Reviewer 2

Summary:

This manuscript leverages data from a long-term agricultural experiment at the Russell Ranch in California and a year of more detailed measurements to explore interacting cover crop and compost effects on subsurface soil carbon dynamics. Authors blend historical measurement of carbon stocks with present day analyses of carbon (bulk C, FTIR), nutrients (Mehlich-III), soil physical properties (aggregation, moisture content, and hydraulic conductivity), and microbial biomarkers (PLFA) at four sampling dates. An ANOVA was used to assess the effect of time, depth, and management, with subsequent separate analysis of differences between management treatments at each of three depths.

Although the experimental design and methods are sound, there is a disconnect between the objective to assess interaction of cover crops and compost and the data analysis. The discussion ties in interesting concepts such as the 'cascade theory' and microbial stress indicators that must be brought up further into the introduction to create a threat throughout the paper.

RESPONSE: We greatly appreciate the amount of time and effort put into this review as evidenced by the extremely constructive comments provided! We will address the reviewer's concerns by reorganizing the manuscript, highlighting the limitations of our data, and drawing on previous research at Russell Ranch as described below..

Below please find my recommendations to reframe the paper and utilize historic data, specific questions and a few line edits for authors. Most edits occur in the first half of the paper, which may help to connect the methods and results into the compelling discussion.

Title: I recommend making this more specific. The final sentence of the discussion states that "care should be taken when applying these results to different soil types and climates"; therefore, adding the soil type or climate (or both) into the title seems prudent.

RESPONSE: Thanks for this helpful comment. We agree and will change the title to "Synergy between compost and cover crops in a Mediterannean row crop system leads to increased subsoil carbon storage"

Abstract:

Throughout the paper, can authors use the treatment names as in the original experimental dataset (Wolf, 2018 page 6): CONV = CMT conventional maize-tomato, ORG=OMT organic maize-tomato, and (page 5) WCC – winter cover crop? I understand that Tautges and Chiartas 2019 used the CONV, ORG notation, but a brief explanation would be helpful.

RESPONSE: As the 2018 Wolf paper was primarily a data paper, and the 2019 Tautges and Chiartas paper is more directly comparable to this manuscript, we chose to use the 2019 abbreviations. We will include a short explanation in the methods to make the difference between abbreviations used in the 2019 Tautges and Chiartas paper and 2018 Wolf paper clearer.

The theory of cover crops providing a macropore system for transport of DOC is interesting, but the data do not support this theory (no measurement of porosity, change in bulk density, or changes in soil hydraulic properties). It is appropriate for a discussion, but I might exclude this as a main finding from the abstract.

RESPONSE: We will adjust the abstract to remove the last two sentences, and replace them with a less explicit statement: "Our results show the potential for increased subsoil carbon storage under compost + cover crops in tilled agricultural systems, and identify potential pathways for increasing carbon transport and storage in subsoil layers."

<u>Introduction:</u>

I appreciate that the abstract and introduction mention soil health, but there is no clear definition or explanation of its importance to the paper. Either simply remove this term and focus solely on soil carbon and microbial processes, or please directly connect soil health and often associated shallow sampling regimes to this "outsized perceived role in ecosystem services".

Response: We agree that inclusion of reference to soil health is not consistent with the topic of this manuscript. We will remove the mention of soil health from both the abstract and introduction. We will also clarify the sentence "outsized perceived role in ecosystem services" to indicate that increased sampling of the top 15 cm of soil marks a growing reliance on the soil surface to answer questions about processes in the entire soil profile, and runs the risk of subsoils being treated merely as "more dilute topsoils" (Salomé et. al 2010).

1. Salomé, C., Nunan, N., Pouteau, V., Lerch, T. Z. and Chenu, C.: Carbon dynamics in topsoil and in subsoil may be controlled by different regulatory mechanisms, Glob. Chang. Biol., 16(1), 416–426, https://doi.org/10.1111/j.1365-2486.2009.01884.x, 2010.

This is a good argument and dataset to support deeper sampling. Authors may also include references summarized by Mobley et al 2015 in their article "Surficial gains and subsoil losses of soil carbon and nitrogen during secondary forest development": Post & Kwon, 2000; West & Post, 2002 review 360 articles on land use change, with only 10% sampling below 30cm.

Response: We will include the suggested references in the introduction.

In this paragraph, please clarify, at what depth are the authors designating topsoil v subsoil for this study?

Response: We will insert a statement that we are designating 0-15cm as topsoil and 15-100 cm as subsoil.

This first paragraph of the introduction discusses "longer C residence times" of deep soil C, which requires further explanation.

Response: We will insert a sentence into the "C chemistry" section in the reorganized introduction explaining the trend of greater C_{14} ages in subsoil layers. The age of carbon below 30 cm generally increases with depth, with estimates of carbon as old as 10^3-10^4 years in deeper regions of the subsoil (Rumpel et. al 2012). This is in contrast to the younger C_{14} ages of 10^2-10^3 years in the top 30 cm (Paul 2001).

1. Rumpel, C., Chabbi, A. and Marschner, B.: Carbon storage and sequestration in subsoil horizons: Knowledge, gaps and potentials, in Recarbonization of the Biosphere: Ecosystems and the Global Carbon Cycle, pp. 445–464, Springer Netherlands, https://doi.org/10.1007/978-94-007-4159-1_20, , 2012.
2. Paul, E. A., Collins, H. P. and Leavitt, S. W.: Dynamics of resistant soil carbon of midwestern agricultural soils measured by naturally occurring 14C abundance, Geoderma, 104(3–4), 239–256, https://doi.org/10.1016/S0016-7061(01)00083-0, 2001.

Overall, the introduction structure can be strengthened by clarifying topic sentences (e.g., specify cover crops L51) and adding updated references. Can you support the Jenny citation with more modern references, even Brady and Weill Nature and Properties of Soils, or USDA technical information "Designations for Horizons and Layers" in Soil Survey Manual – Ch 3 (https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrcs142p2_054253#designations).

Response: We will update the references to include more recent citations related to soil formation.

The introduction structure may flow better using paragraphs separated into chemical, physical and biological controls or layered as (1) depth; (2) chemistry of C inputs and stabilization at depth; (3) management impacts at depth – specifically cover crops; and (4) management interaction with other factors (microbial).

