Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2017-323-AC1, 2017

Biogeosciences Discussions
Open Access
EGU

Interactive comment on "High-resolution digital mapping of soil organic carbon in permafrost terrain using machine-learning: A case study in a sub-Arctic peatland environment" *by* Matthias B. Siewert

MB Siewert

matthias.siewert@umu.se

Received and published: 8 November 2017

Anonymous Referee #1

Received and published: 21 September 2017

This paper discusses a study that developed a high spatial resolution map of soil organic carbon for a sub-Arctic peatland in northern Sweden, using essentially Random Forest algorithms and a suite of environmental variables, including land cover,remotely-sensed vegetation indices, and digital elevation terrain modeling (DEM). The study

Printer-friendly version

Discussion paper

is relatively straightforward, and demonstrates a reasonable approach for modeling/mapping soil carbon in high northern latitude systems. My only major issue had to do with clarification of the resolutions of the various input datasets, and the ultimate resolution provided by the model/map.

That and other minor points are listed here:

1)So, with regard to the resolution of the inputs and outputs, I found it slightly hard to follow, and I think it might help to put all of the resolutions on Figure 1 (right now only the orthophoto/DEM and the final map resolutions are on there). If I am understanding this correctly, the orthophoto is 1m and the DEM is 2m (this is actually slightly misleading in Figure 1, which has the orthophoto + DEM as 1m – but, I guess that the DEM was just "down-sampled" to 1m resolution. The SPOT data are either 10m or 20m, and the minimum size of a land cover classification was 130m 2 , so somewhat consistent with a SPOT pixel, although it's unclear what the range of extents are for land cover regions.

The final map is then generated at the 2m resolution; why not 1m and utilize the more resolute orthophoto information?

Thank your for this very interesting comment that opens up a different perspective to this work. The resolution of the individual products is now mentioned in Figure 1 (Now Figure 3). A resolution of 2 m for the final model was originally chosen as a compromise between the available input variables, output quality, the benefit of higher resolution and processing time.

However, as this point has been mentioned by several reviewers, I suggest to run the model at several spatial resolutions: 1m, 2m, 10m, 30m, 100m and 1000 m for the Total SOC and at 1m for different depth intervals. This would be in line with the suggestions made by reviewer 3 and 4. At this point, I have tested

Printer-friendly version

Discussion paper

these resolutions using the random forest model and got meaningful results. The outcome will be discussed following the reviewers input with regard to the resolution of different input datasets.

2) Figure 6 – Does "Mean decrease in accuracy" indicate the accuracy reduction when that variable is removed from the analysis? If so, make that clear in the figure and caption.

Yes, that is the meaning of this measure. The figure caption has been reformulated to emphasis this:

Fig. 6. Variable importance for the prediction of total SOC measured as mean decrease in accuracy of the random forest model if the variable is excluded. The higher the value the more important is the variable.

3) Also, it's interesting that the most important variable in the analysis was Land Cover, the variable at the coarsest resolution, followed by three SPOT variables. In fact, you don't get a DEM variable until the 5th-most important (Elevation), and even then it's unclear that the information is necessary at the 2m resolution (could be equally useful if aggregated to a coarser resolution). I know from first-hand experience that these systems can be highly variable in space over short distances with regard to SOC; however, it's certainly interesting that most of the variability explained occurs at resolutions of tens of meters, which puts into question the utility of a 2m resolution map. I think this is worthy of some additional discussion in the paper – particularly within the context of what is discussed on Page 13, Lines 1-8,where a fine resolution is necessary to capture the appropriate scale of variability in SOC.

I agree with the reviewer in this point. Indeed, lower resolution input variables seem more important than higher resolution input variables. However, I believe that this is much an effect of the validation rather than true value in the spatial

prediction using the model. Looking at the resulting maps of SOC in Figure 6, it is clearly visible that information from the DEM has a strong influence on the final map. This seems to be more relevant for fine scale and linear landscape features, while larger homogeneous areas are more influenced by lower resolution input data. The discussion will be updated regarding this comment.

