Reviewer 1

Reviewer Summary:

In this manuscript Graham et al. explore the potential of pyrogenic DOM (pyDOM) to be biodegraded. A purely computational approach “substrate-explicit model” and previously published data were used to estimate the energy content, metabolic efficiency, and aerobic decomposition of DOM of pyrogenic and non-pyrogenic “natural” molecules. This study provides a computational explanation of why other recent studies have discovered that pyrogenic DOM can be extensively degraded (consumed) by microbes. This work contributes to the recent paradigm shift on the knowledge on pyrogenic matter’s stability and reveals that pyDOM has a comparable bio-degradability (i.e., biological lability/reactivity/consumability by microbes) to natural (e.g., fluvial) DOM. This indicates that pyrogenic molecules are not as recalcitrant as previously presumed and suggesting that part of the combustion continuum is actively involved in the global biogeochemical cycles.

Reviewer Evaluation and Recommendation: This is an excellent manuscript that contributes greatly to the literature on wildfire biogeochemistry. Very well written, flowed well, smooth read, clear visuals. The bio-degradability of pyrogenic matter is currently a hot topic though there are very few studies exploring it. The present article is a great contribution to this research trajectory and is being submitted for publication (and hopefully soon published) at a great timing for the community. Indeed, as the authors mention, this and the other studies are laying the foundations of a lot of necessary future work on pyDOM microbiology/biochemistry. Computational modeling studies, especially like this one exploring bio-degradability, are generally lacking even in the “natural” DOM world, which makes this work very novel and interesting to readers from various communities.

Unfortunately, there is a major flaw with the design of the study preventing me from recommending it for publication. This flaw can heavily skew the conclusions of this work. However, this flaw can be addressed without too much trouble after which the study would be ready for publication. Thus, I recommend this paper to be published after a major revision. My other comments are minor and can be easily addressed.

We appreciate the positivity, constructive feedback, and overall effort this reviewer put into their review. The comments are really helpful for improving the quality of the manuscript. We also appreciate the chance to address them, and we feel this will result in a higher quality manuscript. Please see our responses to specific comments below.

Major comment:

There are multiple concerns regarding the design of the study, in particular, the choice of natural and pyrogenic data. The conclusions of this study are heavily dependent on the comparison between these two datasets (Figure 2). Thus, complete data comparability must be ensured.

Issues:

1. Pyrogenic data (Table S1) in this study was derived from previous studies which had utilized solvent extractions. For example, a good fraction of the pyrogenic data is extracted from the studies of Oros and Simoneit who analyzed aerosol filters and extracted pyrogenic molecules using dichloromethane. Considering that in a natural
environments wildfire deposits charcoal, which charcoal is then water-leached by rain, there are two immediate questions arising: 1) Is aerosol data representative of charcoal on land? 2) Is solvent-extraction comparable to water-extraction? I am unsure how comparable aerosols and charcoals are, but there is likely literature on this. However, regarding solvent versus water, a recent study showed that water and solvent extracts of charcoal are extremely different (McKenna et al., 2021). By looking at Table S1, 100% of the pyrogenic data used in this study is derived from a solvent extraction. It is highly unlikely that these solvent-extracted molecules are representative of pyrogenic molecules in aquatic environments. In order to discuss bioavailability of pyDOM in aquatic systems (the goal of this study), the pyrogenic data used here must be from water extracts. Given that both charcoals are being leached by rain, and atmospheric aerosols are deposited in waters (Wagner et al., 2021), I would say that using data of water-extracts of both charcoal and aerosols would be appropriate, but emphasis should be onto charcoal contributions as they are likely bigger.

We recognize that listing the extraction type for our list of representative PyOM compounds may be misleading to readers and detract from our overarching goal. Our aim in generating the PyOM list was to identify molecules known to derive from pyrogenic materials, which references to are most often found in the aerosol community, although many of these molecules can also be found in natural waters and sediments (e.g. phenols, anhydrosugars, etc., Norwood et al., 2013; Suciu et al., 2019). While we included the type of extraction done and method measured, many times the solvent extractions are done primarily for preparation for the downstream analysis (e.g. GC-MS) and do not accurately reflect their water solubility (or lack thereof). While many of the references do use solvent extractions for the above reason, the molecules in question are often actually quite water soluble in natural systems (e.g. anhydrosugars, small phenols, etc; Norwood et al., 2013; Cai et al., 2020)– even molecules that are surprising based on their Kow's, due in part perhaps to their affinity and co-soublization with bulk DOM pools (such as PAHs, Wagner et al., 2017; May et al., 1978). Therefore, while the total extractable pools in question may be quite different with solvent vs water extractions as the reviewer notes, we targeted specific molecules from a variety of pyrogenic literature sources to extract known formula that should be present beyond one analytical instruments observation window.

