Answer to questions by RC1: Anonymous Referee #1

RC1: Anonymous Referee #1, 16 Aug 2022

Ågren et al. map the spatial distribution of peat soils and organic layer thickness in Sweden using an existing soil moisture map and national-level field inventory data. The manuscript is well written and mostly sound, but I was left partly confused when reading the manuscript. I have the following major points.

Question: Why did you predict thickness of organic layer based on soil moisture map and not from the original predictor variables that were used to produce the soil moisture map? This seems to be quite odd as there is now double uncertainty in the estimates, as the prediction of soil moisture was already a little uncertain. You should justify your approach better. It would be also interesting to compare different predictor variable sets (e.g. topography variables, satellite imagery etc.) to produce the thickness of organic layer and not just use one existing map.

Response*: There is an urgent need for an improved peat map among the Swedish land managers and practitioners. After the SLU soil moisture map was recently produced at a remarkably high resolution, there has been wide interest how we can understand the relationship between the soil moisture variation and distribution of peat across the country using the SLU soil moisture map. This present study was an immediate response to this applied research need. We understand your concerns regarding uncertainty, however, our comparison with "traditional maps" suggest that the peat delineation based on our approach was satisfactory (Table 3 in our article).*

Question: Based on Figure 4, the fit of the model between soil moisture and organic layer thickness is not very good. This should be discussed in more detail.

Response: Unfortunately, the fit of the model in figure 4 looks worse than it actually is due to the overlap of the points in the figure. We have therefore revised the figure to include the quadrants of the confusion matrix, using ≥50 cm peat as an example. If we take the example of delineating peat with ≥50 cm depth, the model misclassifies peat as mineral soil in 104 instances (FN) and misclassifies mineral soils as peat in 135 cases (FP) But, 427 observations are correctly classified as peat and 2217 are correctly classified as mineral soils. 2644 of 2883 soil pits are correctly classified in this example. This gives the ≥50 cm peat map a higher overall quality (Kappa and MCC) than the topographic and Quaternary deposits maps. We will include this new graph in the revised manuscript and explain this in the discussion.

