

Comment on soil-2021-147

Anonymous Referee #2

Referee comment on "Soil nutrient contents and stoichiometry within aggregate size classes varied with tea plantation age and soil depth in southern Guangxi in China" by Ling Mao et al., SOIL Discuss., <https://doi.org/10.5194/soil-2021-147-RC2>, 2022

The manuscript describes a study of soil chemistry within different soil aggregate sizes at various soil depths across tea plant plantations ranging from 8 to 43 years old. Soil aggregates became smaller over time and with depth. Soil chemistry changes over time were most prominent near the surface and diminished with depth. C, N, Fe²⁺, and Mn²⁺ increased with age, Ca²⁺ and Mg²⁺ decreased with age, and P remained stable through time. Soil chemistry changes with soil depth generally occurred in the opposite direction of the changes with age. Although there were anecdotal differences in chemistry among aggregate size classes (e.g., mass fraction of C tended to increase with aggregate size), the changes in aggregate chemistry with depth and over time tended to follow the same patterns as those in the bulk soil. Changes in soil pH were related to Ca:Mg and Fe:Mn ratios, suggesting that soil acidification could be leading to preferential losses of soil micronutrients.

The manuscript currently contains six tables and nine figures, which seems a bit overwhelming for most readers to follow. I would strongly suggest that the authors attempt to reduce the amount of raw data presented by identifying the most important aspects of the manuscript. Clarifying the objectives and making the hypotheses more specific would help to provide this focus. In my opinion, the changes in the absolute values of soil nutrients are more relevant than the changes in the ratios of the nutrients (stoichiometry). Therefore, my suggestion would be to move Tables 2-6 into the supplemental and replace them with a single ANOVA table to summarize the stoichiometry findings (e.g., the last 5 rows of Table S1).

I believe it is important to provide more details about the site locations and management practices in the materials and methods section. Considering that one of the main aims of this manuscript is to quantify soil nutrient changes through time, it may be necessary to provide some details about the typical annual inputs (e.g., manure, inorganic fertilizers, and litter) and typical outputs (removal of tea for harvest), including approximate annual quantities and nutrient contents. It would also be helpful to provide more information about the site locations – for example, whether the tea plantations are managed by a

single entity or managed independently. The authors may want to consider bringing the Figure S1 map into the main text to help with this, as the map shows that the sites are randomly located in space, which helps to mitigate the concern of pseudoreplication.

The statistical analyses may require some additional considerations to be sufficiently robust. First, the current two-way ANOVA tests the effects of soil depth and time on the variables (i.e., nutrient concentrations or ratios) within each aggregate size class (e.g., > 2 mm). However, the authors draw many comparisons about differences in those variables among the aggregate size classes (e.g., L221-223), but these differences were not tested statistically. Therefore, my suggestion is to add aggregate size class as an explanatory variable in the ANOVA. Second, for each statistical test, comparisons are being made between all soil depths (4) and times (4), for a total of 16 comparisons. However, according to the statistical methods description, no adjustment is currently being made to compensate for these multiple post-hoc comparisons, and therefore the reported p-values are likely too small. To address this, I suggest using an accepted post-hoc adjustment for multiple comparisons such as Tukey's HSD test. Since this will likely change the significance of some effects, the results and discussion may need to be revised accordingly.

The manuscript requires revisions for grammar.

L1, elsewhere: Suggest changing "stoichiometric characteristics" to "nutrient stoichiometry."

L1, elsewhere: Suggest changing "tea-planting age" to "tea plantation age."

L9, elsewhere: "sort of effective way" could be "a tool."

L14, elsewhere: "at the aggregate scales" could be "within aggregates."

L22-24: Leaching was not measured in this study, so it seems overreaching to include this in the abstract.

L24-27: The comparison of C and N to other tea plantations is somewhat arbitrary, as soil types may be drastically different among the plantations. I suggest removing these sentences.

L30-32: What is the cause of soil acidification, and how could it be mitigated?

L30-32: Since Mn²⁺ toxicity was not measured in the study, it does not need to be mentioned in the abstract.

L34: This only provided information about the tea-plantation agroecosystem, not "global terrestrial ecosystems."

L56: The term "lower ratio" is ambiguous. Please indicate what is considered high vs. low.

L66-68: Did this study measure total C or organic C? If there were carbonates present at depth, then the total C:N ratio would be much different than the OC:N ratio.

L85-99: This paragraph seems like it might be more informative at the beginning of the introduction.

L107: "assumed" could be "hypothesized."

L129-133: Most readers will be familiar with space-for-time substitution, so it is probably not necessary to describe the concept in great detail here.

L139-140: "Each of the four tea plantation age groups was replicated in five locations for a total of 20 experimental units"

L142: "space self-correlation" could be "spatial autocorrelation."

L146-148: Strictly speaking, this description indicates that the study measured surface litter (a stock), rather than litterfall (a rate). Measuring litterfall would require keeping the newly falling litter separate from the existing surface litter (e.g., litterfall traps) and measuring over a certain period of time (e.g., 1 year).

L152: What was the surface area of each soil sample?

L156, elsewhere: "subjected to filtration" could be "sieved."

L158-159: Please define all symbols and abbreviations (V and Ø).

L165: Should "vertical" be "horizontal?"

L179, elsewhere: "abstracted" could be "determined."

L190: Are the "parallel specimens" check standards?

193-197: Please clarify how the mean weight diameter of aggregates is being used to infer aggregate stability.

L202: Please state the alpha value (e.g., $p = 0.05$) used to determine significance.

L205-265: P-values should be stated throughout the results section. Any significant interactions should also be noted.

L207, elsewhere: "remarkable" or "remarkably" could be "significant" or "significantly" if the implication is that they are statistically significant.

L224, L249: It is not clear which metric these values were "elevated" above.

L231, elsewhere: Differences in nutrient concentrations and ratios among aggregate sizes classes were not statistically tested. Please include statistical tests.

L233: It is not clear which source of variation is being discussed with the statement "did not show remarkable variation." For example, does this mean variation by age, by depth, or by aggregate size?

L243-244: "mainly distributed" could be "more concentrated."

L256-257: "ratios were evenly distributed in" could be "did not vary among."

L262-265: Should Figure 9 be described here?

L270-273: What are the potential explanations for higher coarse macroaggregates in the 17-year plantations compared to the other ages? Are the younger plantations managed differently than the older plantations in way that would lead to this outcome? Or was there something specific about that age-group that made them different than the others (e.g., more manure was applied for several years prior to this study)?

L318, elsewhere: "organic matters/OMs" could be "organic matter/OM."

L322: Please clarify what is meant by "propelled the causal links."

L341-346: Changes in soil texture are longer-term processes that would not be expected to change over this time period.

L436-438: As an alternative to leaching, how much Ca and Mg might be lost during tea harvest compared to how much is being added by manure or fertilizer? How do these factors change with tea planting age?

L485-503: The conclusions section is nearly identical to the abstract. I would suggest revising the abstract to include a broader opening to provide context for the study (i.e., land use change and tea plantations) and reducing the scope of results given in the abstract.

All figure and table captions: Please specify whether the comparisons among different tea plantations ages (capital letters) are made within each soil depth, and whether the comparisons among each soil depth (lower case letters) are made within each plantation age.

Table 1: Were other nutrient concentrations measured in the litter (e.g., P, Ca)?

Figures 2, 3, and 4: I suggest converting these line graphs into bar graphs (like figures 5-8) for clarity and consistency.

Figure 9: The regression lines could be colored to match the soil depths to improve interpretability.