We plan to update our PyOM reference list to include a broader range of molecular formulas examined as both reviewers suggest, (see responses below), and plan to provide additional information on whether each compound has been identified in water soluble organic matter previously, to address this comment specifically.


By contrast, for “natural” DOM, authors use data for surface waters and sediment water extracts, which data are data are appropriate for discussing aquatic biogeochemistry.

2. Second issue identified is on the type of data used. For pyDOM, authors use known structures whereas for natural DOM authors use molecular formulas from FT-ICR-MS. The authors say that they focused their study on known pyrogenic compounds because “While we recognize that recent research has applied new technologies to inferring PyOM compound presence in environmental samples (e.g., FTICR-MS), there remains high uncertainty in the confidence of formula assignment and structural information with some of these techniques.” However, the authors use FT-ICR-MS data for the natural DOM, and their own argument can be made for their natural DOM dataset. I agree that using known structures is much better, but at present we do not know most of them for pyDOM or natural DOM, so this may require the use of FT-MS (ICR or Orbitrap) molecular data for this study. Using known structures for pyDOM and using molecular formulas (from FT-ICR-MS) for natural DOM creates a major discrepancy in the comparison between pyrogenic and natural DOM. There are multiple reasons: 1) natural DOM samples were extracted with PPL (Garayburu-Caruso et al., 2020), which introduces a specific bias. 2) Only a specific fraction of molecules ionizes in negative ESI, which is the ionization used for the natural DOM dataset (Garayburu-Caruso et al., 2020). Negative ESI introduces another very specific bias onto natural DOM data. These biases are not accounted for in the structural pyDOM dataset. 3) There are entirely different data analysis routines for the natural and pyrogenic datasets (structure vs formula assignment). These routines can also possibly skew the analyzed data differently. For example, it is common to filter assigned molecular formulas following the constraints published by (Stubbins et al., 2010). If such step was done on the natural DOM data, should be also done on the pyDOM data.

Thank you for this comment. We realize the following sentence is misleading, and we plan to delete it from the manuscript: “While we recognize that recent research has applied new technologies to inferring PyOM compound presence in environmental samples (e.g., FTICR-MS), there remains high uncertainty in the confidence of formula assignment and structural information with some of these techniques.” As this reviewer rightly points out, this is an oversimplification of the nuances of our dataset and approach.

We acknowledge there are biases that result from sample preparation for FTICR-MS, as well as in the technology itself. However, it remains the gold standard for characterizing the composition of natural organic pools (Bahureksa et al. 2021). Because we are focused on inferring the bioavailability of OM, we used a reference FTICR-MS dataset of the water-soluble portion of OM which should reflect OM that is most readily available for microbial consumption (Bahureksa et al. 2021, Tfaily et al. 2017). It is well-described that compounds in water-soluble OM as detected by FTICR-MS on surface waters and sediments correspond to microbial decomposition processes (Graham et al. 2017, 2018, Garayburu-Caruso et al 2021). Therefore, we feel the NOM dataset used in this study is an adequate method to describe the composition of bioavailable OM.

The reviewer also correctly notes that PPL cartridges selectively remove small organic molecules and that the detectable mass range of many FTICR-MS protocols also excludes
small molecules. These biases should result in a smaller range of bioavailability of OM detected by FTICR-MS than of NOM pools. Given this and the demonstrated correspondence of FTICR-MS-detected compounds with microbial processes, we feel that we can satisfactorily reach the overarching conclusion of PyOM bioavailability falling within the range of bioavailability of NOM pools.

We will realize the language we used in the manuscript may not adequately reflect these nuances, and we will update the text accordingly. Per our comments below, we will also add some data of PyOM pools as detected by FTICR-MS to our analysis in order to provide a directly comparable dataset.


3. The environmental DOM dataset contains sediment water extracts. How are these relevant to aquatic microbial processes which are primarily occurring in surface waters away from sediments? I suggest removing the sediment DOM and only keeping the surficial DOM.

While we agree that many researchers have historically focused on surface water OM, there has been increasing interest in processes that occur in subsurface hyporheic zones. These zones can contribute up to 96% of respiration in some streams (Naegeli et al. 1997). The microorganisms that catalyze OM decomposition are disproportionately bound to sediments in these zones. Therefore, we feel that sediment OM is an important consideration in the biodegradation of riverine OM. We also note that our analysis of sediment OM is separate from surficial OM and does not detract from that comparison. Based on reviewer 2’s positive comments about including sediment OM, we plan to keep this comparison in the manuscript.

