UF UNIVERSITY of FLORIDA

College of Agricultural and Life Sciences
Department of Soil, Water, and Ecosystem Sciences

2181 McCarty Hall A
PO Box 110290
Gainesville, FL 32611-0290
352-294-3125
swarnali.mahmood@ufl.edu
swarnali.mahmood@gmail.com

Dr. Jose Alfonso Gomez
Topic Editor, *SOIL*
Institute for Sustainable Agriculture, Spanish National Research Council (IAS- CSIC), Córdoba, Spain

August 16, 2026

Dear Dr. Gomez,

Thank you for considering our manuscript, "*Long-term storage of air-dried samples compromises water-extractable organic carbon as a soil health indicator*" for publication in *SOIL*. We appreciate the comments provided by the reviewers in open discussion.

We have addressed all the concerns the reviewers raised. Specifically, we have clarified our soil reanalysis timing, air-drying methods, and residual moisture contents of the cover crop study archived samples. We have also highlighted methodological limitations and interpretational uncertainties associated with POX-C and soil storage as suggested. In addition, we have clarified that the study isolates storage effects, not field temporal trends. We have also emphasized that our study quantifies the magnitude and consistency of the changes in WEOC pool across multiple long-term experiments, which remain underexplored.

Please contact us if you have any questions. Thank you for your consideration of our manuscript.

Sincerely,

Swarnali Mahmood

Swarnali Mahmood, Ph.D.
On behalf of all the authors

COMMENTS FROM REVIEWER 1

The manuscript “Long-term storage of air-dried samples compromises water-extractable organic carbon as a soil health indicator” authored by S. Mahmood and co-workers focused on the evolution with time of soil indicators with time of storage. The data refer to a common practice of “air drying” and storing at “room temperature”. Indeed, room temperatures are varying a lot across different locations and seasons, as for room moisture. In my opinion, air drying is rarely reaching moisture content below 5-10% in soils (at least in my region). This means that microbial communities will still have some residual activities at these moisture levels and room temperature, which is the main reason behind the modification of labile soil health indicators such as water-soluble C and N. For “labile” soil health indicators, other methods of drying (at controlled temperature and moisture) and storage (at colder temperature) are preferably, and this could be mentioned in the manuscript. Therefore, results from this study are relevant, but their novelty is relatively low.

Thank you for your feedback. We agree with the reviewer that residual moisture in air-dried soils can sustain microbial activity and contribute to changes in labile C pools during storage.

In the revised Introduction, we emphasized that “Air-drying prior to analysis and long-term storage is a standard and widely adopted procedure in soil science for both routine laboratory analyses and soil archiving (Kühnel et al., 2019; Stott, 2019). Despite being classified as air-dried, soils typically retain a small amount of residual water (0.3-6%) that remains adsorbed to mineral surfaces and within fine pores, particularly in clay-rich and organic matter-rich soils. Residual moisture content has been shown to be strongly related to soil texture and organic matter content, reflecting water retained at high matric potentials and on reactive mineral surfaces (Poeplau et al., 2015; Wäldchen et al., 2012).” (lines 38-44). We believe these revisions further corroborate the knowledge gap currently listed: “the combined effects of management and multi-year storage of air-dried soils on the repeatability and temporal trajectories of soil health indicators remain poorly understood.” (lines 55-57).

To address the reviewer’s concern, we expanded the Materials and Methods section to provide additional details regarding sample preparation and storage as follows: “Soil samples were air-dried in a climate-controlled laboratory on open trays until a constant mass was achieved before sieving and archiving. Samples were then stored in airtight plastic bags at controlled room temperature (i.e., 24°C) in the lab.” (lines 75-76).

In the Results section, we reported the residual moisture content measured in a subset of air-dried samples from our cover crop study, which “averaged 0.26% (n = 9; range: 0.20-0.32; CV = 16%)” (lines 134-135). We also discussed implications of these measurements: “This value is consistent with expectations for coarse-textured sandy soils, which retain substantially less residual water than fine-textured soils (Poeplau et al., 2015; Wäldchen et al., 2012)” (lines 135-136). We also discussed the potential contribution of residual moisture to variation among archived samples: “Microbial activity may persist even at the low water contents present in air-dried soils, potentially resulting in continued transformation of labile carbon and nitrogen pools during storage.” (lines 44-45). We believe the revisions complement existing discussion on how microbial dynamics, such as cell lysis and continued enzymatic depolymerization, could contribute to the observed accumulation of WEOC (lines 149-152). We also acknowledge that “For these labile indicators, drying at controlled temperature and/or cold storage could be further explored.” (lines 207-208).