Response: We agree and thank you for this suggestion. We will restructure the introduction to provide separate paragraphs for the impact of depth, C chemistry (introducing FTIR and the cascade theory), management impacts (tillage, cover crops, compost application) and microbes (PLFA stress biomarkers, nutrient stoichiometry) on C storage.

The introduction touches upon stoichiometry, a critical highly manipulated factor in managed conventional systems that effects soil C storage. To go further in depth on soil chemistry (e.g., at L40), authors can address changes over time in stoichiometric constraints on decomposition (e.g., see Soong et al 2019 "Microbial carbon limitation: The need for integrating microorganisms into our understanding of ecosystem carbon cycling").

Response: We will include the suggested reference, and insert a short discussion of the importance of stoichiometry in the microbes and C storage introduction section.

Also, authors can mention higher physical disturbances in surface soils (L55), and the types of management associated with cover crops, such as crimping/rolling.

Response: We will include a short discussion of tillage effects in the "management" section of the introduction.

Please also include specific soil type, climate and cropping system when comparing to other studies, otherwise direct comparisons are not particularly informative.

Response: We will add detail about soil type, climate and cropping system whenever we make comparisons to other studies.

Can also cite McClelland et al 2020 "Management of cover crops in temperate climates influences soil organic carbon stocks: a metaanalysis" that analyzed soils only down to 30cm.

Response: We will include this citation in the manuscript.

As for the sampling strategy by depth, can the authors please describe why they separated out into these depths 0-15, "intervening", and the subsurface as 60-100cm? How do these depths compare to the horizons in these two soils? (Looking up the series descriptions Yolo has A horizons down to 66cm and then C horizons, and Rincon has A down to 20, B 20-100cm. Should the analysis be completed on A and B horizons rather than depth profiles?) How do these depths relate to roots of corn (100cm+), tomato (60cm+) and cover crops (variable)?

Can authors please justify why 15-60cm is combined into a single sample in 2018, when historical data had an additional delineation? (Is it simply limited time/costs or another reason?)

Response: The depths used in this experiment were selected for the ability to compare directly to the results of the Tautges and Chiartas (2019) paper. Combining the 15-30 and 30-60 cm depth intervals that were used in Tautges & Chiartas et. al 2019 was originally done due to limitations of time and resources in processing samples. Authors endeavoured to have all time sensitive analyses done within 1 week of sampling, and were forced to limit the total number of samples to make that happen.

The decision to keep 0-15 cm separate instead of combining the 0-15 and 15-30 cm depths was based on preliminary data showing significant differences among treatments in genes coding for carbon degrading enzymes at 0-15 cm. These differences were not present at deeper depths (15-30, 30-60 or 60-100cm) (D. Rath, unpublished data).

The Yolo soil series is a young soil (Entisol) without strong horizonation or changes in texture with depth, which we used to justify using depth-based increments instead of horizon-based increments. We did not detect a significant effect of block (correlated with soil type) on any of the variables measured in the experiment, and so applied the same depth increments to the plots in the Rincon soil series, despite the fact that it has slightly more evidence of clay accumulation (Alfisol).

Please stay consistent with the terms “subsoil” versus “subsurface soil”, as depth is a major component of this study.

Response: We originally used “subsurface soil” to refer to soils between 15 and 100 cm and “subsoil” to refer to soils deeper than 100 cm based on Rumpel et. al (2012). In order to reduce confusion we will use the term “subsoil” to refer to soils at all depths below 15 cm.

The overarching question and hypothesis require further editing to clearly lead into the results and discussion. There seems to be a disconnect between the main question and the methods of this paper. The main question includes “carbon formation” (does that mean microbially processed C? or stabilizedC formation?) and “storage processes” (that obviously includes aggregation, but the carbon content of these size classes was not measured). Also, what is meant by the term “SOC-related indicators”, does that mean SOC stability or reactivity-related indicators? As written the hypotheses are just predictions, there is no description as to the mechanisms behind the described expected results.

Response: We agree that we need to strengthen our question and hypotheses and thank you for your suggestions. Accordingly , we will edit our manuscript to clarify the overarching question and hypotheses and clarify “carbon formation” to mean stabilized carbon formation.

We will also remove the term “storage processes” from the overarching question - we did not do any measurements of actual processes, but our results taken together suggested a mechanism by which carbon stocks were increased in ORG systems. The term SOC-related indicators refers to the connections between C and nutrient input, microbial biomass and SOC formation; we will remove this term and replace it with a more specific statement.

We will modify the hypothesis section to highlight that the main hypothesis of the paper is that “*high concentrations of mobile C and essential nutrients for microbial activity provided by the compost, combined with the easier movement of water downward associated with a history of cover-cropping, helped transport the material needed to build C in the subsoil*” as suggested in a later comment.

An interesting hypothesis arises in the discussion around cascade theory, can authors pull that into the introduction? This can provide a way to integrate the study of carbon chemistry (FTIR) and microbial biomarkers that otherwise are not included in the hypotheses.

Response: We will include a short discussion of the cascade theory in the introduction, under the reorganized “C chemistry” section.

Finally, I agree with the previous reviewer comment, that the treatments CONV (fertilizer), CONV+WCC (fertilizer + cover crops), and ORG (compost + cover crops) do not disentangle the effect of compost. I don't think there is there a treatment in the Century Experiment that was maize-tomato plus compost only or fertilizer + compost, but this should be mentioned as a limitation in the study, particularly in the subtraction of FTIR spectra.

Response: We will include a clear statement in the methods and introduction that there is no compost-only treatment, and make it clear that we cannot compare the effects of compost alone to the effects of cover crops alone due to the experimental design. We will also include this in our discussion of the limitations of the data.

This manuscript covers many aspects of deep soil C and management, no need to emphasize the complicated factors of global change (L87) at the end of the introduction, unless those are also analyzed over time.

Response: We agree and will remove this final sentence from the introduction.

Materials and Methods:

Thank you for a concise description of the site and experiment. I recommend authors also add basic climate data such as climate type, mean annual mix and max temperature, mean annual precipitation, and also specific 2018-19 climate data for comparison.

