SOIL Discuss.,
https://doi.org/10.5194/soil-2020-75-AC4, 2021

Interactive comment on “Predicting the spatial distribution of soil organic carbon stock in Swedish forests using remotely sensed and site-specific variables” *by* Kpade O. L. Hounkpatin et al.

Kpade O. L. Hounkpatin et al.

ozias.hounkpatin@slu.se

Received and published: 24 February 2021

Dear Reviewer,

Please find enclosed the revision of our manuscript “Predicting the spatial distribution of soil organic carbon stock in Swedish forests using remotely sensed and site-specific covariates” by Hounkpatin et al. We thank you for the very competent review of our paper and the productive comments towards the improvement of our manuscript. We agree with almost all of them and revised the manuscript accordingly.

We hope that our paper is now acceptable for publication in SOIL.

Yours sincerely,

Ozias Hounkpatin

Answers to reviewer

The manuscript attempted to compare the model performance of global and local models in predicting humus layer, mineral soil and total SOC stock and to identify the controlling factors for SOC stock prediction. Besides, this study also investigated the effect different combinations of data from site characteristics and remotely sensed variables on model performance. The results from independent dataset indicated that the local models generally had better model performance than the global models. The only use of remotely sensed variables had limited predictive ability while site characteristics had better explanatory strength in estimating SOC stocks. The authors suggest that further work can focus on mapping these influential site covariates. The manuscript is overall well-written with clear objectives and reasonable methodology. However, two major limitation of this manuscript are: (1) in comparing global and local models, I can tell from Table 4 and Table 5 that some local models had higher R2 than that of global models while the remaining local ones had lower R2. Therefore, the conclusion that local models have a comparative advantage over global models is not convincing for me. Again, it is not fair to compare the performance indicators as mentioned in this study for local and global models. Instead, for the global model, authors should also calculate indicators for three regions separately to make them comparable for these in local models; - Author′s Response: The error metrics were actually computed based on the same validation set for both global and local models, this to avoid comparison bias. Table 3 actually shows the error metrics for the global models when the independent validation set for the (different strata) North, Central and South Sweden were merged. Table 4 shows the specific error metrics for the local models against their respective validation set. Figure 2 presents the error metric results of each global model for the three subareas along with corresponding local models (unfortunately repeating error metrics of local models from Table 4 for the independent validation. To ensure clarity, we provided more details in the section (see lines 224 – 231): related to the data splitting procedure.

- Indeed some local models had higher R2 than that of global models while the remaining local ones had lower R2. However, we considered their performance in relation to the specific category of predictors the models were built from. In that regards, our conclusion was that in general local models with site specific covariates only or those built using all the covariates did in most cases perform marginally better than the global models.

Manuscript lines (see lines 224 – 231): The RF models built on data covering the whole area of Sweden are hereafter called "global models". The RF models created for each of the subareas are hereafter reported as "local models". Considering the subareas as strata, the local models were built by randomly splitting the local datasets into calibration (80%) and validation (20%) subset. The training set of the global models was constituted by merging the 80% split of the strata dataset. To avoid comparison bias, both the global and local models were evaluated on the same validation set. Consequently, each global model was evaluated again three validation subset separately corresponding each to the 20% split of the Northern, Central and Southern local dataset. The merged 20% split of these local datasets was then used as validation set at national scale. We trained both global and local models based on tenfold cross validation with 5 repetitions using the R "caret" package (Kuhn, 2015). (2) As mentioned by authors, the site characteristics are only available at the visited plots, therefore the digital maps can not be produced by the models built with site characteristics, which certainly limits the usage of site characteristics. I am afraid that the importance of these site characteristics in this study have been overlooked as the observed site characteristics are directly used in independent validation which certainly ignores the inaccuracy of these site characteristics if they are mapped by certain algorithms. I mean, the site characteristics used in a fair independent validation should come from the predicted maps of these relevant site characteristics, not from observed data. Therefore, I suggest a major revision before it can be published. - Author′s Response: We share the concern of the reviewer that all covariates should actually be maps for DSM of the whole area while the site specific data were only observations. Initially, the focus was mainly evaluating the geodata we now have versus site specific variables and not necessarily making a map. We made the map based on the geodata to give a visual trend of the distribution of the SOC stock. We also agree that the current study neglects to consider mapping uncertainties from potential maps of site specific variables which might actually translate into lower prediction accuracies of SOC stock compared to models based on their observation data. Our interpretation failed to put this clearly into perspective and we have now created a section discussing this issue. To elaborate more on this issue in that section, we also made some preliminary maps (which need further improvements because of data imbalance issues) of the site specific variables and used these maps only as covariate to assess the quality of their predictions.

