SOIL Discuss.,
https://doi.org/10.5194/soil-2017-40-RC1, 2018

Interactive comment on "No Silver Bullet for Digital Soil Mapping: Country-specific Soil Organic Carbon Estimates across Latin America" *by* Mario Guevara et al.

T. Hengl (Referee)

tom.hengl@isric.org

Received and published: 7 March 2018

The paper compares five statistical/machine learning techniques for the purpose of mapping organic carbon stock 0-30 cm for 19 countries in Latin and Central America: SVM, RF, PL, KK and RK. The authors conclude, based on their results, that (P14L13) "there are no silver bullets on digital soil mapping" meaning that no single technique outperforms other in sense of accuracy. Also, authors further conclude that countries need to work on improving the quality and quantity of ground data on soil carbon. Methodology is relatively well explained and I especially appreciate that most of the code is also available for review (https://github.com/DSM-

LAC/NoSilverBulletsForDSM), which is really the best way to publish scientific work. Although a valuable piece of work that brings >20 soil data producers from Latin America, I have some questions concerning the methodology and results (some sections lack clarity and especially methods section needs to be extended), and also some reservations considering the main messages of the article. I hope that the authors (and the readers) will recognize some of these points and that these will help them improve their paper:

1. Weighted overall measure of mapping accuracy maybe more informative than stats per country? The main result of the paper is that neither of the five methods considered (SVM, RF, PL, KK and RK) results in significantly more accurate results. I have however two concerns about the P10L18-20 and Figure 2: (a) since two countries (BRA and MEX; Table 1) have over ca 5 times more points than all other countries together, I think it would be more fair to compare overall accuracy of methods using a weighted overall measure of mapping accuracy (where the weight could be either size of country, or better number of points per country), (b) another weighting factor that could be used is the accuracy of estimated soil carbon stocks (t/ha) since different points come with different accuracy (success of fitting splines to irregular soil horizon data can be estimated so that this information can be provided per point). I believe that providing a weighted overall measure of mapping accuracy for the five methods would give a more objective view of which method is more accurate. At the moment I see that (P10L18) for BRA the best method is RF and for MEX RK, hence these would be clear winners considering that these are based on ca 10,000 points (>85% of all points).

2. Building models with <30 training points can lead to artifacts Table 1 indicates that, from 19 countries, only 7 countries have >100 points available for modeling. Modeling and comparing models for countries with <30 points I would not even recommend as fitting of variograms for data sets with <50 is rather tricky and can lead to artifacts. Oliver and Webster (2014; https://doi.org/10.1016/j.catena.2013.09.006) suggest that one should collect at least 100-150 samples to get a reliable estimates of variogram

parameters. I would not be that strict but at least, any results you get for countries with <30 points should probably be critically evaluated.

4. Is local better than global? Is it justified to stitch maps produced by countries vs using global models? Authors provide comparison of state-of-the-art methods for generating spatial predictions using soil carbon stocks and large stack of environmental covariates. However, I firmly believe that most of the readers would in fact be more interested in finding out whether (a) building N local models per country and then making predictions, or (b) fitting a single global model using all data, is more accurate? So adding a section 3.4b "SOC predictions based on global models" and predicted values to the plot Figure 3, would probably significantly increase value of this paper. Also, it would prove that the FAO GSP's choice to let countries map properties and then stitch maps together, is a better option than to merge all points together and then fit single global models. Read more about this discussion in http://www.pedometrics.org/Pedometron/Pedometron38.pdf "On usability of soil maps (and on global soil data models vs stitching together of individual disparate soil maps)"

5. Evaluation of the accuracy of predictions should ideally be based on e.g. k-fold CV with re-fitting It is not entirely clear from the manuscript how was the model evaluation implemented (section 2.3). I would expect a 5-fold CV with model refitting i.e. "repeatedcv" (https://topepo.github.io/caret/model-training-and-tuning.html#basic-parameter-tuning) - is this the one you used? Note that "repeatedcv" ensures that (a) models are repeatedly fitted and (b) there is enough repetition to get stable results. Please provide half page explaining how exactly is the accuracy / RMSE derived.

6. Github repository missing training data Github repository does not contain all points and grids you have downloaded from ISRIC WoSIS and worldgrids.org. I would at least appreciate if you could put the main regression matrix containing values of the target variable and all covariates. This way we would be able to reproduce your results, as in Nussbaum et al. (2018; https://doi.org/10.5194/soil-4-1-2018).

On the end, I should also mention that I am big supporter on connecting countries and especially researchers and applied specialists in countries around the world to share data on soil (and this paper clearly contributes to this initiative). I am not as big supporter of making issues such as climate change, deforestation, soil erosion, soil carbon and similar, become questions of national sovereignty and/or political debate. Or to quote Neil deGrasse Tyson: "Objective truths are established by evidence. Personal truths by faith. Political truths by incessant repetition." See also: https://www.facebook.com/neildegrassetyson/videos/10155195888806613/ and that is why I have especially high reservations towards letting countries freely choose the "most accurate" method to determine soil carbon stocks.

Other detailed comments in the manuscript and questions are available in the appendix (PDF).

Yours,

T. Hengl https://orcid.org/0000-0002-9921-5129

Please also note the supplement to this comment:
https://www.soil-discuss.net/soil-2017-40/soil-2017-40-RC1-supplement.pdf

---

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2017-40, 2018.