Authors write that the 'horizon information' is available from Wolf et al, but I only can find soil chemistry by depth, not the soil description in that dataset (horizon delineations are online). Can authors add in the horizon depth into the methods for both the Yolo and Rincon soils, and key chemistry such as pH and texture? A table in the materials and methods section could organize all of this soil and climate information for quick reference.

This could also include the other key management notes that will impact DOC transport, such as the conversion from furrow to drip in 2014, as well as information from the 2018-2019 season such as crop planting/harvest dates, total irrigation amount, and the anomalous compost application in September 2019. These details can then be incorporated smoothly into the discussion.

Response: Thank you for the suggestion to include additional data on climate, horizon and management. We agree that the Wolf et al. manuscript does not include soil description (horizons); this was an oversight on our part. We will include tables in the Methods/Appendix section describing:

1) Climate variables (average, 1993, 2018) including MAT, MAP, Min Temp, Max Temp
2) Soil horizon variables including texture, layer depths, pH, and %C estimates from SSURGO data

3) Management practices and dates

The differentiation between the sampling and analysis of the older data and 2018-2019 methods is now clearer. Thank you for the new methods section. However, without hypotheses asking seasonal questions over time – why sample at four time points in a single year? Particularly as the authors state that a single year of data is not sufficient to look at differences at depth (L81-82) to justify use of the historical data. Perhaps authors can create one or two hypotheses for the 2018-19 season, and other for the long-term effects and historical data.

Response: We sampled at 4 time points throughout the year based on the hypothesis that EOC and available nutrients in ORG systems would be higher only at 1 or 2 time points (June and Aug), and to get an idea of how much variation there was during the year. We did not include this discussion of seasonal variation during the year for brevity's sake, choosing instead to focus on what the overall trends indicated about C and nutrient transport. We will modify the manuscript to include a short discussion of seasonal variation in our measurements, and the potential implications for stoichiometry , microbial biomass and C formation.

The use of PLFA and FTIR is not justified from the hypotheses or introduction. The use of these techniques, particularly stress ratios for PLFA needs to be explained within a wider context in the introduction.

Response: We will include short descriptions and justifications for use of these methods in the introduction.

2.7 Please clarify the statement that 9 out of 18 plots were sampled for hydraulic conductivity (those under tomato). Were half of the plots under corn and the other under tomato during this 2018 sampling? That needs to be included in the methods section. Or are you referencing the full 18 plots of all the Century experimental treatments? Finally, why are 8 dates included for soil moisture content, when soils are sampled only 4 times?

Response: To clarify, there were 18 plots total included in our 2018 sampling - 6 ORG, 6 CONV+WCC, and 6 CONV. For each of these treatments, 3 plots were planted with tomato, and 3 plots were planted with corn. This gives a total of 9 plots under corn, and 9 plots under tomato.

For the hydraulic conductivity sampling in August, we only sampled the 9 plots under tomato as we were not able to access the corn plots due to harvesting activities. We will modify the methods section to make this clearer by including the number of replicates in a table.

In regards to the dates for soil moisture content, this was an error on our part: moisture content was measured 8 times, not 4 times. We will correct our methods section accordingly.

2.8 I have some concern over the use of averaging and subtraction of the spectra. What was the variance between the historic soils of 15-30 and 30-60 cm?

Response: The variation between 15-30 and 30-60 cm soils in 1993 was negligible for all three systems. We have included a sample graph of the residuals in 1993 ORG soils to illustrate this point.

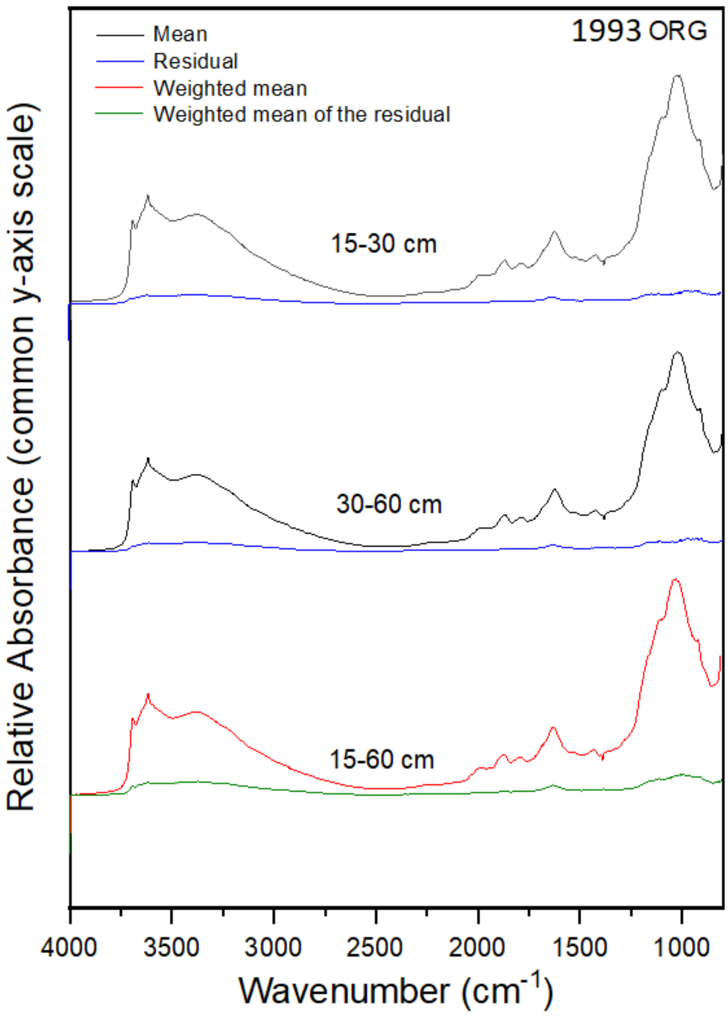
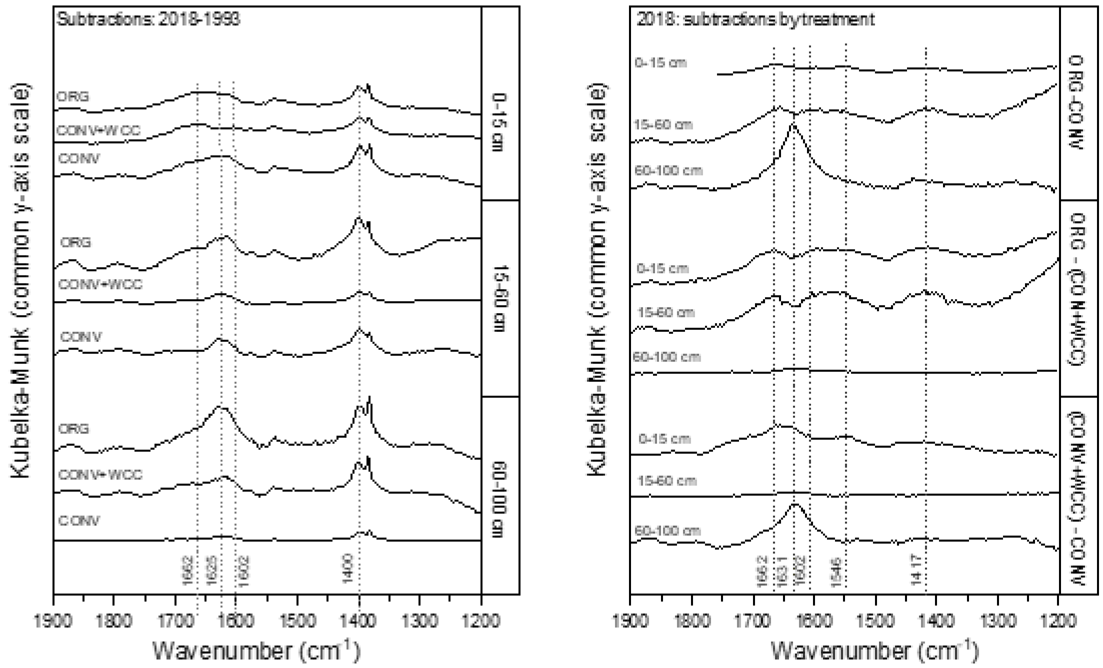