4. The environmental DOM dataset is from only one study (Garayburu-Caruso et al., 2020). The authors compare pyDOM data from multiple studies to natural DOM from one study. Is the data in Garayburu-Caruso et al. (2020) representative of multiple different aquatic environments?

We note that the Garayburu-Caruso dataset was explicitly designed as a global survey and to capture the range of variability in global river OM. It involved a global crowdsourcing initiative that included over 97 river corridors in 8 countries within a 6-week period, from 29 July to 19 September 2019, spanning most major biomes (i.e., desert, tropical, temperate forests) and stream orders with various morphologies. All data in Garayburu-Caruso et al. were collected via standardized materials and sample procedures and analyzed by FTICR-MS at the same time, on the same instrument, using the same settings. Because of this and the challenges that this reviewer notes below in comparing FTICR-MS data from different labs with different procedures and instruments, we feel this is an excellent dataset for describing global OM pools via FTICR-MS.

In summary, the comparison between pyrogenic and natural DOM seems like “apples to oranges” at present. All or some of the reasons listed above may skew the bioavailability of the two datasets to make them falsely appear with comparable bioavailability. This flaw needs to be addressed to be corrected in order to make the datasets comparable.

We acknowledge the discrepancies in comparisons. To address this, we will include additional PyOM data derived from FTICR-MS. We note that the modelling approach we used is directly comparable across FTICR-MS (aside from measurement bias) and known compounds because it only considers molecular formula. Regardless of if a molecule has structural and formula information, as in the case of known molecules, structural information is not considered. This was a key factor in why this modelling approach was chosen. We also note that biases introduced by FTICR-MS should narrow the range of variability in our modelling approach, not widen it, by excluding components of NOM. The FTICR-MS procedure used by Garayburu-Caruso et al. also has been experimentally associated with biological activity. Given this, we feel that with the addition of some PyOM FTICR-MS data for validation, we can satisfactorily reach the overarching conclusion of PyOM bioavailability falling within the range of bioavailability of NOM pools.

Suggestion for fixing this:

For natural DOM: add data from other studies and remove the sediment DOM. There is an overwhelming amount of FT-ICR-MS data published and provided in repositories. I am also sure that many research groups will be completely open to share data with you for this novel study. I do recommend mixing various surficial aquatic systems, primarily rivers, but also hopefully you can add lakes, wetlands, marine, etc. The choice of data will then allow you to determine if you can make claims strictly related to riverine environments or more like the global aquatic environment. Maybe you can compare pyDOM bioavailability to availability of DOM from different aquatic systems?

We note that this is a short-format publication with a strict maximum word count of 2,500 words. This inherently limits the breadth and depth of the analysis we can perform. Because of that, we have limited this analysis to rivers only. We acknowledge that it would be very interesting to
extend this analysis to other aquatic ecosystems in further work. As described above, we carefully chose our reference dataset as the Garayburu-Caruso et al. publication to have a comprehensive comparison with minimal biases as described above. For greater clarity, we will provide a map of sampling locations in the SI.

For pyDOM: I recommend using FT-ICR-MS data for ensuring comparability. Tracing truly pyrogenic molecules in natural systems is at present very challenging, so I recommend using charcoal water-extracts. There is a good number of studies that have published such data: (Chen et al., 2022; Goranov et al., 2020; Goranov et al., 2022; McKenna et al., 2021; Smith et al., 2016; Wagner et al., 2017; Ward et al., 2014; Wozniak et al., 2020; Yan et al., 2022) – just a few of the top of my citation manager.

While we have included some suggested formulae in the current manuscript extracted from tables in published works (n = 67 from Hockaday et al., 2006; Wagner et al., 2017), we agree that this paper would benefit from additional FTICR-MS datasets for the PyOM comparison. Thank you for the useful citations. We will pull vegetation char-derived spectra from these papers and include them for comparison.

I think by doing this you will achieve complete comparability (ensuring data were from PPL extracts and -ESI). I can foresee one issue – if obtaining data from multiple groups, you might get molecular formulas which could be biased by the different software that groups use (ICBM, pyKrev, Formularity, etc.). What I suggest is inquiring for peak lists (m/z and intensity data) and you work up the data yourselves to avoid comparability issues from different processing routines.

We will inquire about getting peak lists as several suggested papers do not have links to their m/z data in the publication. Several of the papers do include tables of assigned formulas with m/z included. Thank you!

