Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2018-15-RC2, 2018

Interactive comment on "A model based on Rock-Eval thermal analysis to quantify the size of the centennially persistent organic carbon pool in temperate soils" *by* Lauric Cécillon et al.

Anonymous Referee #2

Received and published: 27 February 2018

The manuscript presents an important and innovative contribution to the quantitative determination of the SOC pool with centennial turnover times (CPSOC). Quantitative understanding of this pool is highly important, e.g. for model initialization. The paper is well written and the statistical methods used are state of the art and without major flaws, as far as I can judge. However, there is one major point, in which the authors have not yet entirely convinced me that the derived model is capable to predict CPSOC in "new" samples:

1. Machine learning is used to find the best regression model predicting the proportion of CPSOC. A high R2 (0.91) in the calibration dataset is impressive and shows that

Printer-friendly version

Discussion paper

the thermal stability (RE6) can be linked to biogeochemical stability, which has been shown before. Now, the interesting thing is the validation: the authors report the same R2 for the validation set and show that they scatter as nicely along the 1:1 line as the calibration dataset does: Of course, this is the case because the dataset was randomly split, although the samples were not independent but originated from only 4 experiments. So the question is, will the results we similarly good, if for example 3 sites are used to train the model and 1 site is used for validation? This would give a much more honest picture on the validity of the approach. I guess that the prediction would not be as good: According to the Barre et al 2016 paper, at least the three presented thermal stability parameters HI, OI and T50CO2_ox which played an intermediate to important role also in the present study, varied considerably across sites. Also Figure 3A indicates that the 'thermal signature' of the samples is really site dependent. So for me the question is: Can this product really hold what the authors are promising, e.g. in the last sentence of the abstract: 'This model can thus be used to predict. . .'? This has to be clarified and if not the case be discussed with much more caveats. Uncertainties are already huge and they would probably inflate if new samples shall be predicted.

I also have some (partly minor) specific comments:

P5,line 8: Why such a huge intercept in the regression (0.4)?

P6, line 23: Fixed standard deviation for SOC concentration data? Why, and what is it exactly derived from?

P12, line 28: I thought it was 30 RE6 parameters, here it says it was 25?

Discussion:

1. Is very technical and focused on the specific RE6 method and related papers. Authors miss the chance to broaden the perspective and discuss this approach to estimate CPSOC in comparison to other approaches or to establish a clear link to kinetic models.

2. Is very positive about the overall results (and sample set), and although uncertainty was a clearly stated focus of the study it is not really taken up here: Yes, the sample set is truly unique, but this is also the problem: How uncertain will the CPSOC estimation in soils be, which do not have bare fallow treatments?

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2018-15, 2018.