Figure A5. Averaged spectra and residuals for 15-30 and 30-60 cm, and weighted average and weighted mean of the residual for 15-60 cm in 1993 ORG samples.

What information is provided via subtraction of the conventional plus cover crop from the organic spectra?

Response: This subtraction allows us to compare how the C signature in soils under cover crops with mineral fertilizer compare to cover crops with compost. As noted in the right-hand figure below, there is not much difference between the ORG and CONV+WCC spectra at 60-100 cm. This suggests that the difference in the aromatic or carboxylate C content of these two treatments at 60-100 cm is tiny in relation to the ORG-CONV or (CONV+WCC)-CONV comparison.



I am unfamiliar with this subtraction analysis, so I am curious, what information is revealed from subtraction as the reflectance intensity does not represent quantity, but rather soil chemical signature?

Response: While FTIR is not a strictly quantitative tool for identifying specific compounds in mixed samples, it can be used pseudo-quantitatively due to the fact that the absorption of IR light by a specific molecular bond at a specific electromagnetic frequency follows the Beer-Lambert Law (Beer's Law) (e.g., Margenot et al 2016, Smith 2001). Therefore, the height and area of a spectral peak are proportional to the abundance of molecules in a sample (linear relationship), and comparing the presence and absence of peaks and the relative differences in spectral contributions from each peak in a subtraction can suggest differences in C chemistry.

Previous studies with DRIFTS in both the near-infrared (Dalal and Henry, 1986) and mid-infrared regions (Demyan et al., 2012; Margenot et al., 2015; West et al., 2020; Deiss et al., 2021) have shown direct associations between soil organic carbon concentration and absorbance at specific frequencies (depicted as peak height or area of single peaks or peak ratios). The non-linearity of concentration and absorbance arising from DRIFT can be partially corrected using the Kubelka-Munk (KM) function, which we will include alongside the uncorrected spectra in the manuscript, since a KM correction does not fully correct spectra (e.g. Clark and Roush, 1984) .

Assuming that band assignments are correct (they are consistent with literature), the presence of a peak in the subtraction at 1666 cm-1 suggests a difference in aromatic C in the surface horizons, especially for the ORG treatment. We will also update the manuscript to provide peak intensity ratios of aromatic to carboxyl moieties [v(C=C):vas(COO-) (1662 cm^{-1}:1631cm^{-1})] as an additional way of showing the relative difference within each subtraction.

1. Clark, R.N., and Roush, T.L., 1984. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications, *J. Geophys. Res.*, 89(B7), 6329– 6340, doi:10.1029/JB089iB07p06329.
2. Dalal, R.C. and Henry, R.J., 1986. Simultaneous Determination of Moisture, Organic Carbon, and Total Nitrogen by Near Infrared Reflectance Spectrophotometry. Soil Science Society of America Journal, 50: 120-123. https://doi.org/10.2136/sssaj1986.03615995005000010023x
3. Deiss, L., Sall, A., Demyan, M.S., Culman, S.W., 2021. Does crop rotation affect soil organic matter stratification in tillage systems. Soil & Tillage Research 209, 104932.
4. Demyan, M.S., Rasche, F., Schulz, E., Breulmann, M., Müller, T., Cadisch, G., 2012. Use of specific peaks obtained by diffuse reflectance Fourier transform mid-infrared spectroscopy to study the composition of organic matter in a Haplic Chernozem. Eur. J. Soil Sci. 63, 189–199. doi:10.1111/j.1365-2389.2011.01420.x
5. Margenot, A.J., Calderón, F.J., Bowles, T.M., Parikh, S.J., Jackson, L.E.. 2015. Soil organic matter functional group composition in relation to organic carbon, nitrogen, and phosphorus fractions in organically managed tomato fields *Soil Science Society of America Journal* 79:772-782.
6. Margenot, A.J., F.J. Calderon, and S.J. Parikh. 2016. Limitations and potential of spectral subtractions in Fourier-transform infrared spectroscopy of soil samples. Soil Sci. Soc. Am. J. 80:10-26.
7. Smith, B.C., 2011. Fundamentals of Fourier Transform Infrared Spectroscopy. Second Edition. CRC Press, Taylor and Francis Group.
8. West, J.R., Cates, A.M., Ruark, M.D., Deiss, L., Whitman, T., Rui, Y., 2020. Winter rye does not increase microbial necromass contributions to soil organic carbon in continuous corn silage in North Central US. Soil Biol. Biochem. 107899. doi:10.1016/j.soilbio.2020.107899

2.9 Can the authors please describe the details of the ANOVA. Was this a mixed effect model accounting for the block design? Was there an effect of block? (That difference would be interesting to see due to the two soil types). It would be helpful if the authors state that they checked normality of the data prior to ANOVA.

