

Community comment 1 (Anna Cates)

We are very grateful for the time you took to review our manuscript and for the valuable comments that will help improve the work. In this document, we have included each of your suggestions and have attempted to address them as best as possible (text in blue following each comment). All these comments will be taken into account and incorporated into the revised version of the manuscript.

What a fascinating dataset with these distinct phases. The authors do a good job of explain the analyses they did and constraining their interpretations appropriately.

I'd like to see more discussion of tillage as an activity pushing C into the intermediate pool, possibly? Is there something about the reorganization of soil physically that prompts that intermediate turnover? Plus, this evidence is really contrary to other work suggesting that conventionally tilled systems have more C at depth, and no-till systems should accumulate it in the surface (eg Baker 2007).

We appreciate this comment. Using the reference provided by the reviewer, together with additional literature review, we will expand this aspect of the Discussion and attempt to better relate our results to factors associated with the tillage system. Nevertheless, as stated in the manuscript, because several management factors changed simultaneously in 2008, it is difficult to establish a single mechanistic explanation for the decline in SOC stocks observed after that year. In this sense, the conclusions of this study should not go beyond showing that SOC can be restored in degraded sites through the adoption of conservation management, but that this increase is inherently vulnerable.

Regarding the possible negative effects of fully implementing no-till without any sporadic tillage, as stated in the Discussion section: (L456) "... it has been suggested that the incorporation of aboveground carbon inputs into the mineral soil may be slower (Cai et al., 2022; Six et al., 2004). This would result in a significant stratification of carbon inputs if biological activity does not promote its incorporation at a significant rate, thereby reducing the contact between this material and the soil mineral phase. This would lead to increases in surface particulate organic matter (...)" This particulate organic matter would be subject to a higher decomposition rate because of its lower capacity to become stabilized through association with the mineral soil matrix. This process would be consistent with our results. Nevertheless, other processes that may also be partly influenced by changes in the tillage system, such as potential reductions in C inputs and microbial carbon mining, may also be contributing to the different dynamics observed among periods.

It's not clear whether all phases of the experiment are present in all years? Especially if not, but even if so, one caveat I would introduce is whether interannual variability could be a driver? I.e. during the third stage, is it possible that pasture was present in years that were just less productive, so inputs were lower and C turnover higher due to more microbial activity but less to eat? Not having read the other work referenced, I'm not sure whether the authors have annual productivity data, but the sensitivity analysis could definitely get much more detailed. A 20% reduction in biomass inputs may have less effect in a warm + wet year, for example, with high microbial activity year round, rather than a year where decomposition was limited by water or temperature.

The experiment underwent changes in the staggering of rotation phases assigned to plots belonging to the same treatment (i.e., the same cropping sequence). Until 1974, the three plots within each treatment were synchronized, meaning that they went through the same phase of the rotation simultaneously. From that year onward, the rotation phases were staggered across the three plots, so that each plot represented a different and consecutive year within the six-year crop rotation. This was intended to partially dilute the effect of climate on plot productivity, because the same treatment was exposed to different environmental conditions across the three plots.

Nevertheless, the reviewer's point regarding the effect of climate on plot productivity could clearly have been one of the factors determining, at least in part, the observed results. However, in line with our response to the previous comment, this factor is confounded with others that may also have affected the dynamics of C inputs to the soil profile, including changes in the crop sequence, occasional removal of aboveground biomass, and the effects of the complete transition to no-till, among others. For this reason, the aim of this study is not to provide a definitive explanation for the cause of the decline in SOC stocks after 2008. This decline undoubtedly reflected an effective reduction in C inputs resulting from a combination of the factors mentioned above. Rather, we took advantage of the changes in SOC stock and radiocarbon trajectories to evaluate how the C accumulated during the recovery phase (1984–2008) was allocated among compartments with contrasting kinetics and persistence.

The processes responsible for the subsequent destabilization cannot be answered by this study. This is why we performed the sensitivity analysis showing that the changes occurred in 2008 cannot be explained solely by changes in C inputs, but also required a change in compartment kinetics, essentially involving the recently accumulated C in the intermediate pool, to reproduce the observed decline. Our main hypothesis is that microbial C mining of this recently stored C occurred in response to a multicausal reduction in C inputs.