Another aspect is that permanganate oxidizable carbon (POX-C) is proposed as a more stable indicator, as suggested by USDA, but there are many concerns about the methods used for measuring POX-C, and this is not reported in the document. I suggest to include some of the studies questioning POX-C used methods to give a better overview of this indicator, as those by Margenot's group Margenot et al., 2024 (https://doi.org/10.1002/ael2.20124).

> ***We appreciate this important suggestion. We acknowledged the methodological limitations and interpretational uncertainties associated with POX-C in our Introduction. "In contrast, POX-C is generally less sensitive to drying but may still be affected by methodological factors such as soil mass, sieve size, laboratory handling, and SOC content (Gasch et al., 2020; Wade et al., 2020)." (lines 53-55). In our revised manuscript, we presented POX-C more cautiously as a relatively stable but method-dependent indicator highlighting recent studies. We added in our Discussion that "Recent studies have questioned its interpretation as an indicator of biologically active carbon (Christy et al., 2023; Vellenga et al., 2025; Woodings and Margenot, 2023)." (lines 172-174). "POX-C pool remained stable during multi-year storage, underscoring its potential for assessing long-term management impacts in archived soils. This offers new insights into the ongoing debate on the interpretability and practicality of POX-C as a soil health indicator (Gasch et al., 2020; Margenot et al., 2024; Woodings and Margenot, 2023)." (lines 208-211).***

Lastly, the presentation of the experimental design, which comprises sample sets from 5 experiment, should be improved to clarify when samples were collected, when they were analysed the first time, when they where reanalysed and the variation occurred during the years of storage.

> ***We revised the Materials and Methods section to clarify that the comparisons are within-sample (i.e., archived vs. original measurements), rather than across time or management treatments, in the cover crop study. The changes in these indicators were quantified as differences between original and reanalyzed values from the same archived samples, thereby isolating storage effects from field-driven temporal variability. This design is different from those of the other long-term experiments where we could not separate the storage effects from those due to management practices.***

Abstract: line 19 "four long term agricultural sites" is not clear if it is referred to comparison of long-term storage or real field experiments in which WEOC is decreasing. Please revise.

> ***We revised lines 19-20 for clarification as: "In addition, WEOC concentrations consistently declined over time in this experiment along with four long-term agricultural sites in the USA, but their bulk soil organic carbon (SOC) or soil organic matter (SOM) did not." (lines 19-20).***

Lines 43-49: see my comment above about POX-C methodological problems

> ***We highlighted these limitations in the revised discussion as per your suggestions as follows: "Recent studies have questioned its interpretation as an indicator of biologically active carbon (Christy et al., 2023; Vellenga et al., 2025; Woodings and Margenot, 2023)." (lines 172-174). "POX-C pool remained stable during multi-year storage, underscoring its potential for assessing long-term management impacts in archived soils. This offers new insights into the ongoing debate on the interpretability and practicality of POX-C as a soil health indicator (Gasch et al., 2020; Margenot et al., 2024; Woodings and Margenot, 2023)" (lines 208-211).***

Lines 51-54: Why Florida and Midwest experiments are listed separately instead of simply saying "five long term experiments"? This will make things easier to understand, in my opinion.

The Florida experiment follows a different design from the other four long-term experiments. As discussed earlier, we were able to isolate storage effects in the Florida experiment by analyzing the same samples before and after storage. In the other experiments, however, storage effects were confounded with management effects.