Response: We will include this missing detail in the methods section. We first used a mixed effect model with block as a random effect. Since the block was not significant for any of the variables measured, we removed it from the model. We also checked normality and assumptions of the linear model prior to ANOVA. Due to the limited number of samples we included all measurements in our statistical tests, including potential outliers.

If variability was high for certain metrics (hydraulic conductivity), it seems there may be some outliers, how were those assessed?

In the case of the hydraulic conductivity measurements, these outliers potentially represent the contributions of macropores. It is well recognized that hydraulic conductivity measurements can vary widely across fields and landscapes (Rahmati et al. 2018) and often do not account for the presence of macropores (Brooks et al. 2004). Ksat values can span 3 orders of magnitude within a short distance (Øygarden et al. 1997). Thus, although our measurements do not reveal statistically significant differences between the treatments, the scattered high-permeability zones in the cover-crop treatments represented by the outliers are likely to contribute to rapid moisture redistribution and may play an important role in explaining the elevated deep moisture.

1. Brooks, E. S., Boll, J. and McDaniel, P. A.: A hillslope-scale experiment to measure lateral saturated hydraulic conductivity, Water Resour. Res., 40(4), 4208, https://doi.org/10.1029/2003WR002858, 2004.
2. Øygarden, L., Kværner, J. and Jenssen, P. D.: Soil erosion via preferential flow to drainage systems in clay soils, Geoderma, 76(1–2), 65–86, https://doi.org/10.1016/S0016-7061(96)00099-7, 1997.

3. Rahmati, M., Weihermüller, L., Vanderborght, J., Pachepsky, Y. A., Mao, L., Sadeghi, S. H., Moosavi, N., Kheirfam, H., Montzka, C., Van Looy, K., Toth, B., Hazbavi, Z., Al Yamani, W., Albalasmeh, A. A., Alghzawi, M. Z., Angulo-Jaramillo, R., Antonino, A. C. D., Arampatzis, G., Armindo, R. A., Asadi, H., Bamutaze, Y., Batlle-Aguilar, J., Béchet, B., Becker, F., Blöschl, G., Bohne, K., Braud, I., Castellano, C., Cerdà, A., Chalhoub, M., Cichota, R., Císlerová, M., Clothier, B., Coquet, Y., Cornelis, W., Corradini, C., Coutinho, A. P., De Oliveira, M. B., De Macedo, J. R., Durães, M. F., Emami, H., Eskandari, I., Farajnia, A., Flammini, A., Fodor, N., Gharaibeh, M., Ghavimipanah, M. H., Ghezzehei, T. A., Giertz, S., Hatzigiannakis, E. G., Horn, R., Jiménez, J. J., Jacques, D., Keesstra, S. D., Kelishadi, H., Kiani-Harchegani, M., Kouselou, M., Jha, M. K., Lassabatere, L., Li, X., Liebig, M. A., Lichner, L., López, M. V., Machiwal, D., Mallants, D., Mallmann, M. S., De Oliveira Marques, J. D., Marshall, M. R., Mertens, J., Meunier, F., Mohammadi, M. H., Mohanty, B. P., Pulido-Moncada, M., Montenegro, S., Morbidelli, R., Moret-Fernández, D., Moosavi, A. A., Mosaddeghi, M. R., Mousavi, S. B., Mozaffari, H., Nabiollahi, K., Neyshabouri, M. R., Ottoni, M. V., Ottoni Filho, T. B., Pahlavan-Rad, M. R., Panagopoulos, A., Peth, S., Peyneau, P. E., Picciafuoco, T., Poesen, J., Pulido, M., Reinert, D. J., Reinsch, S., Rezaei, M., Roberts, F. P., Robinson, D., Rodrigo-Comino, J., Rotunno Filho, O. C., Saito, T., et al.: Development and analysis of the Soil Water Infiltration Global database, Earth Syst. Sci. Data, 10(3), 1237–1263, https://doi.org/10.5194/essd-10-1237-2018, 2018.

The lack of differences in the field may simply be due to low power with only three field replicates. Rather than splitting the data by depth to do comparisons between treatments, can the authors run an analysis that accounts for autocorrelation over depth? On that same note, do authors need to account for repeated measures across sampling dates in 2018-2019 and within the historical data?

Response: We will attempt to do the suggested autocorrelation analyses. One note is that for comparisons of the historical data, we only used the harvest time point in 2018 (TP3) and the harvest time point in 1993 to avoid potential issues with seasonal variation and repeated measures.

I appreciate access to the data and code used for this analysis. Thank you for supporting transparency in data analysis.

Response: Thank you!

3.1 The cumulative inputs over 25 years are useful, but would be more comparable to other studies if averaged per year. This data also may be well suited for a table including all C inputs and nutrient inputs over the 25 year period (transform Fig 1 to Table 1 using Mg/ha/yr). Perhaps with the level of detail from the Century Experiment on all organic inputs, the statistical analysis could incorporate the treatments as continuous variables (amount of mineral/organic N input) rather than categorical variables?

