Authors response to Reviewer #2 (RC 2) for soil-2021-81

This manuscript presents the distribution of SOC and its labile fractions predicted using parent material, land use and soil properties in Southwest Germany. The results indicated that soil properties were clustered by parent materials and soil texture rather than land use. In general, mixed-effect model gave better predictions than bivariate regression. They compared "global model" with "local model" to show that the application of global model on local dataset introduced poorer predictions. Also, the explained variance generally decreased from bulk SOC to its labile fractions.

In general, the objectives were clear and relevant while the scientific value is sufficient. The large sample size contributes to a robust prediction. However, there are several concerns to be addressed.

One concern is the distribution of the sampling points. As mentioned in L47-48, soil formation is also controlled by climate and topography. The clustered locations of the four parent materials are likely to introduce differences in topographical and climate conditions. As climate and topography factors were not included in the models, their effects might be recognized as the effects of parent materials, texture or land use in the predictive models. (Details in comments for Fig. 1)

See answer to comment on Fig. 1

Another concern is that the usage of "global/local scale", "global/local model", "global/local cluster" and "global/local/entire dataset" may confuse readers because they were used without necessary explanations. In addition, the words "global" vs. "local" give the impression that the study aimed to compare SOC distribution on global vs. local scale, but no investigation on global scale was given in this study.

Thanks for your comments. We fully agree that the term 'global' may be confusing. Therefore, we decided to replace 'global' by 'total' to avoid this misunderstanding.

In addition, in some parts of the manuscript, R^2 was used to estimate whether models are well-fitted, which is not proper. Also, the Results and Discussion can be improved by splitting them into sub-sections and better re-organizing. Finally, the readability of the manuscript can be improved by revising long-complexed sentences and vague expressions.
As similar mentioned to Reviewer #1, R-square is used to show the explained variance, this manuscript aims to show how much mineral phase parameters and their different combinations are able to explain the variance of SOC, HWEC and MBC. Notwithstanding, we fully agree that the root mean square error is a much better measure to determine the model performance. Therefore we added it to the text.
Title: (1) Although "soil organic matter" is used in the title, the main part of this manuscript is mostly talking about "soil organic carbon". Please be consistent in using them because soil organic matter contains not only organic carbon but also other elements such as nitrogen.
We agree that a consequent use of the terminology is required. Therefore we changed the title to "soil organic carbon"

(2) It is advised to add restrictions on the area/location because the study was performed in western Germany and will not be necessarily applicable in other places.
We agree that local areas were sampled that were all located in the larger region of Trier in southwestern Germany. Anyhow, it was not our primary intention to characterize a specific region. Instead, the sampling region was selected because it covers soils with identical land use types (arable and grassland) and similar climatic/pedoclimatic conditions but substantially different parent material, and thus different soil mineral phase properties. Our prior aim was to show that SOC of local clusters is better assessed using local models. It is clear that at larger scales (nation-wide and larger) differences in pedoclimate add to the factors explaining SOC (and labile fractions). However, the climatic impact is generally not relevant for local areas independent of where they are. At the same time what the reviewer states is exactly would we found and suggest: A specific assessment of local areas with a local model is preferred. Such a local model is not transferable on to one to another local area.

L14, L18 and L21: It is confusing to mention "local scale", "global/local cluster" and "global/local dataset" in abstract without further explanation. The usage of "local" vs."global" gives me a feeling that this study compares SOC distribution on local vs. global scales. Apparently, the distribution of the sampling sites represents a local or sub-regional scale. It is suggested to either give them definitions when they are mentioned for the first time or replace them with more suitable words.
Thanks for your comment, we now define what is meant by local and total (previously global) cluster/dataset and we replaced the scale-related terminology.

L21: As only regressions were performed in this study, it is recommended not to use both correlation and regression in the text.
We revised the text accordingly.

L21-23: It is difficult to understand this sentence. It is not clear between which factors the correlations are significant. What does "partially low" mean? Splitting this sentence into simple ones may help.
For a better understanding of the sentence, we followed the advice to split it. 'Partially low' means that some of the correlation coefficients (R^2) only showed a small explained variance. Such vague terminology was replaced by more objective wordings.