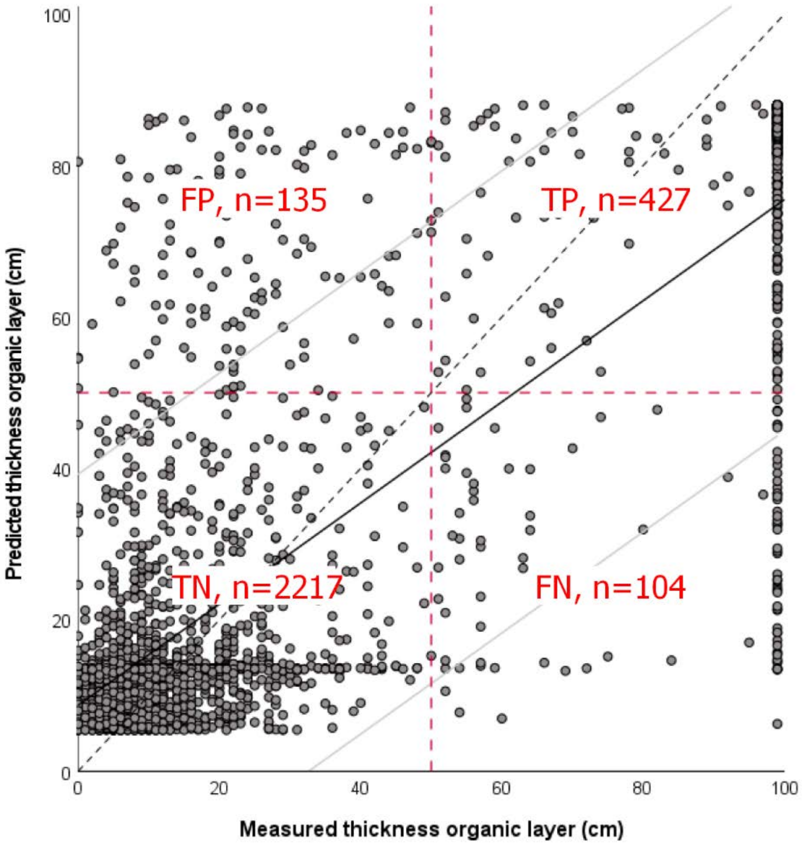
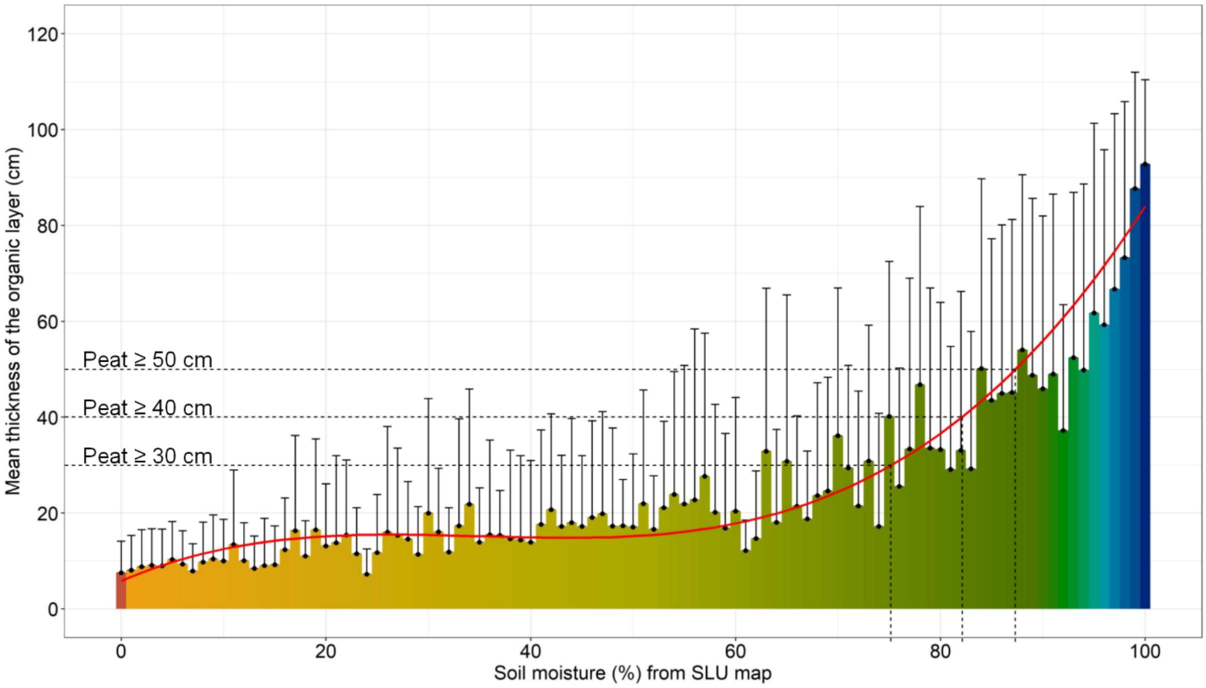


Figure 4 from the article with the quadrants for the confusion matrix highlighted in red, here using ≥50 cm peat as an example. Predicted vs measured continuous thickness of the organic layer in the evaluation dataset (n = 2 883). Dashed line indicate the 1:1 line and black line a linear regression (R^2 = 0.67, p<0.001), grey lines indicate 95% CI. Field measurements of peat thickness that were 99 cm or above were reported as 99 cm, hence the many over lapping data points.

Question: Actually, there has been a lot of discussion that R2 should not be used for nonlinear regressions. You should justify why you evaluated model performance with R2 and you should also report how you calculated R2.