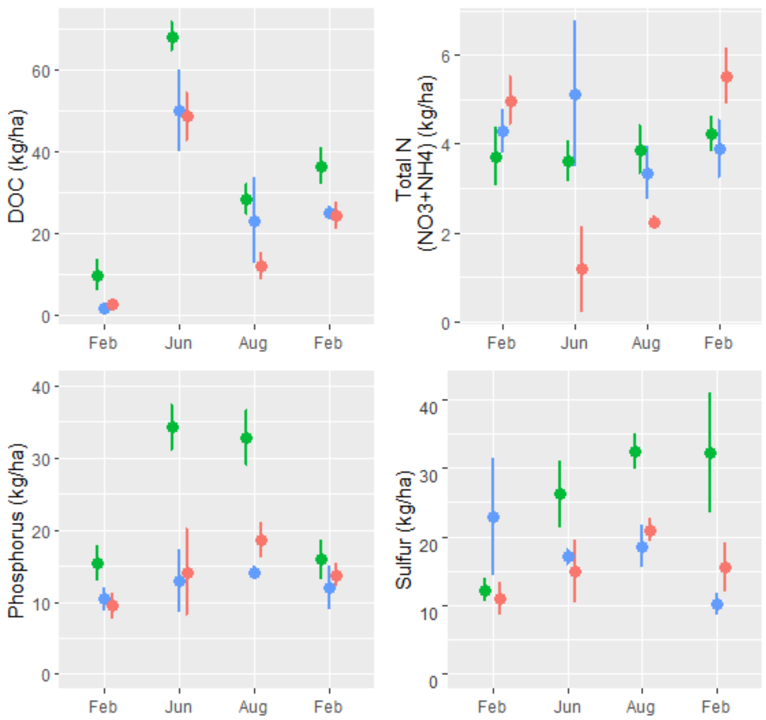
Response: Thanks for this idea. We will convert this graph to a table, and re-run the analyses using the amount of C and N added as a continuous variable.

L227 If a result is non-significant, than I would remove any interpretation of 'increase'.

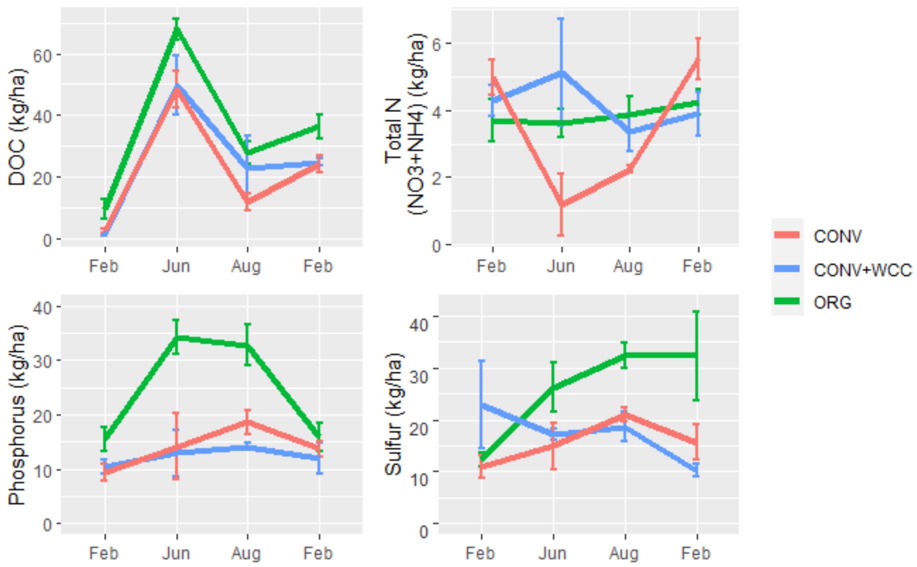
Response: We will remove any interpretation of "increase", but will still include the graphs in our results and discussion sections, as we believe they are indicative of an overall trend in the data.

Fig 2. Extremely clear pattern here. Can the significant differences be noted in some way on the figure? I would remove the lines between the points, as there are no actual measurements there, and the trends are obvious.

Response:We appreciate this comment but believe that having lines linking data points helps visualize the trends and makes interpreting the graph more intuitive than just having the points. We have included sample plots for nutrient stocks at 60-100 cm to illustrate this point: We will add indications of significance at $p<0.05$ to the figure in the updated manuscript.



Sample Plot 1. Nutrient stocks from 60-100 cm displayed using only means and standard error bars



Sample Plot 2. Nutrient stocks from 60-100 cm displayed using means and error bars, with lines connecting points.

Fig. 4 I would change the layout of this figure. You can zoom in on the y-axis and add precipitation and irrigation events. Otherwise, a simple average across the time and bar graph or box plot would tell the story more clearly, since the statistical analysis was not over time.

Response: We will make the suggested changes to the y-axis.

L270 Why do authors state "largest seasonal variation" in nutrient data was in June, when only mineral N and DOC were highest in June? S and P were higher in August.

Response: The statement referenced above is

"Nutrient values showed large seasonal variation, with the highest levels of carbon, nitrogen and sulfur observed during the June timepoint. DOC, mineral nitrogen, and sulfur values were lowest during the winter (Nov - Feb), which coincided with the period of highest rainfall. Phosphorus levels increased slightly throughout the 2018-2019 year."

We will correct this statement to state that DOC and mineral N were highest in June, whereas S was slightly higher in August. P was highest in Feb 2019.

3.7 Authors must introduce microbial stress indicators earlier in the introduction and hypothesis. How does this relate to stoichiometry and soil C stability?

Response: We will include a description/importance of stress indicators in the introduction and hypotheses. PLFA stress ratios are indicators of limited carbon and water availability to microbes. They represent an overall shift away from the thinner, more permeable cell membranes associated with Gram - bacteria and monounsaturated fatty acids towards the more tightly packed, less permeable cell membranes associated with Gram+ bacteria and saturated fatty acids (Silhavy et. al 2010).

These stress ratios are also connected to the availability of soil C. An increase in the Gram+:Gram- ratio has been associated with a decrease in easily available water and carbon (Fanin et. al 2019, Fierer et. al 2003, Bossio et. al 1998), while an increase in the saturated: unsaturated ratio and cy17:pre ratios are associated with dehydrated conditions (Moore-Kucera et. al 2007).

We have cited these stress indicators as supporting the idea of increased soluble carbon and water content in ORG systems, but have avoided making predictions about what they mean for carbon stability or how they respond to changes in nutrient stoichiometry due to inconsistencies in the literature. Making functional predictions about microbial carbon usage using PLFA are difficult, as many PLFAs are not specific to distinct species and instead distributed across taxa (Ruess et. al 2010). Additionally, the relationship between nutrient stoichiometry and PLFA biomass and ratios is inconsistent, and changes over time (Huajun et. al 2016, Ng et. al 2014).

