**RC2**

**General Comments**

This article reports the effects of ten years of experimental warming on soil carbon fractions at three depths from Blodgett Forest. Subsoil (80-90 cm depth) particulate organic matter (POM) mass was reduced and bulk soil carbon shifted composition toward more recalcitrant compounds (lignin and aromatic bonds), while mineral-associated organic matter (MAOM) remained unchanged, indicating that POM is more responsive to warming than MAOM.

The study has important implications for impacts of global change on soil carbon cycling, and represents a novel contribution to the field by reporting specific changes in carbon compound and organic matter composition with sustained warming.

However, there is a concerning over-interpretation of statistically non-significant results. The confidence intervals reveal massive uncertainty, which should preclude strong conclusions about subsoil carbon loss at specific depths. In addition, the reported marginally significant results are hard to justify within the confidence intervals that span a huge range of negative to positive values. At the very least, these parts of the results section should be reworded to reflect the high uncertainty instead of the supposedly "marginal significance." The statistical power of the comparisons is weak with n=3 samples throughout, and there were no reported corrections for the multiple comparisons made.

There is weak evidence for the bulk SOC composition shifts (Section 3.3, Fig. 3), which are based largely on visual PCA inspection. There are no formal statistical tests for whether warmed vs. control groups differ significantly in composition.

There are a few sentences in the results section that start with the word "basically," which detracts from the meaning of the sentence. Recommend removing or rephrasing.

Throughout the discussion section, my suggestion is to separate the pattern description from statistical inference.

The authors do a good job of putting the results from their study in the context of previous studies.

Throughout the manuscript, the authors explain topsoil patterns by inferring warming-induced increases in plant inputs (e.g., lines 377-384, 432-436, 485-487).

However, plant inputs were not measured in this study and the Ofiti et al. (2021) study they reference reported decreased root biomass at this site under warming, contradicting the proposed mechanism. Alternative explanations (moisture effects, substrate limitation, redistribution) receive insufficient consideration and the Abstract and key sections present this inference as established rather than hypothetical. I recommend revising to: (a) clearly identify increased inputs as a hypothesis rather than observation, (b) address the contradiction with Ofiti et al.'s root data, (c) discuss alternative mechanisms with appropriate weight, and (d) acknowledge this limitation explicitly. Consider tempering mechanistic conclusions in the Abstract and Conclusions to reflect this uncertainty.

Overall, the claim that MAOM remained stable under warming, across depths, is well-supported. The claim that POM is more responsive than MAOM to warming is somewhat supported, with a consistent pattern but mixed statistical significance. The claim that subsoil carbon loss is driven by POM depletion is overstated based on the evidence presented. While the bulk SOC interaction is significant, the individual subsoil depth reductions are not significant and the confidence intervals include substantial gains, not just losses. The claim that bulk SOC shifted toward lignin/aromatic bonds in warmed subsoils is also not well-supported, since it is based on visual interpretation of PCA only and Fig 3 shows trends but overlapping confidence ellipses.

We are deeply grateful for your thorough reading of the manuscript. Your critical and constructive feedback, and detailed, practical suggestions will significantly improve the quality of our manuscript. We will implement your suggestions accordingly.

We fully agree with your comments on wording in the results section, overestimating of certain statistical results, and will reword our interpretation of the potential increased plant inputs to the system from direct observations and instead highlight them as hypothesis. Accordingly, we will revise the manuscript based on your suggestions, find alternative explanations, and explicitly address the problems of uncertainties by small sample size, and the limitation of the study. Thanks for your comments.

We agree that adding statistical analysis on functional groups data would be beneficial. However, because DRIFTS data is only semi-quantitative, the reliability of statistical test would also be challenged, which is instead typically presented in principal

component analysis. However, to address your comment we will add a table into the supplementary information where we directly test the area under the curve values.

We also acknowledge that limited replication (n = 3 blocks) reduces statistical power, and results in our studies, meaning that marginal effects, should be interpreted with caution. However, large manipulation experiments such as the Blodgett Forest whole-soil warming experiment are typically limited in replication number. Because post hoc comparisons were only conducted following significant interaction terms and were restricted within depth increments, we did not apply additional multiple-comparison corrections, consistent with recommendations for hypothesis-driven mixed-effects modeling. We have chosen conservative statistical methods, which led to larger p-values compared to some other statistical methods to help support this.

**Specific Comments**

Were there changes in soil moisture with warming and by depth? This is important for the discussion of decomposition dynamics.

We will add this to discussion for example from Riley et al. (2025) and Soong et al. (2021). Briefly, warming significantly decreases soil moisture, and this reduction is most pronounced at topsoil (10-20 cm) and deep soil (80-90 cm; Pegoraro et al., 2024). The soil moisture at surface fluctuates magnificently between dry summer and moist winter (Soong et al., 2021).

L257-258: This could be more clearly worded to indicate that it is the interaction effect that is significant. Perhaps "Warming reduced SOC concentration in the subsoil but not the topsoil (warming × depth interaction: $p = 0.002$)."

We will simplify this sentence and other part of results accordingly.

L259-262: Reporting these statistics as marginal effects may be inaccurate, since the confidence interval is very large and the p-value close to non-significant. Recommend rephrasing to reflect the high uncertainty instead. Within that confidence interval, the effect could be anything from a huge loss to a moderate gain.

We will rephrase to highlight the uncertainties as follows:

At 60–70 and 80–90 cm, mean SOC concentrations were lower by 54% (CI: -83 %, 26 %) and 56% (-84 %, 20 %), respectively, but the wide confidence intervals and

marginally significant p-values ($p = 0.099$ and $p = 0.086$, respectively) indicate substantial variability rather than a definitive loss.

L271: Awkward/unclear wording “affected distinguishably”

We will clarify the wording as follows:

Depth significantly reduced fPOM concentrations ($p < 0.001$), while the impact of the treatment differed across depths ($p = 0.001$).

L271-273: Again, the large confidence intervals detract from the significance that is claimed. A non-significant increase could actually be a decrease, increase, or major increase, but the directionality of it is highly uncertain.

We will rephrase the sentence as follows:

At 10-20 and 40-50 cm, warming on average increased fPOM concentration respectively by 56 % and 97 %, but these increases were weak due to large standard errors.

L292-294: Unclear what is meant by “main effects” in this sentence.

We will simplify it by deleting „and main effects “

L314-318: "warmed subsoils displayed a trend towards increased AUC values" – this is descriptive, not statistically validated

We will now support this with statistical evidence in the supplementary information.

L334: Specify which PCA plot is being referenced.

We will implement the change.

L348-349: Reword for clarity.

We will reword as follows:

We did not find significant effects of warming, but significant effects of depth ($p < 0.001$; Fig. 5) on the bulk soil DRIFT stability index (DSI).

DSI analysis (Section 3.5) shows no significant warming effect for bulk soil despite claims.

We will change the title accordingly.

L355-361: Any mention of significance or non-significance should be accompanied by a p-value or p-value range.

We will add the p-value range to non-significance.

L382: If the relationship wasn't significant, it should not be interpreted as though it was. It is also unclear whether the non-significance is related to just one or more of those relationships listed.

We will rephrase and focus on the significant results.

L434: There is nothing in the code/data availability section. However, the authors are commended on their inclusion of significant amounts of data in the supplemental.

Thank you for your comment, we haven't added this yet. However, we can confirm that all data will be made available through the ESS-dive website as is customary for data published in collaboration with the Blodgett Forest whole-soil warming project.

**Technical Revisions**

L26: period instead of comma or else missing capitalization

L375: Significant not significantly

L396: Subsoil not subsoils

We will implement all the three suggestions.