Response: *In our selection of model (Figure 3) we first disregarded the non-significant models (based on P-values). Then we used R2 to rank them because that provided the best accuracy metrics for the curve estimation procedure in SPSS we used. In regression analysis, the more parameters you add the more you increase the R2, however, you risk over-parametization of the model. Normally, we would never use a cubic model for predictions as it often overfits the model and introduce bias. However, in this case it was the only function that could capture the variability of the data with a sharp increase at the low end of soil moisture, followed by a plateau and then a sharp increase on the high end. However, it could not fully capture this pattern, hence field measurements ranged 0-99 cm while predictions ranged 6-88 cm. If our focus had been on predicting the thickness of the organic layer then we would have investigated a more complex non-linear model that might have captured this better. However, in this study the thickness of the organic layer is just a means to an end, as we focus on the horizontal delineation of peat. The red curve in the graph fits the data "well" in the range* *we focus on (ca 70-90% soil moisture), in the green color-range.*



The sharp increase on the low end is explained by the fact that the driest sites are generally found on crests and ridges characterized by rock outcrops with very thin organic layers. The plateau in the middle represents typical forest soils (mostly podsoils) and the sharp increase at the high end is explained by the formation of peat in the wettest areas. So in this case we believe it is an appropriate use of the cubic function.

Question: Precision was lower in your map than in the topographic map. It seems that your mapping approach seems to overpredict peatlands, at least to some extent. This seems to be the case also when visually interpreting the material in Figs. 5 and A1 and looking at the information in Table 3.

This should be accounted for and discussed in more detail. You could discuss e.g. why your approach seems to overpredict the extent of peatlands and overestimate thickness of thin organic layers. The overestimation of thickness of thin organic layers is probably due to the selected cubic model. You could potentially also use other (non-linear) regression models and discuss the pros and cons of different models.

	TP	**TN**	**FP**	**FN**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**Specificity (%)**	**Kappa**	**MCC**
Peat ≥50 cm map	*427*	*2217*	*135*	*104*	*91.7*	*75.9*	*80.4*	*94.3*	*0.73*	*0.73*
Peat ≥40 cm map	*417*	*2133*	*165*	*114*	*90.1*	*71.7*	*78.5*	*92.8*	*0.69*	*0.69*
Peat ≥30 cm map	*537*	*2013*	*199*	*134*	*88.5*	*72.9*	*80.0*	*91.1*	*0.69*	*0.69*
Topographic map	264	2318	39	266	89.4	87.1	49.8	98.4	0.58	0.61

Quaternary deposits map	363	2227	130	167	89.7	73.6	68.5	94.5	0.65	0.65

Response: *Mostly it is the topographic map that under-predicts peatlands. For example, the topographic map only over-predicts 39 instances but under-predicts 266 instances, a clear bias toward under-prediction (with an overweight of 227 misclassified soil pits). So the high precision for the topographic map is driven by the low number of FP, however, note that the recall for the topographic map was below 50%. Our ≥50 cm peat under-predicts peat in 104 instances and over-predicts peat in 135 instances, a fairly balanced distribution of errors with only a slight overweight (n=31) towards over-predictions. But you are correct, that this problem is bigger for the peat ≥30 cm map, with an overweight of 65 plots towards overprediction. This illustrates how easy it is to misinterpret results when using measures that only focus on parts of the confusion matrix. The best measure of the overall performance of the map quality (taking into account both over- and under-predictions) is the MCC which shows that our predicted maps outperform the topographical and Quaternary deposits maps and that the ≥50 cm peat has the highest quality. We suggest to clarify this further in the discussion in the revised manuscript and we will also add text on the slight over-prediction of peat on the ≥30 cm maps.*

As for the visual interpretation of the map, we were also guilty of this interpretation at first because we are so used to believe that the traditional maps show the truth. However, the evaluation of the map in Table 3 shows that our predicted maps locate more peatlands (higher recall rate), without over-predictions of peatlands (for peat ≥50 cm). It took us some time to “recalibrate” our brain and not trust the classical topographical maps “blindly” when interpreting the maps. And that we are considering as a success of our proposed methodology for peatland delineation.

Additionally, I have the following more detailed but mostly minor comments:

Abstract:

Comment: Please remove the first sentence. It is not necessary.