Manuscript lines (see lines 585 – 615): see section 4.4 Implication and limitations of the study DSM relies on existing maps for building regression models and ultimately prediction for mapping. The quality and accuracy of predictions depend as discussed earlier on choosing the most relevant covariates in relation to the target to be predicted. The present study revealed that covariates which were available as maps did contribute to the MDA, but site characteristics were more prominent in relation to the SOC stock in Sweden. This might suggest that mapping these variables that were more decisive for SOC stock prediction and including them as covariates for mapping may improve accuracy. Since the primary focus of the present study was mainly to evaluate the GoC data versus the SSC data at different scale of modelling and not primarily for making a map, the observed data of the latter were used. However, since only the observed data of the site characteristics were considered, this study fails to consider that the mapping of these site characteristics would involve modelling errors and the propagation of these errors into the final maps of these site variables might actually reduce their predictive power. For completeness, we carried out a preliminary mapping of the site characteristics (SI-11) using additional soil inventory data, random forest and the GoC as predictors. These mapped site characteristics (mSSC) were then used as covariate for predicting the SOC stocks.

Table 6 presents the error metrics after independent validation for both local and global models along with the percentage margin of the RMSE in relation of the models based on GoC. First, the local mSSC based models still recorded lower RMSE as compared to global mSSC models. Compared to the GoC models, the overall positive percentage margin of the RMSE for the independent validations indicated that the mSSC models recorded the lowest RMSE. However, when assessing the RMSE margin between the global models of GoC and SSC, negative percentages were mainly recorded for the Northern Sweden independently of the depth. This indicated that the mSSC based global models were less accurate at predicting SOC stock locally in the Northern Sweden while the mSSC based local model did present a better prediction for Northern Sweden for the humus layer and mineral SOC stock. The mean SOC predictions based on the mSSC showed a stronger increasing gradient from Northern to Southern Sweden (SI-12) as compared to the pattern observed with the GoC maps. However, the uncertainty distribution was of similar magnitude (SI-12) as those observed for the maps based on GoC probably due to error propagations as these covariates were used to make the site specific characteristics maps. This suggests that the SSC should still be supplemented for improvement at this stage with other covariates different from the GoC such as multi-temporal spectral (e.g. normalized difference vegetation index) data that are able to capture vegetation dynamic in forests. Notwithstanding the fact that was obviously error propagation, the study of which was beyond the scope of this study, the preceding results tend to confirm the potential of the high resolution maps of the site characteristics to contribute to the improvement of SOC stock prediction as compared to using only the GoC data. Given that the preliminary mappings of the SSC recorded low kappa (0.17 – 0.48) values (SI-11) at this stage further improvements are still necessary to improve their predictive ability and associated coefficient of variations.