Detailed Review and Specific Comments:

Abstract: Excellent. Gives a comprehensive overview of the study, information is succinctly presented. I only recommend adding one sentence somewhere around lines 29-31 that explains that this model is a computational approach using molecular formula data from mass spectrometry or molecular formulas of known DOM structures. Readers who are not familiar with this “substrate-explicit model” will likely be confused, so enhancement in clarity is needed.

Thank you. We will do this.

Intro: Very good. Establishes the importance of wildfires and pyrogenic DOM, provides background on the modeling approach and authors identify the gaps in our knowledge of pyrogenic DOM. Authors also establish a clear objective for the study. Some minor revisions are needed:

- Terminology throughout the manuscript:
  - Line 45: River corridor biogeochemistry. From the abstract and title of the paper it seems that the study is going to make claims regarding broad DOM
across various aquatic systems, not just fluvial ones. I suggest replacing with “aquatic biogeochemistry” throughout the manuscript.

Thank you. We will do this.

- pyOM is commonly used for particulate OM whereas pyDOM is used for dissolved OM. I advise using pyDOM as using pyOM for an aquatic study is confusing.

Thank you. We recognize that the nomenclature surrounding pyrogenic organic matter is quite varied and often misleading (Zimmerman and Mitra, 2017). We will revise the nomenclature throughout the manuscript regarding pyrogenic dissolved organic matter, and define terms and their usage in the text upon first instance.


- This term needs to be more clearly defined. Authors should also consider using “biolability” as one may argue that all molecules surrounding microbes are available to the microbes, but only some molecules can be uptaken and consumed (i.e., labile to biodegradation).

Thank you for this comment. The term ‘bioavailability’ is used throughout biogeochemical literature to include both microbial access to OM and the ability for it to be consumed. We recognize biolability is preferred by some fields of research, so we will include both terms at first mention along with a clear definition of ‘bioavailability’.

- BC – having BC and pyOM in the same manuscript is confusing as many readers view them as synonymous. I recommend removing BC entirely. In studies employing molecular data (from FT-ICR-MS, etc.) using a carbon term is also not appropriate as the data is reported for the whole molecules (i.e., on matter-basis) and not just for the carbon backbone (i.e., not on a carbon-basis). In simple terms, FT-ICR-MS measures DOM, not DOC. I recommend using terms that directly correspond to the structure (i.e., condensed aromatic compounds, ConAC, polycyclic aromatics, PCA, oxygenated PAHs, OPAH, or others that are used among different research groups).

Thank you for this comment. As outlined above, we recognize the considerable variability in terms used for pyrogenic derived organic matter and the need for more standardized vocabulary across studies. We agree with the reviewer that BC is not synonymous with PyOM, yet many readers may assume that it is. By adding in FTICR-MS datasets, we do not feel that we can assign structural information to pyrogenic assignments from this technique as the reviewer is suggesting, but we do agree we can be more specific with the continuum of the types of molecules this diverse pool contains. We will modify the language throughout the manuscript with these comments in mind, and will define every term utilized for PyOM up front to reduce confusion.

- “Natural” DOM. I recognize the necessity to use this term, the problem with it is that some pyrogenic molecules exist in detectable quantities in natural DOM (Goranov et al., 2022). Or they share the same molecular formulas but are different isomers.
Maybe authors should acknowledge this and clearly state in the intro that natural DOM corresponds to environmentally ubiquitous molecules obtained from fieldwork whereas the pyDOM dataset is from laboratory experiments and extractions to ensure these molecules are truly pyrogenic. One side question – is there any overlap (i.e., common formulas) among the natural and pyrogenic DOM datasets that the authors used? Are those common formulas removed or they were kept in? Please clarify this in the methods section.

We will update the introduction to reflect this point: “clearly state in the intro that natural DOM corresponds to environmentally ubiquitous molecules obtained from fieldwork whereas the pyDOM dataset is from laboratory experiments and extractions to ensure these molecules are truly pyrogenic”.

We will also report any overlaps in the molecular formula between NOM and PyOM. We did not previously assess this, so overlapping molecules were not removed.

- Lines 66-68. Authors make the claim that “Yet, there has been no systematic evaluation of the bioavailability of different constituents within the heterogeneous compounds that comprise PyOM”. This was actually recently done by two separate research groups (Bostick et al., 2021; Chen et al., 2022; Goranov et al., 2022) who looked at various constituents of pyrogenic DOM using various analytical approaches. Those studies also use “true” pyDOM from biochar extracts similar to Norwood et al. 2013, but of more wildfire-representative temperatures.

We agree there should be more nuance to this statement. We will update the wording to be more accurate.