Response*: The first sentence will be deleted.*

Comment: l17-19: it is not needed to report the results from an existing study. Please rephrase and shorten the sentence

Response: *“with high accuracy (Kappa = 0.69, MCC = 0.68)” will be removed from the sentence.*

Comment: l25: please report also the precision results

Response: *We choose not to focus on the precision, but we will motivate better in the discussion why.*

Comment: l28: "peatlands visible from airplanes" could be written "peatlands that can be visually detected from aerial imagery"

Response: *We will follow this suggestion*

Comment: l29: delete "most importantly"

Response*: "most importantly" will be deleted.*

Introduction:

Comment: l80-87: the direct quote is unnecessarily long. Do you need to include it?

Response: *We believe it's important to be clear about this definitions and would like to keep it as it reads.*

Comment: l113: is the fourth objective necessary to include?

Response*: Yes, we believe so. According to "Digital mapping of peatlands - A critical review" by (Minasny et al., 2019) the review showed that while it is common to delineate peat extent, studies rarely perform validation or calculations of the uncertainty of the predictions. We believe it is a good practice in research to provide uncertainty estimates of model results.*

Comment: l114: write "study provides a guide to map…"

Response*: We followed this suggestion*

Methods:

Comment: l142-155: this paragraph could be shortened as it describes results from an earlier study, not the methods of this study.

Response*: While you asked us to shorten this section, the editor asked us to expand this section. We decided to add some more background on the soil moisture map which probably will aid the reader to better understand this article. Please see our earlier response to the editor.*

Comment: Did you account for spatial autocorrelation when e.g., constructing the model and dividing the calibration and validation datasets?

Response*: No, as the closest distance between soil pits in the datasets are 500 m and autocorrelation in peat depth is usually detected on a smaller scale, we could not account for the autocorrelation. Hence the data was just split randomly into two datasets. Although this may be interesting to investigate further in the future, we argue that autocorrelation may not produce better results because of the sample design of the field dataset.*

Comment: l195: This is difficult to understand. Does it mean that 1:25 000 map covers 1.7% of Sweden and so on?

Response: *Yes, we now propose to remove the brackets in hope that this makes it clearer.*

"However, there are various scales with different coverages for the quaternary deposit maps in Sweden, such as 1:25 000 covers 1.7% of the area, 1:50 000 covers 2.7%, 1:100 000 covers 47%, 1:200 000 covers 1.4%, 1:250 000 covers 21.2%, 1:750 000 covers 33.6% and 1:1 000 000 covers 100%."

Comment: Why did you include the used accuracy metrics? Kappa has been heavily criticized (see e.g., https://doi.org/10.1016/j.rse.2019.111630). You could also have included F-score.

***Response**: Yes, Kappa has been heavily criticized, but since it has been used extensively previously and is still often used in map-quality studies we present it here for easy comparison with other studies. However, due to the criticism towards Kappa we also present MCC which is considered a better measure (Chicco et al., 2021).*

We choose MCC instead of F-score as MCC has been shown to have advantages over the F-score (Chicco and Jurman, 2020). However, as we publish the raw confusion matrix with this manuscript, it is possible for the interested readers to get an estimate F-score (along with other metrics of choice) if deemed necessary. We believe it should be a standard to publish the raw confusion data (TP, TN, FP, FN) as that will enable future meta-studies to calculate all possible evaluation metrics needed. This is unfortunately quite often neglected in the literature today.

We will also add a sentence in the discussion to explain this further; "Out of the two measurements kappa and MCC, MCC is considered the most informative measure (Chicco et al., 2021)."

Comment: Section 2.6: How were the field inventory datasets upscaled? Does this simply mean that you calculated national level statistics from the datasets using different methods?

***Response**: Yes, it is simple statistical upscaling performed by the experts in charge of the national surveys. They are expert statisticians and have a deep understanding of the sampling design and how to best calculate national estimates based on the survey data. In the revised manuscript, we will now explain that we use statistical upscaling and that is was performed by the experts.*

*"To compare the predicted peat soil estimates from the maps with other estimates of peatland coverage in Sweden, we also calculated peatland coverage by **statistical** upscaling from the national inventories, i.e. NFI and SFSI (SLU, 2021) to derive a complete coverage for the Swedish forest landscape. From these inventories, peat coverage was estimated by the **statistical experts** at NFI **(Fridman et al., 2014)** and SFSI **(Stendahl et al., 2017)** in 6 different ways".*

We reference Nilsson et al., 2018 and Hånell, 2009 where the upslacling calulations are described more-in depth. We now also refer to the articles by (Fridman et al., 2014), and (Stendahl et al., 2017) that describes the NFI and SFSI in more detail.

