



EGUsphere, referee comment RC2
<https://doi.org/10.5194/egusphere-2022-1034-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on egusphere-2022-1034

Anonymous Referee #2

Referee comment on "Shapley values reveal the drivers of soil organic carbon stock prediction" by Alexandre M. J.-C. Wadoux et al., EGU sphere,
<https://doi.org/10.5194/egusphere-2022-1034-RC2>, 2022

Shapley values reveal the drivers of soil organic carbon stocks prediction

This title is too methodological and provides no meaningful insights regarding what this study is reporting. I think the title will be meaningful if written as:

Elevation, vegetation and temperature determine the spatial variation of French SOC stocks

Identifying relationships between environmental factors and SOC stocks is an important topic of scientific investigation. In this study, authors used soil samples from 2206 sampling sites and data of 23 environmental factors from France to predict the spatial variation of SOC stocks of 0-50 cm depth interval. Authors investigated how the correlations between SOC and environmental factors vary across the prediction points of the study area using "shapely values". Authors reported that topography, reflectance property of vegetation (NDVI), and temperature primarily explain the spatial variation of French SOC stocks. I think authors are attempting to address an important topic, but this manuscript needs substantial revision before it can be published.

There has been a number of SOC stock studies previously published from France, which have reported relationships between environmental factors and SOC stocks. Authors should compare their findings with previous studies and explain how and why their result is different and novel. Authors should also report whether they used the soils samples used by previous studies, and which findings are new in this manuscript. To merit for publication, authors should explain what are new findings in this study that is not available in previous studies from the same study area.

I am not comfortable in authors using "process-based" modeling phrase repeatedly in this manuscript. In this study, authors did not use any process-based model, nor they report any new soil carbon regulating process, so it's just a pure distraction. There is a long and rich history of SOC process-based modeling literature where studies attempt to predict the temporal dynamics of SOC under changing land use and climate, which is not within the scope of this manuscript.

In summary, I found this manuscript as prepared in rush, and does not report any interesting mathematical relationships between environmental factors and SOC stocks, which can be used to predict the SOC stocks. The manuscript is not focused and sentence structures need substantial revision before it can be published. My comments below are intended to improve the quality of this manuscript.

Abstract:

I am not aware about the word limitation in the Abstract for this journal, but currently this abstract is more than 350 words and could be reduced substantially by deleting unnecessary texts. The abstract is not structured and should be rewritten. By reading the abstract, I couldn't understand what was the relation between RF and shapely values, and why both are used in this study.

L4-11: These sentences describe the methods used in this study. Please replace these sentences and describe your methods briefly in 1-2 sentences.

L7: Please define what is shapely values, and why someone should care about it?

L8-9: "what is the". This sentence is not correct. The relationships shown in Figure 3 are relationships between "shapely values and environmental factors", and not the "relationships between environmental factors and SOC stocks", which are not the same. Authors need to clarify this statement.

L10-12: In my understanding, this study reports correlational findings which may or may not be related to any soil carbon regulating processes, so I am not sure what "Results were validated both in light of the existing and well-described soil processes mediating soil carbon storage" means?

L13-16: Again, these relations are based on correlations and does not provide any process-based understanding.

L16: "This shows..." I think this sentence does not report anything and not relevant in Abstract.

Introduction:

Introduction section should properly cite and discuss recent and relevant studies in this topic. I assume there are a number of studies which have attempted to explain the control of environmental factors on SOC stocks. Discussing the findings of these studies will strengthen this manuscript:

Mishra et al. 2022. Empirical relationships between environmental factors and soil organic carbon produce comparable prediction accuracy as the machine learning, *Soil Science Society of America Journal*, doi:10.1002/saj2.20453.

Gautam et al. 2022. Climate change may release over 1.8 petagrams of soil organic carbon from topsoils in the United States by 2100, *Global Ecology & Biogeography*, 31, 1146-1160, doi: 10.1111/geb.13489

This is a spatial prediction study with no contribution to process-based modeling. So, the texts refereeing to process-based modeling is not relevant in this study and should be removed. I suggest discussing findings of additional studies which have reported mathematical relationships between environmental factors and SOC stocks.

L35-36: "Dynamic modeling...". This sentence is not relevant to the content of this manuscript.

Materials and Methods

Figure 2& 3: Please level the Y-axis in both figures, and provide units in both X and Y axis. Figure 3 does not provide any information regarding the relationships between **environmental factors and SOC stocks**, and I am not sure the scientific merit of these plots. Are these relationships additive, and can be used to predict the SOC stocks?

L 391: "We found". This sentence suggests there were no new meaningful insights in this study.

L 444: Delete "Varied" from the sentence.

L450: Delete "full stop" from the middle of sentence.

References:

This is not a "literature review" manuscript, therefore, I encourage authors to give priority to recent literature of SOC stocks. For example, using studies published in the last 10 years in this topic unless the study is published from the same study area or have used the same set of samples.