EGUsphere, referee comment RC1
https://doi.org/10.5194/egusphere-2022-562-RC1, 2022

Comment on egusphere-2022-562

Enrico Balugani (Referee)

Referee comment on "Biotic factors dominantly determine soil inorganic carbon stock across Tibetan alpine grasslands" by Junxiao Pan et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-562-RC1, 2022

Dear Authors,
I read your manuscript with great interest. In my opinion, this article provides interesting data and insights on SIC dynamics, and thus fits the SOIL journal aim and scope. The article is well structured and organized, and written in an acceptable English. The study is about the results and analysis of a large-scale soil sampling campaign in the Tibetan plateau, aimed at investigating the soil inorganic carbon density in topsoil (0-10 cm) and subsoil (20-30 cm) and their correlation with various explanatory variables selected by the Authors. As such, I think the study should be published in SOIL, after some fixing of the manuscript, especially of the materials and methods and result section for what concerns the analysis of explanatory variables for SIC dentisty.

Major comments.

I think that the introduction is quite clear, but could be improved by: (i) giving an idea of the relative relevance of SOC/SIC pools, just to put things in perspective for readers; (ii) give some definition of "top" and "sub" soil to the reader (different researchers may divide the soil profile in different ways).
The materials and methods section is mostly good, but I have a few comments/reservations: (i) did the Authors think that taking part of the samples in July, part in August, and part in September, may have had an effect on the results? For example, maybe soil pH and microbial abundance vary during summer, and thus there is another explanatory variable not taken into account (temporal variation). I suggest to include the information on date of sampling in Table S1, and discuss this issue in the Discussion section. (ii) The subsection "Statistical analyses" need improvement, in my opinion; more specifically, it needs to be more rigorous. First of all, a clear list of all the explanatory variables taken into account should be given, and a clear definition of which goes into edaphic, microbial, plant, and climate - and also biotic/abiotic. Then, there is a question: why didn't the Authors study the correlation index for each variable with respect to the target (SIC)? Spearman and Pearson correlations could be used, and give a clear picture to the reader in a simple table. Then, instead of selecting the most relevant explanatory variables to build the multi-linear model, the Authors decide to create a large model with all explanatory variables; this I can understand, but the reason for this choice

vs the former should be given. As far as I understand, the Authors create a "theoretical" multilinear model with all explanatory variables the Authors identified, and then the Authors assess the relevance of each explanatory variable. In my experience, this is often done using Global Sensitivity Analysis techniques, to take into account joint effects and different orders of sensitivity (see Saltelli 2008 sensitivity analysis a primer). However the Authors use another method, called Varation Partitioning analysis - that is okay with me, but this method should be explained further.
The results section has the same problems of the previous section: it needs to be more rigorous on the statistical part. Please avoid confusing statements as "positively associated" and "negatively correlated" - if the Authors study correlation, then both are correlated, either positively or negatively. The significance of a correlation should be given (it is in the figures 2 and 3, but not in the text). It is very important that statistical techniques are important to have a common ground, thus my focus on rigor, but they do not give clear-cut answers: the difference between topsoil and subsoil relevance of explanatory variables is not as big as it appears from the text of subsection 3.3, as it can be seen from the figures 2, 3 and 4. Figure 2 and 3 show that the SIC values are mainly clustered by "grassland type" (i.e. AM, AS and AD), so much so that probably the best predictor would be just to consider the grassland type, something that could be done using Remote Sensing (not very informative about processes, I admit). This should at least be discussed in section 4. Figure 5 is very confusing, since the PVA method was not explained I cannot decript it myself, the caption is not very informative: how did the Authors get the percentages in the caption? What do the Authors mean with "unique effect" and "common interception"? What are the percentages in the figure? What do the Authors mean with "residuals"?
The discussion section has also a problem with rigorousness: the Authors should clearly divide between the results (which show statistical correlation between variables) and their interpretation and speculation (the cause-effect relationship, which has not been studied here). Also, the Authors should make it clear that the results are valid only withih the study area, and every extrapolation to other areas should be done very carefully.
The conclusion could be improved by directly answering the two questions at the end of introduction at the beginning, and clearly dividing results (of the statistical analysis) from interpretation.

Minor comments:

line - comment
15 - start with "The"
23 - "associated" do the Authors mean "correlated"?
51 - substitute "where" with "which"
84 - create a new paragraph for lines 84-91
122 - "the rest of the samples, about 700 g, were also [...]"
123 - "other soil properties" which ones?
127 - substitute "into" with "for"
128 - substitute "gravels" with "material"
166 - "[...] terminated. The primer [...]"
167 - substitute "Then" with "Finally"
219-220 - "(Fig.s 2 and 3 for topsoil and subsoil, respectively)"
220 - again "associated". Not clear what the Authors mean, and how it relates with "negatively correlated"
241 - substitute "afford" with "study"
244 - substitute "Since" with "Due to"
282 - substitute "more" with "larger"

293 - "association"?
304 - "[...] fungal gene abundance are correlated with SIC stock [...]"
308 - "and the increase in acidity is neutralized through [...]"
350 - substitute "less roles" with "a lesser role"
355 - substitute "significance" with "effect"
356 - substitute "maintains" with "is important to maintain"
357-358 - "[...] biotic factors are correlated with SIC stock in the Tibetan plateau [...]"