Comment: the heading of 3.4 could be changed. Should it be "visual interpretation of peatland maps"?

***Response:** Good suggestion, we will change it accordingly.*

Discussion and conclusions

Comment: l375-376: This is misleading as you used ALS data very indirectly.

Response*: We kindly disagree here, the whole foundation of the soil moisture map is ALS data. Out of the 28 input data, the most important variables for predicting soil moisture were the digital terrain indices derived from ALS-data, not the traditional maps or the runoff. We believe this to be clearer now in the revised manuscript that we have expanded the methods section on the SLU soil moisture map.*

Comment: l460: you write multiple times that the map should not be taken literally. It is not necessary to mention this multiple times.

Response*: We will address this throughout the manuscript.*

Comment: The section "The novelty of the developed maps" could be shortened and merged with conclusion section. Some text can also be moved to other parts of discussion.

Response*: We will consider this in our revised manuscript.*

Comment: l509: delete "coarse", "global mapping" is sufficient.

Response*: We will follow this suggestion*

Comment: l504-510: Sentinel-2 has 10 m resolution and it surely can be used for quite detailed planning. There is also other remote sensing than just ALS data that can be used in detailed planning.

Response: *Yes, sentinel-2 data can be used to detect open peatlands, however, it shares the same limitation as aerial photos that the canopy cover makes it difficult to detect the soils in large areas of the Swedish forest landscape. A benefit of our method of mapping peat using mostly ALS data is that the ALS "sees through the canopy cover". For more detailed planning, for example designing a riparian zone (which in Sweden is on average 4 m wide), or finding suitable crossings across streams, high resolution aerial photos (often 0,5 m resolution) in combination with ALS data is usually preferred by practitioners. Sentinel-2 data may be too "pixelated" for such planning. When we developed the soil moisture map (Ågren et al., 2021), we evaluated the CORINE land use data which is based on Sentinel-2. However it was excluded due to low contribution to the models. See also our answer to editor's questions.*

References

Chicco, D. and Jurman, G.: The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, Bmc Genomics, 21, 2020.

Chicco, D., Warrens, M. J., and Jurman, G.: The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment, Ieee Access, 9, 78368-78381, 10.1109/Access.2021.3084050, 2021.

Fridman, J., Holm, S., Nilsson, M., Nilsson, P., Ringvall, A. H., and Stahl, G.: Adapting National Forest Inventories to changing requirements - the case of the Swedish National Forest Inventory at the turn of the 20th century, Silva Fenn, 48, 2014.

Minasny, B., Berglund, O., Connolly, J., Hedley, C., de Vries, F., Gimona, A., Kempen, B., Kidd, D., Lilja, H., Malone, B., McBratney, A., Roudier, P., O'Rourke, S., Rudiyanto, Padarian, J., Poggio, L., ten Caten, A., Thompson, D., Tuve, C., and Widyatmanti, W.: Digital mapping of peatlands - A critical review, Earth-Sci Rev, 196, 2019.

Stendahl, J., Berg, B., and Lindahl, B. D.: Manganese availability is negatively associated with carbon storage in northern coniferous forest humus layers, Sci Rep-Uk, 7, 2017.
Ågren, A. M., Larson, J., Paul, S. S., Laudon, H., and Lidberg, W.: Use of multiple LIDAR-derived digital terrain indices and machine learning for high-resolution national-scale soil moisture mapping of the Swedish forest landscape, Geoderma, 404, 115280, https://doi.org/10.1016/j.geoderma.2021.115280, 2021.

Citation: https://doi.org/10.5194/egusphere-2022-79-RC1