Based on your comment and those of the other reviewer, we clarified our rationale in the Introduction: "we retrieved archived soils from a cover crop management study in Florida and compared the values of WEOC, WEN, Cmin, and POX-C before and after multi-year storage. This Controlled Storage Experiment directly quantified the storage effects by comparing original and reanalyzed measurements from the same archived samples. We hypothesized that multi-year storage would increase the concentrations of WEOC and WEN. We then tested whether these storage effects could be detected in other systems in a Multi-site Validation Experiment. Using archived soils from the Florida study and four long-term agricultural trials, we compared the temporal trajectories of WEOC with those of bulk C in samples collected throughout each experiment. Because bulk C is robust against long-term storage (Blake et al., 2000; Kühnel et al., 2019), we hypothesized that storage effects would cause WEOC trajectories to diverge from those of bulk C." (lines 58-66).

Sections 2.2 and 2.3 and elsewhere: It is not clear the set-up of the re-analysis of stored samples. Which were the reference values for the samples stored from the other 4 long-term experiments? The comparison was done between the re-analysis values of stored samples against the freshly analysed soils?

Section 2.2 and 2.3: We clarified in the revision that WEOC concentrations from the four long-term experiments were "determined using the same extraction protocol in 2023" in the same laboratory" (lines 109-111).

Line 101: sampling time or storage duration?

As shown in Figure 2, "in the Multi-site Validation Experiment, we compared the temporal trajectories of the SOM and WEOC values reported in Nyabami et al. (2024) using Pearson correlation analysis between sampling time and soil properties. We repeated this analysis for SOC and WEOC in archived soils from the four long-term agricultural studies" (lines 116-119).

Line 116: archived sample sets would better explain the setup of the experiment

In my opinion, section 3.2 and 2.2 should clearly state when archived samples where reanalysed, mentioning the storage time elapsed after sampling, in a scheme. I found that the experimental design is not fully explaining if the samples from long term experiments were collected in year X and reanalysed for this study. In this case, can the authors compare the WEOC data from the first time they were analysed and the reanalysis here? The increasing or decreasing trends can be driven by the land use over years, not only storage.

We appreciate the reviewer's feedback regarding the lack of clarity in the sampling design. As discussed in our previous responses, we systematically revised the Materials and Methods section to clearly indicate when the archived samples were reanalyzed and to explain how the sampling design differed between the Florida study and the other long-term experiments. The Florida cover crop experiment was uniquely designed to isolate storage effects by comparing original and reanalyzed measurements from the same archived samples, thereby isolating storage effects from field-driven temporal variability. In contrast, our intention in including the four long-term experiments was not

to isolate storage effects independently from management history, but rather to evaluate the practical implications of using archived air-dried soils in long-term soil health assessments.

line 139: why post-rewetting studies are cited to justify WEN decrease during storage?

We clarified in the revision that measurements of WEN involve rewetting archived soils with water, which is why we pointed out its dissimilarities with the post-wetting studies (lines 152-154).

Lines 147-150: as I said above, a quick mention to doubts about POX-C methods and "lability" should be mentioned here.

Besides our Introduction, we highlighted the methodological limitations and interpretational uncertainties associated with POX-C in our revision as follows: "Recent studies have questioned its interpretation as an indicator of biologically active carbon (Christy et al., 2023; Vellenga et al., 2025; Woodings and Margenot, 2023)." (lines 172-174). "POX-C pool remained stable during multi-year storage, underscoring its potential for assessing long-term management impacts in archived soils. This offers new insights into the ongoing debate on the interpretability and practicality of POX-C as a soil health indicator (Gasch et al., 2020; Margenot et al., 2024; Woodings and Margenot, 2023)" (lines 208-211).

Line 163: Again, the use of "temporal trends" could refer to trends driven by the agronomic practices inducing shifts in soil indicators or by the storage time at room temperature that is causing re-arrangements of C pools. I suggest to use different wording to avoid this confusion.

We fully understand the reviewer's suggestion to clarify the drivers of these trends. Although we found it challenging to identify terminology that clearly distinguished the different effects, we revised the Results and Discussion sections to explicitly identify the drivers of each trend and avoid potential confusion. These revisions also complement our changes to the Methods section by providing a clearer explanation of the sampling design and analytical approach.

Line 181: "controlled conditions" does not equal to optimized condition to minimize soil microbial activity. WEOC can be analysed in stored samples if the moisture is kept very low and the storage temperature is low (but not on frozen samples). But I agree that measuring WEOC as soon as possible is the best option.

Thank you for your positive feedback.

