Review Report

Recommendation: Major revision

This manuscript addresses an important and timely question in coastal wetland restoration: whether commonly used *Spartina alterniflora* management practices have unintended consequences for deep soil carbon stability. The study is potentially valuable because it goes beyond the surface soil and examines a 0–100 cm profile, compares two restoration interventions (plastic mulching and deep tillage), and attempts to integrate hydro-salinity, reactive Fe fractions, microbial biomass, bacterial community composition, and path modeling into a mechanistic framework. The topic is relevant to restoration ecology, blue carbon accounting, and post-invasion management.

Overall, I find the manuscript promising and potentially publishable after major revision. The comparison between PM and DT is interesting, and the depth-explicit perspective is a clear strength. However, the current version has several important issues related to experimental interpretation, internal consistency, and the strength of mechanistic inference. In particular, I think the authors need to more carefully separate what is directly shown by the data from what is hypothesized, and they need to resolve several inconsistencies between the abstract, methods, results, figures, and interpretation.

My first major concern is the experimental framing. The study compares CK, PM, and DT 18 months after treatment, but CK is described as an unremediated invaded stand, whereas PM and DT both involve prior vegetation removal and subsequent intervention. As a result, the present design does not cleanly isolate the effect of mulching or tillage alone; rather, it combines invasive plant removal and post-removal management. This should be acknowledged more explicitly in the interpretation. Relatedly, the methods describe a "randomized complete block design," but the blocks are not clearly defined in the text. The figure caption also refers to "three sampling sites," whereas the methods describe five replicate plots per treatment. The experimental layout therefore needs to be clarified much more precisely.

My second major concern is that the manuscript's strongest mechanistic claims about **deep microbial reactivation** are not fully supported by the microbial community data presented. The hypotheses explicitly focus on deep soils (30–100 cm), yet the bacterial sequencing results shown in the paper are only for the 0–10 cm and 10–20 cm layers. The discussion nevertheless uses these surface/subsurface patterns to infer the identity and role of deep microbial groups under PM. In its current form, the manuscript demonstrates increased deep-soil MBC under PM, but it does not directly demonstrate how bacterial community structure changed in the deep layers themselves. This gap should be acknowledged much more clearly, and the mechanistic language should be softened.

A related issue is the handling of microbial necromass and amino sugars. The abstract and conclusion state that "amino-sugar biomarkers indicated coherent shifts in microbial necromass C with Fe–OC dynamics," suggesting an important line of evidence for the paper's interpretation. However, in the methods and main results provided here, I could not find a dedicated amino-sugar methods section, analytical description, or corresponding results presentation. If these data are central to the argument, they need to be fully described and reported. If not, these statements should be removed or greatly reduced.

My third major concern is the carbon stock calculation and the interpretation of net carbon loss. The methods state that SOC stocks were calculated using both fixed-depth and equivalent soil mass approaches, but only the fixed-depth equation is shown, and the results section mainly discusses SOCS-FD. If bulk density differs among treatments, the ESM approach becomes especially important and should either be reported clearly or removed from the methods. In addition, SOC is estimated by loss-on-ignition whereas TC is measured by elemental analysis; the manuscript would benefit from a stronger justification of why LOI-based SOC is appropriate for these estuarine wetland soils and how the authors ensured comparability between the different C metrics. A supplementary worked example of the stock calculation would also improve transparency.

My fourth major concern is that the PLS-PM analysis is interpreted too strongly as evidence of causality. The authors use PLS-PM to argue that PM altered physicochemical stress, destabilized reactive Fe pools, and thereby amplified microbial decomposition and SOC/Fe-OC loss. This is a reasonable conceptual model, but the data remain observational within a single post-treatment sampling campaign. The discussion itself acknowledges that there are no pore-water chemistry data, no direct GHG fluxes, and no direct measurements of iron-reducing activity, so the present evidence is not sufficient to distinguish microbial mineralization from leaching/export or redistribution processes. I therefore encourage the authors to treat the path models as inferential and heuristic rather than strongly causal.

The manuscript also contains several internal inconsistencies that should be corrected before further consideration. For example, the methods describe figure 2h as MBC and 2i as MBN, but the results text refers to deep-soil MBC increases in Fig. 2i. Likewise, the text discussing Fed and Fep appears to cite figure panels inconsistently relative to the figure caption. In the alpha-diversity paragraph, the text states that PM reduced Shannon and Chao1 "($P < 0.05$)" but then says "although no significant differences were detected," which is contradictory. I also suggest revising some of the more dramatic wording, such as "catastrophic collapse" and "failure of the Iron Gate," to better match the tone expected in a research article.

Despite these concerns, I think the manuscript has several strengths. The depth-resolved framework is important, the comparison of PM and DT is potentially informative for management, and the authors are asking a meaningful question about whether restoration practices may mobilize legacy subsoil C. With substantial revision, particularly in clarifying design, tightening the evidence chain, and improving internal consistency, this study could make a useful contribution.

Specific suggestions for revision

1. Clarify the experimental design and the meaning of "randomized complete block design," including how many independent plots were sampled and how the layout in Fig. 1 corresponds to the methods.
2. Explicitly acknowledge that the design compares invaded CK with two post-removal management treatments, which limits strict attribution of treatment-specific effects.
3. Revise the deep microbial interpretation so that it is based on measured evidence; avoid inferring deep community composition from shallow sequencing alone.
4. Either fully describe and report the amino-sugar/necromass analysis or remove those claims from the abstract and conclusion.
5. Report the equivalent soil mass approach if it was actually used, and provide more transparent stock calculations.
6. Tone down causal language around PLS-PM and around unmeasured mechanisms such as Fe reduction, mineralization, and lateral export.
7. Carefully check all figure references, variable labels, and contradictory significance statements.