L66: In general, organo-mineral associations are considered contributing to the formation of stabilized fractions (not labile fractions) and therefore the accumulation of SOC.
We agree that the formation of organo-mineral associations leads to the stabilization of SOC. Additionally, such accumulation of SOC goes along with increasing contents (not stabilization) of labile fractions such as DOC that are only weakly retained through other mechanisms in the presence of pedogenic oxides. This is what we wanted to say. We changed the sentence as follows to make it more clearer: "Organo-mineral associations are highly relevant for stabilization and accumulation of SOC and also for the accumulation of its labile fractions (Lützow et al., 2006)."

L72: …leading to SOC sequestration…
We adapted this line.

L85: Please check if surnames and given names are misplaced in this reference.
Checked, there is no misplacement of surname and given names. Surname is Jian-Bing, given name is Wei. See https://link.springer.com/article/10.1007/s10661-005-9158-5#article-info

L90: "Local vs. global models" are confusing. Do they mean models on local vs. global scales?
We wanted to indicate that it is necessary to apply models on local clusters/datasets instead of on one global (total) dataset to best explain SOC (more precisely its variance).

L102: Is the “entire dataset” equivalent to the “global dataset”?
A global (now total) dataset is defined as a dataset encompassing the large majority of the dataset. Therefore, next to the entire dataset, the clusters of arable, grassland and loam act as total (previously global) dataset.

Materials and Methods
L104: It is recommended to add more information about the study area. In general, most studies show readers climate factors (e.g. annual precipitation and average temperature), soil type/classification and composition of vegetation/crops.
We added additional information regarding the study area.

L119: Please explain why soil samples were stored either at -20 °C or air-dried. For different analyses?
Samples were stored until they were analyzed. Storage was done in a uniform way for all samples. One part of each sample was air dried for subsequent chemical and physical soil analysis, another part was kept moist and was frozen for subsequent soil microbial analyses (MBC, MBN or respiration). This is now clarified in the text.

L139: More information of the incubation is appreciated. How long the samples were incubated before sampling? What was the temperature? Did you sample for only once or multiple times?
Samples were preincubated at room temperature for one week (7 days), measurement was conducted for 24 hours at an interval of one hour. The information was added to the text..

L146: Please give more information of linear regressions. For example, indicate that they only have one predictor. Did you check the normality of residues?
We added respective information. Normality of residues was checked.

L147: and after: What are the reasons for performing mixed-effect models? Why parent materials, texture group and land use are selected as random effect variables? In general, random effects are used when samples are only a small subset of the group or when limited groups are included. Does it aim to make predictors on a larger scale using the limited dataset?
West, Welch, & Galecki suggest in their book “*Linear mixed models*” that such models can be applied to clustered data. We decided to use mixed effect models to capture the effect of soil properties applied as fixed effects. Developed on different parent material or under different land use management the soils showed a further source of variability. Furthermore, we selected these variables as random factors. It is aimed to remove their bias from the specific levels of the applied random factors. To this end, sampling sites were specifically selected that covered the factors parent materials, texture group and land use.

L162-163: Why was response variable transformed but not predictors?
Transformation of the response variable is common and was applied to achieve a normal distribution of the residues. The predictors of the mineral phase determine the variability of SOC and its labile fractions. Therefore we tried to keep them as they occur in the environment/our dataset. SOC (or its labile fractions) as variable part of the soil was consequently transformed to achieved normally distributed residues.

Result

Overall: The readability can be improved by dividing this section into a few subsections due to a large content in this section.
Thanks for this valuable advice, we divided the results in subsections.
3.1 Soil properties and cluster identification
3.2 Bivariate relationships of mineral phase and SOC and its labile fractions
3.3 Estimation of SOC and its labile fractions by mixed effect models
3.4 Comparison of total and local explained variability.

L170: What are “soils and topsoil properties”? Consider revising.
Line was revised. Topsoil was separately mentioned due to the fact that our study is focused on agricultural topsoils. To avoid confusion or misunderstanding we decided to use only the term ‘Soil properties’.

L177-178: Are they significantly different or different by looking at means/ranges?
Differences were mostly statistically significant differences. Here we solely wanted to mention that a higher proportion of organic substance was found in grassland soils compared to arable soils.

L190: “Somewhat different” is vague.
It was changed accordingly.

L205: and after: This paragraph is comprised of isolated points, which makes it difficult to follow. A suggestion is to describe Table 3 in a well-organized way to shorten this paragraph. For example, you can follow the order of entire dataset --> land use --> parent materials --> texture, or you can introduce them by the types of predictors. Also, focusing on your key findings helps.
We will try to rephrase it paragraph. Anyhow we include subsections.