4) Abstract, Line 10 – add “for SOC quantification” after “evaluated”

Changed

5) Abstract, Line 16 – change “surprising” to “surprisingly”

Changed

6) Abstract, Line 19 – add “s” to “scale”

Changed

7) Page 2, Line 2 – specify “Northern” high latitudes

Changed

8) Page 2, Line 8 – to what depth is the ∼1300 Pg SOC estimate?

It is now specified that this includes “soils to a depth of 3 meters and other unconsolidated deposits “. The reader can get more information on this under the specified reference.

9) Page 2, Line 10 and throughout – be consistent, either hyphenate “permafrost affected” or not – probably should hyphenate

Hyphenate is now used throughout

10) Page 2, Line 24 – remove “a” before “commonly”

Changed

11) Page 3, Line 1 – remove the hyphen from “higher-latitudes”

Changed

12) Page 3, Line 15 – I think LCC has not been spelled out yet in the paper

LCC has been spelled out on page 2.

13) Page 4, Line 19 – How long were the transects (i.e. what was the distance between sampling points)?

The following information was added: “between 50 to 300 m (Fig. 2)”. This should enable the reader to understand the sampling layout.

14) Page 4, Line 29 – “deeper soil horizons were sampled in 5-10 cm intervals” – what actually were the intervals, and what determined them?

To be more specific it was added “depending on horizon thickness”

15) Page 5, Line 4 – change “were” to “where”

Changed

16) Page 5, Line 6 – should the notation be “>2 mm,” if you are referring to the coarse fraction, or are you referring to the soil that is not the coarse fraction?

Changed to > 2mm.

17) Page 5, Line 13 – add “SOC” before “stored”

Changed

18) Page 8, Lines 8-10 – I understand the overestimation of SOC values due to the absence of sample point from bare ground surfaces, however, I just want to clarify the justification for using 0 as the quantity of SOC. First, I’m not sure I know what a “blockfield” is – maybe that’s just me, but I think a definition/description would be good. Also, one cause of bare ground in northern high latitudes is cryogenic disturbances (i.e. cryoturbation), and in many cases, these were once vegetated areas that can have quite a bit of SOC. Are these generally uncommon in your study area? In other words, are the dominant bare ground features these blockfields and stone beaches that I imagine have very little SOC?

The following has been added to clarify this:

“Originally, all models overestimated SOC contents for bare ground surfaces. These areas include exposed bedrock, blockfields (extensive areas covered by shattered rock fragments with little or no fine substrate; Fig. 3 b) and stone beaches along lake shores. Alpine heat tundra with minimal soil development and cryogenic features form a separate class. “

19) Page 9, Line 4 – add “ed” to “collect”

Changed

20) Page 9, Line 10 – remove one “s” from “miss-“

Changed

21) Page 11, Line 3 – add “be” after the first “to” and remove the 2nd “to”

Changed

22) Page 11, Line 5 and throughout – Sphagnum should be capitalized and Italicized

Changed

23) Page 12, Line 26 – don't capitalize "Geographically"

Changed

24) Page, 12, Lines 28-29 – I'm not sure that I understand the statement that "very strong environmental gradients" would "suggest low spatial autocorrelation." I would think that strong environmental gradients would lead to high spatial autocorrelation.

Rephrased to "In Abisko, sharp transitions in SOC storage between different land covers are found and suggest low spatial autocorrelation at local scale."

25) Page 18, Line 2 – change "adoptions" to "adaptations" – I think that's what you are meaning to say?

Changed

Thanks, that's clearly what I meant.

26) Page 18, Line 3 – need to reword "release them into the carbon cycle" – even if a carbon pool is stable for a long period of time, it's still in the carbon cycle.

The sentence was reworded to avoid this construct:

Rapid future permafrost degradation in peatlands may lead to erosion of organic sediments. This would transfer presently stored carbon into lakes and potentially into the atmosphere.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-323, 2017.

Printer-friendly version

Discussion paper