- Lines 69-72. I think it would be good to explain the reason for this discrepancy – most research to date assumed that pyOM/pyDOM are only comprised of condensed structures, and the latter are indeed bio-refractory (Bostick et al., 2021), but the rest of pyOM/pyDOM is likely not.

Thank you for identifying this. We will provide clarity.

- Please expand on the substrate-explicit model text after line 82. This is the first time to my knowledge that this modeling approach is used in the wildfire biogeochemistry literature – most readers will have no background in it. Tell us about the model output: on line 99 you mention ΔG<sub>Cox</sub>, λ, and CUE for the first time without providing any info on what they are. There is a lot of great text in the materials and methods that does not read like M&M text – move some of it to the intro to provide a foundation on what these parameters are.

We will reorganize the manuscript to move some of the information from the M&M about modelling into the introduction.

Results and Discussion: Excellent. Results are properly discussed in the context of previously existing literature. Several comments:

- We know that behind a m/z value (or a molecular formula) there could be multiple isomers with the same elemental make-up (Leyva et al., 2019). Different isomers will have different thermodynamic properties. Do the mathematics behind the model
consider this? This needs to be discussed in the last section where authors discuss the limitations.

We agree that there are often multiple isomers with the same formula. The modelling approach we used was designed specifically with this limitation of FTICR-MS data in mind. It only considers molecular formula and does not incorporate structural information. To our knowledge, the is the only extant modelling approach to predict aerobic metabolism rates from FTICR-MS data. It is now incorporated into the US Department of Energy’s KBase modelling ecosystem and seeing increasing usage for providing meaningful predictions from FTICR-MS data in other contexts. We will clarify this in the text.

- Does the model account for potential toxicity of molecules? Some pyrogenic molecules are toxic (Smith et al., 2016), but not sure if a mathematical proxy for toxicity can be extrapolated from previous studies and incorporated into your computation approach. This is probably something to consider and discuss in your last section too.

Toxicity is not considered in the model. It is an interesting topic but is beyond the scope of this paper.

- Figure 1. Plotting \( \lambda \) versus \( \Delta G_{\text{Cox}} \) is not very intuitive to a broader audience. Though I recognize that these plots are useful, someone who is not experienced with these parameters will be confused. I recommend improving the clarity in visualizing the model output. I recommend complementing these plots (or substituting them) with something more recognized such as van Krevelen diagrams. I suggest plotting van Krevelens for river DOM, sediment DOM, and pyDOM and color coding the markers based on lambda and/or \( \Delta G_{\text{Cox}} \). This way different compound classes can be easily identified in the H/C vs O/C space and we can see how energy content and metabolic efficiency vary per compound class (condensed molecules, phenols, lipids, etc.).

We agree with the suggestion to add more common representations of data. We will add extra panels to Figure 1 and add supplemental figures as needed to show Van Krevelen plots.

- It will be good to compare the model output *computational* results with previously published bio-degradation *experimental* studies. Bostick et al. 2021, Chen et al. 2022 and Norwood et al. 2013 provide degradation rate constants for pyDOM. Can the authors compute a similar bulk degradation rate constant using the output of their model and see how computational and experimental results compare? This will strengthen the conclusions of this manuscript.

Thank you for the suggestion. The suggested papers report change or percent change in concentration through time, not a true rate constant (i.e., accounting for non-linearity through curve fitting). This discrepancy in units makes empirical comparisons difficult. We discuss the findings of Norwood et al and we will add conceptual comparisons to Bostick et al 2021 and Chen et al 2022 to the discussion.

Methods: Excellent text, but a lot of it reads like an introduction. Please move a lot of this text to other sections.

Thank you. We will do this.
Graphical Abstract: Absolutely gorgeous design. I would only recommend enhancing font sizes (difficult to read) and not using gray color – the labels (pyOM pool, Biomass, etc.) and other subfigures (e.g., biomass particles with tails) are difficult to read/see. I suggest just converting everything to bold black and increasing the sizes (just like CO2). Also why is “Biomass” looking like bacteria? I recommend replacing Biomass with “Microbes”.

Thank you. We will increase the font size and adjust the colors. The “biomass” is reflective of microbial biomass. We will clarify this label.

Title: It is a representative title, but at present it reads a bit awkwardly. I read it as “bioavailability resembles DOM pools”, which is odd. Consider rephrasing into something like “Computational modeling reveals that molecules in pyrogenic and natural dissolved organic matter pools have similar bio-lability” or something like that.

We will revise the title to read more clearly.

References


Molecular Comparison of Solvent Extraction and Water-Soluble Fractions of Biochar by FT-ICR Mass Spectrometry. Analytical Chemistry.