1. Bossio, D. A. and Scow, K. M.: Impacts of carbon and flooding on soil microbial communities: Phospholipid fatty acid profiles and substrate utilization patterns, Microb. Ecol., 35(3), 265–278, https://doi.org/10.1007/s002489900082, 1998.
2. Fanin, N., Kardol, P., Farrell, M., Nilsson, M. C., Gundale, M. J. and Wardle, D. A.: The ratio of Gram-positive to Gram-negative bacterial PLFA markers as an indicator of carbon availability in organic soils, Soil Biol. Biochem., 128, 111–114, https://doi.org/10.1016/j.soilbio.2018.10.010, 2019.
3. Fierer, N., Schimel, J. P. and Holden, P. A.: Variations in microbial community composition through two soil depth profiles, Soil Biol. Biochem., 35(1), 167–176, https://doi.org/10.1016/S0038-0717(02)00251-1, 2003.
4. Moore-Kucera, J. and Dick, R. P.: PLFA profiling of microbial community structure and seasonal shifts in soils of a Douglas-fir chronosequence, Microb. Ecol., 55(3), 500–511, https://doi.org/10.1007/s00248-007-9295-1, 2008.
5. Ng, E. L., Patti, A. F., Rose, M. T., Schefe, C. R., Wilkinson, K. and Cavagnaro, T. R.: Functional stoichiometry of soil microbial communities after amendment with stabilised organic matter, Soil Biol. Biochem., 76, 170–178, https://doi.org/10.1016/j.soilbio.2014.05.016, 2014.
6. Ruess, L. and Chamberlain, P. M.: The fat that matters: Soil food web analysis using fatty acids and their carbon stable isotope signature, Soil Biol. Biochem., 42(11), 1898–1910, https://doi.org/10.1016/j.soilbio.2010.07.020, 2010.
7. Silhavy, T. J., Kahne, D. and Walker, S.: The bacterial cell envelope., Cold Spring Harb. Perspect. Biol., 2(5), https://doi.org/10.1101/cshperspect.a000414, 2010.
8. Huajun, Y., Phillips, R. P., Liang, R., Xu, Z. and Liu, Q.: Resource stoichiometry mediates soil C loss and nutrient transformations in forest soils, Appl. Soil Ecol., 108, 248–257, https://doi.org/10.1016/j.apsoil.2016.09.001, 2016.

Discussion:

Authors list the key finding of increased SOC and then write what I perceive as the hypothesis of the paper: "that high concentrations of mobile C and essential nutrients for microbial activity provided by the compost, combined with the easier movement of water downward associated with a history of cover-cropping, helped transport the material needed to build C in the subsurface." Having this in the introduction will help to set up the statistical analysis, results, and discussion. However, this hypothesis was not supported by the aggregation data or the hydraulic conductivity data.

Response: We agree with the reviewer that including this hypothesis in the introduction will better set up the rest of the paper, and will do so in the revised manuscript.

We will adjust our discussion to reference previous data from Russell Ranch that supports the idea of increased infiltration under cover crops. Work carried out at Russell Ranch over the last 30+ years has shown that the cover crop mix of hairy vetch, faba beans and oats used in this study increased infiltration and DOC input into the soil profile (Mailapalli et al., 2012). It also increased soil moisture-holding capacity during saturated conditions (Joyce et al. 2002) with no significant differences in bulk density (Colla et al. 2000) and reduced soil surface strength (Folorunso et al. 1992).

Proposed mechanisms behind the increased infiltration and moisture storage include increased aggregation, reduced crusting, and increased macroporosity. It is important to note that these root-induced soil alterations are highly localized and dependent on the root architecture of the cover crops. Specifically, cover crops with prominent tap roots (faba bean) are effective at creating continuous bio-pores, while fibrous roots (oat and hairy vetch) are effective at forming aggregates (Oglive et al., 2021). Therefore, the mixture of cover crops planted at the site likely resulted in widely variable aggregation and pore connectivity effects.

In the context of these previous observations, we attribute the increased variability of our hydraulic conductivity measurements and the increased moisture storage observed in ORG and CONV+WCC plots to the presence of WCCs. Our inference of greater soluble carbon input to subsoil layers under compost application is also supported by the multiple lines of evidence that we present in the manuscript:

1) More soluble C in ORG subsoils
2) Observations of higher water infiltration and storage under cover crops
3) Greater amounts of soluble organic carbon in compost, and
4) Reduced subsoil microbial stress indicators under ORG systems (attributed to more C availability)

1) Colla, G., Mitchell, J. P., Joyce, B. A., Huyck, L. M., Wallender, W. W., Temple, S. R., Hsiao, T. C. and Poudel, D. D.: Soil physical properties and tomato yield and quality in alternative cropping systems, in Agronomy Journal, vol. 92, pp. 924–932, American Society of Agronomy, https://doi.org/10.2134/agronj2000.925924x, , 2000.
2) Folorunso, O., Rolston, D., Prichard, P. and Louie, D.: Cover crops lower soil surface strength, may improve soil permeability, Calif. Agric., 46(6), 26–27 //calag.ucanr.edu/archive/?article=ca.v046n06p26, last access: 6 May 2021, 1992.
3) Joyce, B. A., Wallender, W. W., Mitchell, J. P., Huyck, L. M., Temple, S. R., Brostrom, P. N. and Hsiao, T. C.: INFILTRATION AND SOIL WATER STORAGE UNDER WINTER COVER CROPPING IN CALIFORNIA'S SACRAMENTO VALLEY, Trans. ASAE, 45(2), 315–326, 2002.
4) Mailapalli, D. R., Horwath, W. R., Wallender, W. W. and Burger, M.: Infiltration, Runoff, and Export of Dissolved Organic Carbon from Furrow-Irrigated Forage Fields under Cover Crop and No-Till Management in the Arid Climate of California, J. Irrig. Drain. Eng., 138(1), 35–42, https://doi.org/10.1061/(asce)ir.1943-4774.0000385, 2012.
5) Ogilvie,C.M.;Ashiq,W.; Vasava, H.B.; Biswas, A. Quantifying Root-Soil Interactions in Cover Crop Systems: A Review. Agriculture 2021, 11, 218. https://doi.org/10.3390/ agriculture11030218

Please go into more detail on how no differences in aggregation "rule out" increased pore space as the increase in water content. What is the alternative explanation? Is this just an issue with statistical power?

Response: We agree that "rule out" was not the best choice and will reword that sentence - We believe that the increase in moisture content is not due to an increase in pore space from increased aggregation. Instead, we make the argument that the increase in moisture content was due to an increase in root-related macropores.

