Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2017-323-RC1, 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

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 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)

Printer-friendly version

Discussion paper

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? 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. 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. 4) Abstract, Line 10 – add "for SOC quantification" after "evaluated" 5) Abstract, Line 16 – change "surprising" to "surprisingly" 6) Abstract, Line 19 – add "s" to "scale" 7) Page 2, Line 2 – specify "Northern" high latitudes 8) Page 2, Line 8 – to what depth is the $\sim$1300 Pg SOC estimate? 9) Page 2, Line 10 and throughout – be consistent, either hyphenate "permafrost affected" or not – probably should hyphenate 10) Page 2, Line 24 – remove "a" before "commonly" 11) Page 3, Line 1 – remove the hyphen from "higher-latitudes" 12) Page 3, Line 15 – I think LCC has not been spelled out yet

in the paper 13) Page 4, Line 19 – How long were the transects (i.e. what was the distance between sampling points)? 14) Page 4, Line 29 – "deeper soil horizons were sampled in 5-10 cm intervals" – what actually were the intervals, and what determined them? 15) Page 5, Line 4 – change "were" to "where" 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? 17) Page 5, Line 13 – add "SOC" before "stored" 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? 19) Page 9, Line 4 – add "ed" to "collect" 20) Page 9, Line 10 – remove one "s" from "miss-" 21) Page 11, Line 3 – add "be" after the first "to" and remove the 2nd "to" 22) Page 11, Line 5 and throughout – Sphagnum should be capitalized and Italicized 23) Page 12, Line 26 – don't capitalize "Geographically" 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. 25) Page 18, Line 2 – change "adoptions" to "adaptations" – I think that's what you are meaning to say? 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.

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