Review of Houston et al. "Old Carbon, New Insights: Thermal Reactivity and Bioavailability of Saltmarsh Soils" (Biogeosciences; https://doi.org/10.5194/egusphere-2024-3281)

Synopsis

This is a revised version of a previous manuscript focusing on Ramped Oxidation (RO) ^{14}C activities (reported as "percent modern" or pMC) and $\delta^{13}C$ values for organic carbon (OC) from a set of saltmarsh soil cores from the Skinflats saltmarsh in Scotland, UK. Based on my comments and those of a second, anonymous reviewer, the authors have made considerable changes to the revised version. I believe this revised version represents a significant improvement, but I still have several issues, particularly related to the treatment and presentation of the inverse model results. I highlight these in detail below---it is not clear to me that the authors fully comprehend what is being calculated and reported in these inversions. Only after implementing these further changes would I then support publication in *Biogeosciences*. I believe this will require one more round of review. Please do not hesitate to contact me regarding any questions on this review.

Sincerely,

Jordon Hemingway
jordon.hemingway@eaps.ethz.ch

(there are no line numbers given for the abstract, so I will just write my comments here and the authors can find the relevant lines).

- "…driven by the net contribution from the older fraction…": What does this mean? Are the authors saying that preserved OC in salt marshes is generally low in ^{14}C activity?

- "We also present the first evidence to supprt…": I'll admit that I'm not very well-versed in the MRV side of CDR, but it seems wild to me to claim that thermally labile OC that is currently preserved in salt marshes could count as *additional* CO_2 removal. This is carbon that is already sequestered. I understand that draining and disturbing these salt marshes would lead to *remineralization* of this OC, thus increasing CO_2 emissions, but doing nothing will not lead to any *additional* CO_2 being removed from the atmosphere. What am I missing here?

(here beings the line numbers)

L27: Hemingway et al. (2017) used an oxidizing carrier gas and should thus be cited along with Plante et al. and Stoner et al., not with Rosenheim et al.

L30: "CO_2 evolved at low temperatures is deemd to be from… pools with greater thermal lability than CO_2 evolved at higher temperatures". Yes, of course it is---this is the definition of thermal lability! I don't see what sentences like this are adding.

L34 (and throughout): change to "^{14}C *activity*", as it is a radioactive isotope.

L57-59: "Crucially, the biological availability… depends on… thermal reactivity". This is not true. Biological availability may *correlate* with thermal reactivity, but it does not depend on it *per se*.

L108-109: "stream of high purity oxygen": I suppose I didn't realize in the first round of review that this is indeed a pure O_2 stream (I now dug into the Gartnett et al. 2023 paper). Given this, do the authors think this difference in carrier gas will impact thermogram shape relative to other systems? Have they compared a reference material using their setup vs. using the setup at NOSAMS, ETH, etc.? It would be really nice to see the inter-laboratory reproducibility of this instrument (not just internal reproducibility, which looks quite nice in Gartnett et al.). For reference, most other systems use O_2 in He. As a starting point for this comparison, the authors could look into Bolandini et al. (2025) *Radiocarbon* (https://doi.org/10.1017/RDC.2025.6), who investigated the impact of O_2 flow rate on thermogram shape for the ETH instrument.

L151-153: Fair point to the authors in their response that this is indeed how most labs used to do a ^{13}C correction prior to AMS instruments including a ^{13}C cup---I did not realize this is still how things are doing at SUERC. Still, I note that the authors mis-cited McIntyre et al. (2017) in their response, who indeed used the internal ^{13}C correction of the MiCaDaS system, as is common practice at ETH (that paper instead focuses on the in-line EA-IRMS-AMS for single-analysis %OC, δ^{13}C, and F^{14}C analysis).

L159-162 (and thorughout): Please use μ_E, σ_E, and $p(0, E)$ nomenclature.

L162-166 (and Table A2): I don't understand this---one can simply tell the software to perform a blank correction or not using the "blank_corr" flag (see documentation). How do the "^{13}C values generated [vary] significantly from our IRMS measured values"? If the software does not perform a blank correction, then the $\delta^{13}C$ values used by the software are simply identical to the ones inputted by the user---there is nothing to be "modelled" here. I don't understand how the numbers in Table A2 were generated.

Fig. 1-2/Table 1-4: (I'm not sure exactly where to put this comment, so I will put it at the first place that I think is relevant, which is Fig. 1.) Here, the authors need to report much more information related to the inversion before these results can be interpreted. For example:

- what regularization values, λ, were used for each sample?
- How do the resulting $p(0, E)$ distributions for each sample look? The authors show the thermograms, but never show $p(0, E)$ distributions. It is difficult to judge results without seeing the distributions themselves. This is particularly the case since some of the samples appear to not reach baseline at high temperature (e.g., T3 5.5, T2 15.5). This is important as it is known that the inversion is sensitive to boundary effects, so baseline must by reached or forced (see documentation).
- It would be incredibly useful for the subsequent discussion to know what fraction of each sample is contained within each thermal window. For example, is the 150-325 °C fraction 10% of total C? 20%? This sould be added, e.g., to Table 1 or 2.
- I am again missing the a comparison of mass-weighted RO results vs. bulk measured results. That is, if you simply sum the $F^{14}C$ or $\delta^{13}C$ values for each thermal window weighted by the fraction of total carbon within that thermal window, do you recreate the measured bulk values within uncertainty? This is again an important "sanity check" and can easily be added, e.g., to Tables 1-2 (or Table A3).
- In Table 3, what does the $p(0, E)$ column mean? $p(0, E)$ is a probability density function---that is, a distribution whose integral is equal to unity. It is not a single scalar number. I don't know what 0.02, 0.02, 0.01, etc. refer to. Is this the maximum value in the $p(0, E)$ distribution? But this is arbitrary and depends on the size of the discretized energy step, ΔE…
- Table 4 (and throughout): please update the nomenclature so that it is clear to the reader when the authors are referring to μ_E, σ_E, and $p(0, E)$ of the entire sample vs. for a given thermal window; following previous studies, I recommend μ_E, σ_E, and $p(0, E)$ when referring to the bulk sample and $\mu_{f,E}$, $\sigma_{f,E}$, and $\Pi_f(E)$ when referring to a given thermal window, f.

L222-223: "…no significant changes in μ_E, σ_E, nor activation energy distribtuion ($p(0, E)$)". But $p(0, E)$ is a distribution, not a scalar value, so how can it be compared across samples in the same way as the other metrics?

L232: "…$\mu_{f,E}$ [here using my recommended nomenclature]… increased sequentially…we therefore infer that the thermal recalcitrance of RO fractions is greater at higher temperatures…". As for my comment on L30, of course it is---this is the definition of thermal recalcitrance! Higher temperature thermal fractions will have a higher $\mu_{f,E}$ value by definition---there is nothing to infer!

L252: What is meant by "…wider μ_E ranges…compared to the first three RO fractions"? $\mu_{f,E}$ [here using my recommended nomenclature] is a mean value so it cannot have a "range". Do the authors mean that the *difference* in $\mu_{f,E}$ is greater between the highest two thermal windows than between the lowest three? But if so, then this is simply a function of the chosen temperature windows and doesn't say anything inherent about the OC being combusted.

L253-254: "…may have been caused by non-first order decomposition of carbonates". How would this cause "wider μ_E ranges… and increased bond strength diversity"? Non-first order behavior implies that the resulting thermogram (and thus $p(0, E)$) shape depends on the mass of sample loaded into the instrument (c.f., Fig. 4d of Hemingway et al. (2017) *Biogeosciences*).

L255: change "*rampedpyrox* model" to "distributed actvation energy model"; *rampedpyrox* is simply the name of the python package.

L298-299: "…although the thermal reactivity of OC decreases with ^{14}C content…" The cause-and-effect should be flipped here: ^{14}C activity decreases with decreasing thermal reactivity (i.e., thermal reactivity is the independent variable).

L344-347: Here the logic seems to be: (i) low-*E* components are consumed prior to deposition in the Skinflats; this leads to (ii) thermally recalcitrant material being deposited and thus (iii) no change with depth in the salt marsh (i.e., due to no further remineralization). But this somewhat contradicts the thermograms shown in Fig. 1 (and presumably the corresponding $p(0, E)$ distributions, if they were shown), which shows a fair amount of carbon in the ~200-400 °C range. This is quite thermally labile. In fact, one *does* observe a decrease in the peak height of the ~250 °C peak with depth (relative to the ~450 °C peak). This instead points to a continued remineralization of this thermally labile material with burial depth. This is one instance where it would be very useful to know what fraction of total carbon is contained within each thermal window, as this could then be easily quantified. An alternative appraoch is to use the fraction of total carbon contained in "low-*E*", "middle-*E*", and "high-*E*" bins, as was done for example in Hemingway et al. (2018) *Science* (https://doi.org/10.1126/science.aao6463).

L360-361: "…more energy is required …to decompose older…carbon than younger…carbon". Careful with statements like this; thermal activation energy is merely an analytical tool to separate carbon; there is no requirement that older carbon necessarily has a higher thermal recalcitrance.

L365-266: Schmidt et al. (2011) do not mention thermal reactivity. Also, biological turnover time does not *depend* on thermal reactivity *per se*; the latter is merely an analytical tool.

L380-382: This statement is well-known in the RPO literature (see, e.g., some of the initial papers from the Rosenheim group that focused on using RPO as a means of dating sediments).

L389-390: I don't think you can say that, "…the biologically evolved CO_2…was therefore not from a thermally labile OC pool." It very well could have been from a labile pool if said pool was composed of several compounds of different ^{14}C ages, as is likely.

L399-401: I don't know what this sentence is trying to say. Reword.

L432-433: This relates to my comment in the abstract, but can this *really* be considered as additional C storage for MRV? This carbon is already naturally sequestered…

L485-487: Same as previous comment.