SOIL
Discussions

Open Access

EGU

Interactive comment on "Machine learning and soil sciences: A review aided by machine learning tools" *by* José Padarian et al.

José Padarian et al.

jose.padarian@sydney.edu.au

Received and published: 27 November 2019

Thank you for your feedback. Regarding your comments:

Page 15 lines 3-10. Another way to assess the importance/significance of the covariates in ML is to run the model several times excluding at each time one covariate. This may in some cases be more efficient and robust that just counting the number of times they are used in the model. Some ML tools such as GBM also offer the possibility to visualize the effect of a covariate on a target property, this may be useful too to interpret the results.

We purposely left forward, backward and stepwise selection out of that list since they are covariate selection methods. They compare different models and, of course, that

is useful to understand the modelling exercise as a whole but not to interpret a model in isolation.

In terms of GBM, we know about 2 methods to interpret the model. The first method (permutation) is very similar to the sensitivity analysis reported in Ng et al. (2019) and the second one (relative influence) is estimated based on the average reduction of the error at each split. These methods can be applied to other models as well, not just GBM. We will modify the text to add the relative influence method to the list.

Page 18 Lines 13-15. This consideration about how to open the 'black-box' of CNN or NN is very important. I am disappointed to find it just in the conclusion. I think it merits a section in the main text explaining how this black box has been opened and with corresponding cites and references.

Opening the black-box is a current problem in many disciplines and a research topic by itself. We mention that in the last paragraph of Page 14. In that paragraph, we reference 2 important publication in the area of Machine Learning. In soil sciences, there is no much work on the interpretability of NNs. We reference Ng et al. (2019) as an initial approach to interpret CNNs used in soil spectroscopy. As we mention in the conclusions, we foresee that a large number of studies will focus on this topic.

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2019-57, 2019.