Specific comments are listed below: Line 49: which are nonspatial environmental covariates? If they are not spatial, how they can be used in DSM? - Author′s Response: Sentence corrected with removal of the word non-spatial Manuscript lines (see lines 47 – 50): DSM aims at estimating the spatial distribution of soil classes or soil properties by coupling field and laboratory observations with spatial environmental covariates via quantitative relationships. Line 58: Authors miss at least two papers on comparing local and global models in DSM. Piikki, K., & Söderström, M. (2019). Digital soil mapping of arable land in Sweden–Validation of performance at multiple scales. Geoderma, 352, 342-350. Song, X. D., Wu, H. Y., Ju, B., Liu, F., Yang, F., Li, D. C., ... & Zhang, G. L. (2020). Pedoclimatic zone-based three-dimensional soil organic carbon mapping in China. Geoderma, 363, 114145. - Author′s Response: Thanks for mentioning these papers. The suggested papers have been included. Manuscript lines (see lines 58 – 60): However, models could be calibrated separately for subareas and their predictions can then be combined to cover the whole area (Somarathna et al., 2016; Piikki and Söderström, 2019; Song et al., 2020). Line 66: SCORPAN also includes soil information compared to Jenny's soil forming theory. - Author′s Response: We agree with the comment and included soil information in addition to location position for additional elements from SCORPAN compared to Jenny′s soil forming theory. Manuscript lines (see lines 65 – 67): Building on the soil state-factor (climate, organisms, relief, parent material, age) equation developed by Jenny (1941), McBratney et al. (2003) introduced the conceptual framework for DSM referred to as SCORPAN which complemented the former with the inclusion of soil information and location coordinates. Lines 89-90: It seems to me that this manuscript does not take the SOC changes during three periods into account. I would add relevant statements with supporting references. - Author′s Response: The reviewer is right that the manuscript does not focus on SOC changes over the three inventory periods. We rather focus on SOC stock by considering the averaged SOC at plot level as a good representative of the plots which we indicated in lines 107 – 110. We therefore explained clearly our focus and added further references for the dataset. Manuscript lines (see lines 89 – 93): The NFSI runs concurrently every year with the NFI and consist in repeated survey of forest vegetation and soil chemical and physical properties (Stendahl et al., 2017;Ortiz et al., 2013). Data from the following inventory periods were considered in the present study: 1993 – 2002, 2003 – 2012 and 2013 – 2015. However, the present paper did not focus on SOC changes over these three inventory periods but on SOC stock using plot scale as a unit. Manuscript lines (see lines 110 – 112): Since potential SOC stock change is very small compared to the entire SOC stock the averaged SOC stock between the inventories was considered representative of the plots and was therefore considered for all computations and modelling in order to reduce variability between plots

Lines 103-104: Since mineral soil is sampled at 0-10, 10-20 and 55-65 cm depth, which kind of interpolation is used to harmonize them at 0-50 cm? Please provide more details. - Author′s Response: The carbon stocks for each layer with data is calculated. Thereafter the carbon pool is assumed to change linearly between measured layers. Even if the deepest sampled horizon (55-60 cm) is deeper than the top 50 cm it is utilized in the interpolation but the carbon stock below 50 cm is not counted in. Manuscript lines (see lines 107 - 110): The total mineral SOC stock down to 50 cm depth for each site is calculated using the SOC stock of measured layers with empirical model for bulk density (Nilsson and Lundin, 2006), corrections for stoniness (Stendahl et al., 2009) and linear interpolation between measured layers Line 158: RF is commonly used instead of RFR. - Author′s Response: RFR has been replaced by RF throughout the document. Manuscript lines (see lines 65): The Random Forest (RF) algorithm was selected for SOC stock prediction. Line 216: Which method is used in recursive feature elimination in caret package? Random forest? Author′s Response: Random Forest was used as method in recursive feature elimination in caret package. We now added this method in the referred line. Manuscript lines (see lines 220 – 221): . . .and (2) the recursive feature elimination (RFE) using RF as method to select the optimal set of covariates for each RF model. Line 247: What does p mentioned here? - Author′s Response: The p stands for probability value. We corrected and put the p in bracket after probability. - Manuscript lines (see lines 250):

The latter is the graphical representation of the proportion of time the actual values of SOC stock fall within a series of probability (p) of prediction intervals (PI) limited by (1-p)/2 and (1+p)/2 quantiles.

Please also note the supplement to this comment:
https://soil.copernicus.org/preprints/soil-2020-75/soil-2020-75-AC4-supplement.pdf

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2020-75, 2020.