after storage at 4°C for up to a year. We avoided freezing samples to prevent unwanted flocculation of DOM and any possible subsequent alteration of chemical composition as it left and re-entered solution during freezing and thawing. We recommend storage decisions be informed by direct experimentation and reference to published literature on DOM with similar expected composition, but DOM from different settings may have different requirements in this regard.

> Line 119: If the samples are frozen, microbial activity should not play role. How do the authors store the samples prior to analysis?

As noted above the samples were not frozen to prevent possible flocculation, and were instead stored at 4°C, so preventing direct microbial activity required filtration. Extracellular enzyme activity may still be possible in these conditions but we assumed it would be minimal, and available substrates quickly exhausted without replenishment by the microbial community.

> Line 146: It would be nice to see a picture of this set-up or some schematic drawing. This information could be provided in the suppl. material.

We agree, a photograph is now included in Appendix A (fig A1).

> Line 166: The authors explained in the response to the reviewers why this pH was chosen. Nevertheless, I would appreciate a comment here how a pH of 3 corresponds to the expected pH of the samples that were analyzed in this study.

The pH was not chosen to represent the expected pH of similar DOM in environmental settings, but to prevent pH from changing over the course of the experiments themselves, which interferes with the fluorescence kinetics measurements that are the heart of our data. We explain our reasoning in lines 130-135.

> Fig. 1 I would mention in the figure caption that solid lines represent emission spectra and dotted lines excitation spectra.

We appreciate the feedback and have amended the figure caption.

> Line 258 onwards: Why aren't the components 3 and 4 referred to respective F520 and F450 from here on in the manuscript? This would make it easier to compare to the

previous Murphy et al. (2018) study.

This is an excellent suggestion but we worried about miscommunicating. While the similarities between our components and those in the Murphy et al. 2018 study are strong (and supported by high TCC scores), they are not equally similar for both our components. Component 4 in particular had a lower excitation spectrum congruence to F450. As there is no clearly established hard threshold for comparing component identities, we thought it was important to be clear that we were using our independently-fitted PARAFAC components, not projecting our data onto the model fitted to one of the data sets in that paper. There is a real need to clarify standards for nomenclature when discussing results from multiple PARAFAC models!

I find the structure of the Results and discussion a bit confusing. I would have expected the authors to first start with the method evaluation and comparison, i.e. comparing bulk DOM and PPL extracts, etc. instead of directly starting with the photodegradation experiments. Before the results of the experiments can be discussed I would first want to see that the method is robust. Also, from the wide number of experimental set-ups described in the methods, it was hard for me to directly understand which experiment the authors are discussing first. The results and discussion would benefit from a more structured presentation. If they want to keep the current order an introductory paragraph explaining what will be described in the following sections would be helpful.

As noted above, we agree and have changed the structure of the introduction and results to improve the clarity of our work.

Line 281: why is Fig. 10 cited here, when next in line of figures should be Fig. 3?

This was confusing but should be remedied with the revised manuscript structure.

Section 3.1.4: Can the authors refer to how their results relate to other studies that compare PPL and bulk DOM extracts?

Fig. 7: Please define what is meant with SL (= semi labile) and L (= labile), it is hard to find this information in the text and it would thus be helpful if this abbreviation is defined in the figure caption.

These definitions were presented most explicitly in a much earlier paragraph, in line with some of the other structural issues that made the manuscript's organization confusing, so it should now be easier to refer to this in the text. The definitions have also been added to the figure caption.

Fig. 8: Please clarify what DF, FN and QB are referring to. This figure only shows the time series on the whole water samples, did the authors also do this experiment for the PPL samples? Do they change if they are stored over longer periods of time?

Noted above – these are site codes identifying different wetland locations. We agree these designations are not meaningful in this context, distracting the reader, and have modified the caption. The storage experiment was only performed for whole water samples, as longer storage periods of solid phase extract methanol eluate at -20°C has been shown to produce consistent sample composition (1H NMR) before and after storage in another study, and effects of storage of methanol eluate at -20°C on DOM chemical composition was thoroughly investigated by Flerus et al (2011) and found to be acceptably stable using concentrated seawater DOM.

> Lines 431 onwards: Mention and evaluate again in this paragraph the potential for the (quantitative) loss of compounds during PPL extraction and how this may affect the outcome and interpretation of photodegradation experiments. It does seem like this loss of compounds does not affect the PARAFAC model, but do the authors think that this is universal for all different types of samples from different environments or may the loss of compounds have more profound effects on variations in FDOM?

We think this is an area ripe for further research. We found that RO SRNOM and SRNOM PPL photodegradation time series produced indistinguishable PARAFAC models, but the differences in kinetics suggest the extraction is altering the composition of the DOM in some way that subtly affects its photochemistry. The clearest differences in the biexponential model parameters were in kSL for both PARAFAC components modeled. The reverse osmosis treatment (and subsequent clean-up with cation exchange) used to concentrate the SRNOM when it was collected removes some matrix constituents, but sodium is concentrated during the process, and the RO material contains metals and likely contains silica oxides and sulfate. The solid-phase extraction process should remove many of these, so they may have an effect on photodegradation that is reflected in our results. (Green 2014, Kuhn 2014). Our work demonstrates the need to better understand the effects of strictly DOM chemical composition vs. matrix composition influences on DOM photochemistry, and our use of solid phase extracts reflects our desire to parse these phenomena more carefully.