Line 190: see my comment about POX-C. I think it has potential to became a soil health indicator, but since the last 15 years the method has marginally improved and actually is a worst indicator than SOC, for example.

We highlighted the caution against the use of POX-C for its methodological limitations and interpretational uncertainties in the revised Discussion, as described earlier.

COMMENTS FROM REVIEWER 2

The authors investigated the effect of long-term storage on various soil-health-related indicators of air-dried soil samples. The topic should be interesting and will attract the readers. However, the quality of their manuscript needs substantial improvement for further processing. The worst point to revise is the fundamental lack of information regarding preparation in their air-dried samples. There is no information about the environmental conditions during air drying and the final soil moisture after the complete air-drying, which should substantially affect the characteristics of air-dried soils.

Thank you for highlighting those issues.

In our revised Materials and Methods section, we clarified that the soils from the cover crop experiment "were air-dried in a climate-controlled laboratory on open trays until a constant mass was achieved before sieving and archiving. Samples were then stored in airtight plastic bags at controlled room temperature (i.e., 24°C) in the lab." (lines 75-76). For the LTAR/LTER soils, we included in our revision that "after sampling, these soils were air-dried under laboratory room-temperature conditions to constant mass prior to archival storage" (lines 91-92). We reported the residual moisture content measured in a subset of air-dried samples from our cover crop study, which "averaged 0.26% (n = 9; range: 0.20-0.32%; CV = 16%). This value is consistent with expectations for coarse-textured sandy soils, which retain substantially less residual water than fine-textured soils (Poeplau et al., 2015; Wäldchen et al., 2012)." (lines 134-136). Previous studies reported residual water contents ranging from approximately 0.3–6% in air-dried mineral soils, with the lowest values generally observed in sandy soils and the highest values in clay-rich soils (Wäldchen et al., 2012; Poeplau et al., 2015). For example, Wäldchen et al. (2012) observed water contents as low as 2.8 g kg^{-1} (0.28%) in air-dried sandy soils and demonstrated that residual water is primarily controlled by clay content and adsorption within fine pores. We also discussed the potential contribution of residual moisture to variation among archived samples. "Microbial activity may persist even at the low water contents present in air-dried soils, potentially resulting in continued transformation of labile carbon and nitrogen pools during storage." (lines 44-45).

Furthermore, the scientific strategy and meaning of the additional soil analyses from four long-term agricultural experiments are very unclear, even after I have access to their data sets available for review. Fig 2 shows the results for those additional analyses. The data, however, likely reflect the long-term trends in soil characteristics during the agricultural experiments, rather than the effect of long-term storage on air-dried soils. If the authors still believe the analysis should work for evaluating the effect of long-term storage on air-dried soils, more detailed and dedicated descriptions are essential in the manuscript. Without fully resolving those issues, it is impossible to evaluate the relevance of the discussion and its conclusions properly. Thus, there is no meaning in achieving publication.

We appreciate the reviewer's thoughtful comments and fully understand the concern regarding the interpretation of the four long-term agricultural experiments.

In the revised manuscript, we systematically clarified the study design and the distinct objectives of the Florida experiment and the four long-term agricultural experiments. Specifically, we emphasized that the Florida experiment was uniquely designed to isolate storage effects by comparing original and reanalyzed measurements from the same archived samples, thereby isolating storage effects from field-driven temporal variability. In contrast, the four long-term experiments were not intended to isolate storage effects independently from management history. Rather, they were included to evaluate the practical implications of using archived air-dried soils in long-term soil health assessments, where storage effects are inevitably intertwined with management-induced changes. Our intention in including the four long-term experiments was not to isolate storage effects

independently from management history, but rather to evaluate the practical implications of using archived air-dried soils in long-term soil health assessments.

We revised the Results and Discussion sections to clarify the interpretation of the four long-term agricultural experiments. Specifically, we emphasized that, although long-term management practices influence soil properties, a common pattern emerged across all four archived datasets: WEOC declined with increasing storage duration, whereas bulk SOC and SOM did not. "Our finding indicates that labile carbon indicators may be more susceptible to storage-induced changes than bulk carbon measurements and underscores the importance of accounting for sample storage history when interpreting long-term soil health trends from archived soils" (lines 203-205).