L207-208: The items “global cluster” and “local cluster” are explained here but they appear in previous parts (e.g. L18 and L193). Please give explanations when they appear for the first time.

Thanks for this hint, we mention the definitions earlier.

L208 and L94: Please be consistent for “parent material” or “parent rock material”.
We changed it. Now it is consistent.

L224: What is “a sufficient extent”? Please specify.
Applying “sufficient” is not objective enough, therefore we rephrased this line and specified it by giving the respective level of R^2.

L237 -242: Please indicate that they are from Table 3.
The reference to Table 3 was added.

L240: How to know “weight of samples” is equal? Why does it act as global cluster?
We mean the statistical weight of the samples. We changed the sentence as follows: “The clusters of both land use types largely overlapped and contained a similar proportion of samples from each parent material. Therefore they can be regarded as total clusters.”.

L250: It is not clear how to compare R^2 between bivariate regression and mixed-linear model. By the means of each cluster?
Next to the comparison of the explained variance we showed the RMSE to give a measure for model performance. This is now better clarified in the revised text. "By the mixed effect models, R^2_{cond} reach higher explained variance for SOC (R^2_{cond} = 0.39-0.89, RMSE = 0.21 – 0.42%) compared to the bivariate regressions (R^2 = 0.00-0.73, RMSE = 0.27-1.12%)." Further we added some information at section 2.3

L257-258: DCS sites look different from LBS and DLS.
We clarified it. "Models using parent material or texture as random effect mostly showed minor differences for predictions of SOC, HWEC or MBC. Anyhow, for some local clusters (e.g. DCS, LBS and DLS) distinct results were found. Models using land use as random effect were partly distinct, though, indicating the different influence of land use on SOC and its labile fractions (Table 4).
"

L279-286: Please indicate related Tables and Figures. It is hard to follow.
We now refer to the considered Tables.

L282 & L287-288: This gives me a feeling that you are estimating whether the models were well-fitted. If this is true, comparing R^2 does not make sense. Large R^2 means more variation is explained by predictors. Instead, you have to look at the distribution of residue using e.g. root mean square error (RMSE).
It is aimed by this study to show how well the specific models with their specific parameter combination explained the variance of SOC, HWEC and MBC. Therefore, we rephrased this sentence. We agree that in order to show the goodness of the model fit RMSE is the correct measure. We added this information.

Discussion

L304-305: "for the in total very sandy soils …of LBS". Try to revise this sentence.
Sentence was adapted.

L309: "…SOC in soil" --> "in soil"
It was changed accordingly.

L314-315: "ECEC, Ca and Mg are suitable predictors for SOC in this study"; L317-318: "The minor ability of ECEC (Ca+Mg) to explain SOC.." They look like contradictory. Also, I missed a point that whether you are talking about entire dataset or specified cluster. Table 3 showed that the predictions using ECEC and (Ca+Mg) are largely dependent on parent materials and texture cluster. A possible explanation is that DCS soils had more sands and lower pH, so that Ca and Mg do not contribute to SOC stabilization, whereas DLS and PSS soils had higher pH, so that Ca and Mg bridging play a role in SOC stabilization (see your cited paper). Please consider re-organizing this part.
We clarified these sentences. (L314-315)"The minor ability of ECEC and $(Ca+Mg)_{ECEC}$ and the higher ability of pedogenic oxides to explain variance of SOC and its labile fractions indicated in this study for several cluster (total and local) by bivariate regressions (Table 3), corresponds to findings of Rasmussen et al. (2018)."……
(L317-318)" Ability of ECEC and $(Ca+Mg)_{ECEC}$ was further strongly dependent on the observed parent material or texture cluster. By the mixed effect models, $(Ca+Mg)_{ECEC}$ were more frequently identified as relevant to explain SOC and its labile fractions. Thereby it is shown that by a collective approach of several soil parameters more driver explain a larger part of the variability than by bivariate approaches. As example ECEC and $(Ca+Mg)_{ECEC}$ was found as relevant for the clusters of DLS and PSS, while for DCS it show a minor importance."

L328-333; Grassland had higher SOC contents than arable land, but the PCA showed that they were largely overlapping. This is a good point for discussion. Some explanations will be appreciated.
We amended the Discussion accordingly. "In comparison, mineral phase soil properties clearly separate the dataset while composition of SOM was less enabled for this purpose. Consequently, a broad scatter of the land use clusters was obtained by PCA, suggesting to treat the land use clusters as total datasets as well."