L39-40 are there other important milestones along the way, like perhaps pre-industrialized ag in early 19th century? Also, does this metric include grassland forage or grazing?

We understand the reviewer's point regarding other important milestones. However, this reference was used simply to illustrate the substantial increase in agricultural land area over the course of human history. We consider this information to be merely contextual and illustrative, and we consider that further detail on this historical context is not necessary.

In response to the reviewer's question, Hu et al. (2020) report that croplands occupy 14.3% of the land surface.

L102-3 word missing or mistaken? Should and be “an”?

We appreciate the reviewer’s comment. There was indeed an error, “an” should have been used instead of “and”. This will be corrected in the revised version of the manuscript.

Fig 1 I think this figure might be easier to read if the boxes stretched over a number of years, or if year number was specifically labeled. It’s particularly confusing to understand whether the sorghum or wheat were seeded with the pasture the previous year, or if a one-year-old pasture was torn out to plant an annual crop again?

We appreciate this comment, which is entirely pertinent. Indeed, the figure is confusing for properly understanding the different phases of the rotation without a time reference. The figure will be modified in the revised version of the manuscript to address this comment.

Fig 3 This is a nice representation of model fit, particularly the insets

We appreciate this comment and we are pleased that the figure is visually appealing and useful for understanding the results of the study.

L 307 “The better ability of the three-box model to represent the data (Fig. 4c,d) occurred because most of the SOC stock variations were explained by C stock changes in a pool with intermediate kinetics between the compartments represented in the simpler model”. I’m not following this sentence. Most of the change was in the intermediate pool, right? So how was that represented in the simpler model?

What the sentence is intended to describe is that the three-pool model had a greater capacity to represent the data because changes in total C stock were driven by changes in a pool with intermediate kinetics, between those of the fast and slow pools represented in the two-pool model. This intermediate compartment is precisely the one that is not represented in the simpler model.

As stated in L443 with reference to the two-pool model: “In contrast to the more complex model, the two-compartment model failed to reproduce the full SOC and radiocarbon dataset, as the fast pool cycles too fast to preserve bomb C, while the slow pool strongly dilutes the new inputs into a ¹⁴C depleted pool.”

It is possible that the sentence cited by the reviewer has some wording issues, which we will address in the revised version of the manuscript in order to improve its clarity.

Table 1: is bolding intended to highlight something?

We appreciate the comment. The reviewer is correct that the use of bold formatting would not be appropriate in this case. This will be changed in the revised version of the manuscript.

L384- 395. I don't follow this explanation of the results of the input sensitivity analysis. To my eye, Fig 6 shows worse data fits than Fig 5 at both levels, particularly in stage 3, where observations actually decline while models held steady. How would reducing inputs in stage 3 leave you higher SOC values than maintaining inputs from stage 2? Is there more we need to understand re: rates of loss from each pool as well? Can the authors clarify more?

We appreciate this comment from the reviewer, which helps clarify what we intended to demonstrate with this sensitivity analysis. The question we seek to answer with this analysis is the following: Is the decline in SOC content after 2008 due to changes in the system's internal dynamics (changes in decomposition rates and the transfer scheme between compartments)? Or can it be explained solely by a possible decrease in C inputs?

In L 236, in the Materials and Methods section, we explain what was done: "To assess possible reductions in C inputs after 2008 due to the implementation of no-till that can affect SOC and radiocarbon dynamics, we performed a sensitivity analysis by extending the period-2 parameterization to the end of the experiment (periods 2 + 3) and imposing two reduced-input scenarios during period 3 (20% and 40% reductions in C inputs)."

So, what we did was to leave the parameters of Period 2 (recovery phase) unchanged during Period 3 (after 2008) and test the two C input reduction scenarios. Since the measured dynamics of C stock decline could not be adequately represented and no significant changes were observed in the modeled radiocarbon trajectory, we concluded that it is indeed necessary to consider a different set of parameters for Period 3, as the internal dynamics of the system changed for some management-related reason. Therefore, in response to the reviewer's question, the point of Fig. 6 is to show that the fit is poor after 2008 because the tested reduction in C inputs is insufficient and the system dynamics must also be altered to accurately represent the measured data. The system model with a new set of parameters for period 3 is shown in Fig. 4 and represents the results of the study itself.