L335 This also seems like a great candidate sentence for another hypothesis: "Due to the fact that tillage in all systems would likely eliminate differences among them in the top 30 cm, we would expect any differences in macroporosity and infiltration among treatments to be most affected by those roots that extend below the 30 cm plow layer". This is the first mention of tillage depth. Please specify the depth of disking in the methods, and if this was applied to the conventional fields as well.

Response: Tillage is part of the management of all three farming systems sampled in the manuscript. We have avoided hypotheses related to tillage in the manuscript as the difference in the amount of tillage between our experimental systems is not very large (an additional 1-2 passes/year in CONV+WCC and ORG plots to incorporate cover crops). Additionally, the historical data does not include the amount of tillage per plot. The quoted sentence was intended to highlight that all of the WCC mix used in the RR plots have roots that extend deeper than 30 cm.

We now include disking depth in the tables added to the Methods section.

L340-345 This paragraph on cascade theory describes why FTIR analysis was necessary. This also should be included, or at least alluded to, in the introduction. This is a really interesting discussion (L350-355), and could also be a good place to bring up the variability in the conductivity data.

Response: We will include a short introduction to the cascade theory in the introduction.

L371 Figure referenced should be Fig 9.

Response: We will correct this figure reference.

L375 Support with values from the results. The nutrient values may all be better represented by tables, although the graphs show dynamics across the season, I would argue that depth, not season, is the key factor in this analysis.

Response: We will include rough CNPS ratios in our discussion to support this point. We did not include actual nutrient ratios in our discussion as the reference nutrient stoichiometry values reported in Richardson et.al 2014 were calculated using total nutrient values (digestion) as opposed to the available nutrient values used in this paper. We will include this caveat in our discussion.

We agree that depth is a more important factor than seasonal variation in this manuscript. However, EOC in Mediterranean systems has been shown to vary greatly across the course of a single year (Steenwerth et. al 2008), and showing data across multiple timepoints demonstrates that our observation of increased EOC in ORG systems is not constrained to a single sampling date or depth. This is not true for mineral N and S however, which indicates that future projects involving available N or S in this system should carefully select timepoints for measurement if they mean to compare results to the literature.

We will include a discussion of the importance of seasonal variation in our dataset, and highlight our reasoning behind sampling multiple times throughout the year in the manuscript revision.

1. Richardson, A. E., Kirkby, C. A., Banerjee, S. and Kirkegaard, J. A.: The inorganic nutrient cost of building soil carbon, Carbon Manag., 5(3), 265–268, https://doi.org/10.1080/17583004.2014.923226, 2014.
2. Steenwerth, K. and Belina, K. M.: Cover crops enhance soil organic matter, carbon dynamics and microbiological function in a vineyard agroecosystem, Appl. Soil Ecol., 40(2), 359–369, https://doi.org/10.1016/j.apsoil.2008.06.006, 2008.

L380 Consider rewriting this section title, as there was no direct comparison to a compost treatment alone.

Response: We agree and will reword this section title, and restructure the manuscript by including a clear statement in the methods and introduction that there is no compost-only treatment, and make it clear that we cannot compare the effects of compost alone to the effects of cover crops alone due to the experimental design. We will also include this in our discussion of the limitations of the data.

L382 Is the microbial processing near the surface based on the FTIR data? Please reference.

Response: The conclusion of increased microbial processing near the surface in ORG systems was based on our observations of increased microbial biomass in 2018 as well as increased carboxylate carbon in ORG systems at 0-15 cm from 1993-2018. This increase in carboxylate carbon from 1993-2018 was attributed to microbial oxidation of carbon inputs. We will reference both our biomass and FTIR data in this section (Figure 8a, 9a).

L388-L390 This paragraph seems speculative. Please input FTIR data that supports these ideas (C chemistry from this dataset).

Response: This paragraph is meant to present hypotheses and will be reworded as such. This hypothesis was based on observations of similar EOC levels in CONV vs CONV+WCC, even though the CONV+WCC Gram+:Gram- ratio was lower than in CONV. Cover crops have been shown to increase DOC inputs in the literature (Steenwerth et. al 2008), and so a lower stress ratio in CONV+WCC could suggest a small trickle of DOC, enough to promote priming, but not enough to increase stocks as they may have done in the ORG system. We will include citations of our results and further literature to provide more evidence for our hypothesis of why there was no increase in C in the CONV+WCC subsoils.

1. Steenwerth, K. and Belina, K. M.: Cover crops enhance soil organic matter, carbon dynamics and microbiological function in a vineyard agroecosystem, Appl. Soil Ecol., 40(2), 359–369, https://doi.org/10.1016/j.apsoil.2008.06.006, 2008.

L388 What does "high variability of soil C measurements" refer to? Dry combustion measurements of total C are very consistent.

Response: We were referring to the variability in C measurements across a field, and during a growing season. Total C measurements in upper soil layers can vary by as much as 8% as part of normal seasonal variation (Wuest 2014), and detecting even large changes in subsoil SOC stocks can require a prohibitive number of samples (Kravchenko and Robertson 2011) due to the inherent spatial variability of subsoil SOC.

1. Kravchenko, A. N. and Robertson, G. P.: Whole-Profile Soil Carbon Stocks: The Danger of Assuming Too Much from Analyses of Too Little, Soil Sci. Soc. Am. J., 75(1), 235–240, https://doi.org/10.2136/sssaj2010.0076, 2011.
2. Wuest, S.: Seasonal Variation in Soil Organic Carbon, Soil Sci. Soc. Am. J., 78(4), 1442–1447, https://doi.org/10.2136/sssaj2013.10.0447, 2014.

Conclusion:

L406-407: "This was facilitated by increased soil macropores created by cover crop roots leading to higher rates of transport of soluble C". Macropores were not analyzed in this study, and no increases were found in hydraulic conductivity or aggregation, please clearly delineate quantified results versus hypotheses in this conclusion.

Response: We will reword the conclusion to reflect the data that we directly measured and also to better reflect the changes we have made to the paper - this will include acknowledgement of the limitations in our data, and drawing on previous research at Russell Ranch. We will more clearly delineate statements that are quantified results vs. inferences drawn from those results.

We would like to thank the reviewer for their detailed, helpful comments, and hope that we have addressed their concerns in our response.