L334-336: "Several studies with…" has only one citation?
Thanks for this hint, we added further studies.

L351-352: Previous explanations are good reasons for using multiple parameter models. However, the reasons for using mixed-effect linear model are not well mentioned. For example, why not multiple fixed-effect model or partial least square regression? My recommendation is to stay in a safe way.
As mentioned above, West, Welch & Galecki suggest in their book "*Linear mixed models*" this type of model for clustered data. Soil parameters (e.g. pedogenic oxides, texture) have an influence (with differing strength) on SOC, HWEC or MBC. Furthermore, there is an effect by factors such as the parent material or land use. To further capture this effect, we decided to use mixed effect models.

L373-374: To be prudent, I would say models of parent materials explained more variation of SOC because we don't if the model-fitting was better than others (see comments on L282). The same for L374-375.
We changed the sentences to highlight the explained variance.

L379 and after: A major finding of this study is that the overall explained variance decreased in the order SOC>HWEC>MBC. Some explanations for this would be appreciated.
Ok, we added further explanations.

L395: Please be consistent with "mixed effect model" and "mixed parameter model".
Ok, thanks for this hint. The relevant line was changed to mixed effect model

Figures and Tables:
Fig. 1 The clustered locations of the four parent materials are likely to introduce differences in topographical and climate conditions. For example, DCS and LBS sites are mostly located on the top of the mountain/hill, whereas PSS sites are located in a flatter area. The difference may affect soil formation and SOC accumulation. Also, the different altitudes between DCS and PSS sites may cause differences in climate conditions. Therefore, it is possible that the variation caused by climate and topography factors was explained by parent material or land use in this study. I just wonder whether something has been performed in experimental design, statistics or anything else to deal with this problem.

It was aimed by this study to estimate the effect of the mineral phase. Selected sampling region covers soils with identical land use and similar climatic/pedoclimatic condictions but the parent material is substantially different. Consequently soil mineral phase properties differ largely between the local sampling clusters. Our prior aim was to show that SOC of local cluster is better explained by local models, at larger scales we fully agree that differences in pedoclimatic conditions were factors needed to explain SOC (and its labile fractions). For local areas the climatic factors is generally not relevant independent of where they are.

Table 1: What does the unit for respiration mean? As suggested for L146, more information of the incubation is needed.
There was a typo in the unit. We corrected it: [µg CO_2-C/(g dry matter h)]
We added more information regarding the incubation.

A suggestion for Fig. 2: Why not combining Fig. 2 and Fig. S1 if you want to show the readers that parent material and soil texture make good separations while land use make an insufficient separation?
We tried this option. However, with three plots on one page the readability of the individual plots was poor. So we decided to leave it as is with a focus on the two plots showing differences between clusters

Fig. 2 and 5: The shape of the font might be improved as some of them are narrow but others are wide.

Ok, we adapted the fonts of these figures.

Fig. 3 It looks like that the residue of MBC is less normally distributed compared to SOC and HWEC. Particularly, MBC in grassland soils is underestimated. Also, HWEC has a similar but less obvious trend. My questions are: (1) Is the model prediction of MBC less reliable than others due to the skewed distribution of residue? (2) Are there any reasons for the underestimation of MBC in grassland soils?
For SOC, HWEC and MBC there is a trend of underestimation for grassland sites which increases from SOC to the labile fractions. We assume that additional soil properties (e.g. content of fine root biomass) affect the organic matter here. Since this study is focused on mineral phase parameters and did not consider further biological properties models were less suited to explain SOC and its labile fractions in grassland soils.

Fig. 4 What do “dataset”, “DCS”, “sand” and “arable” on the left mean?
It shows which cluster is shown there with its models. We added information to the figure in order to clarify it.

Table 5: Does the “model” before “global model to local cluster” mean local model?
Yes, this model means the R^2 which is received by the consideration of predicted vs measured data of the models for the specific clusters/datasets. We clarified it

Fig. 6: Is it a part of Table 5? Is there any reason to make it a new Figure? Maybe try to combine Table 5, Table S3 and Fig. 6 into a good shape, or move unnecessary information to supplementary.
Fig. 6 shows the performance of the total model and the respective local model, when both are applied to the same local dataset. This was tested by comparing measured and modelled data based on simple linear regression. This yielded a pseudo R^2. This information is contained in Table 5. We added information regarding RMSE at Table 5 and Figure 6.