The text beginning at line 471 of the discussion helps clarify this point:

"Nevertheless, a sensitivity analysis in which the parameterization of period 2 was held constant during period 3, while simultaneously reducing C inputs after 2008 by 20% and 40% in this last period, showed that such potential reductions alone cannot explain the observed decline in SOC stocks and have virtually no effect on explaining the decrease in $\Delta^{14}\text{C}$ (Fig. 6). Correctly reproducing the observed $\Delta^{14}\text{C}$ dynamics requires the predominant destabilization and loss of bomb- ^{14}C -enriched C from the intermediate pool, reinforcing the robustness of the finding that highlights the high vulnerability of C gains during the recovery period (1984-2008)."

L406-08: minor, but I'd recommend swapping the word order in the sentence to highlight, "it is possible" before giving all the experimental details

We appreciate your comment and agree with your suggestion. We will make the change in the revised version of the manuscript.

L 491-500 What about soil and climate relevance of these meta analyses to the context of this LTE?

Although none of the sites where the reported studies were conducted (He et al., 2024; Rui et al., 2022; Xu et al., 2020) have exactly the same climate and soil conditions as those evaluated at La Estanzuela LTE, these studies found similar findings under vastly different environmental conditions. In this regard, Rui et al. (2022) represent a cool, temperate, and humid site; Xu et al. (2020) represent a warm, temperate, and semi-arid site; while He et al. (2024) represent a cold climate with low air pressure typical of the Tibetan Plateau.

In this regard, we did not use these studies as exact representations of the La Estanzuela site, but rather as a general conceptual framework that shows that the processes discussed in the paragraph have been observed under various environmental conditions. The diversity of environments represented in the studies cited in that paragraph leads us to believe that they constitute an appropriate knowledge framework for discussing the results of this study.

L510-516 It's unsatisfying to claim we should promote practices that increase SOC in stable pools, without any concrete examples; instead saying that all studies have found it's difficult to increase SOC in stable pools. Can you say more about this challenge, or acknowledge that we don't know what those practices are, or discuss what you think they might be? I think the elements highlighted in the next paragraph, i.e. maybe the stage 3 practices didn't keep the microbial necromass flowing, might be a good place to start.

We appreciate the comment and we agree that the text of this section of the discussion should be modified to limit it not only to the results of this study but also to the state of the art regarding the challenge of increasing or maintaining high SOC stocks. So, we will modify this text and other parts of the discussion to limit these statements to what is actually known. We will then state that the literature indicates it is indeed difficult to increase the SOC stock in highly persistent compartments at rapid rates (i.e. over the time scale of decades associated with land management change in most cases) and that there is a high degree of uncertainty regarding agronomic practices aimed at this goal. Finally, the reviewer's suggestion to deepen the discussion regarding the implications of the management changes implemented in period 3 on microbial growth, the generation of microbial necromass, and their implications for SOC dynamics is very interesting. We will improve the review related to this aspect and include it in the discussion. This point is particularly relevant because, although there is no absolute consensus, this line of reasoning emerges as one of the most promising mechanistic interpretations for explaining the SOC and radiocarbon dynamics observed in our data.

L535 I'm surprised that no-till is on your list of practices to increase C, given that your evidence showed the no-till stage had decreasing soil C. Can you explain why this still looks like a valuable practice in light of your evidence?

The reviewer is correct in noting that this text does not fully align with the study's results and contains logical inconsistencies with the rest of the study. Based on the study's results, it can be interpreted that high rates of C input are necessary to sustain the previous increases in SOC stocks during the recovery phase. However, the results do not demonstrate that the shift to no-

till caused the decline in SOC after 2008, but neither do they support the conclusion that this management change alone was sufficient to prevent it.

We appreciate this comment and will modify the text in the revised version of the manuscript, as it is overly general and not sufficiently restricted to the results of this study.