



Comment on soil-2020-100

Anonymous Referee #1

Referee comment on "Estimation of soil properties with mid-infrared soil spectroscopy across yam production landscapes in West Africa" by Philipp Baumann et al., SOIL Discuss., <https://doi.org/10.5194/soil-2020-100-RC1>, 2021

The manuscript evaluated the potential of mid-infrared (MIR) soil spectroscopy in estimating a comprehensive list of soil properties in West Africa. The cross-validation of partial least square regression (PLSR) indicated that 11 soil properties can be accurately estimated ($R^2 > 0.75$) while the predictions of other soil properties were less accurate. The manuscript is overall well written with solid methodology (in terms of model parameter optimization and validation strategy). However, the manuscript does not provide new insights to the community as MIR has been studied a lot in soil spectral prediction across scales and the PLSR model is a commonly used linear model. I suggest the authors redefine the objectives of this study in order to highlight the knowledge gap that this manuscript deals with. In addition, I would like to see more discussions about the pros and cons among lab MIR, lab vis-NIR, and in-situ spectroscopy as I am quite sure that the measurement of MIR mentioned in this study is still time-consuming (preparation of fine ground potassium bromide powder). Please find my detailed suggestions below.

Line 67: The map of soil sampling sites is missing. This is a really helpful message to readers.

Lines 116-118: I have a big question here that is it really necessary to predict total elements by MIR spectra as you used the reference measurements from energy dispersive X-ray fluorescence spectrometry (ED-XRF) which does not require much more cost or time than MIR spectroscopy. So why not directly use the measurements from ED-XRF which have better quality?

Line 144: 128 measurements? I am confused here as there are only 94 milled soil samples.

Line 153: I recommend adding the use of a non-linear model (e.g., Cubist, Random Forest) to better demonstrate the predictive ability of MIR.

Line 172: The definition of S_y is missing in equation 2.

Lines 174-155: Until the end of this manuscript, I did not see any results relevant to the uncertainty analysis mentioned here.

Lines 206: As shown in Figure 1, a large difference in soil properties were observed from four regions. Maybe it would be an interesting part to discuss whether the model build from three regions can be applicable to the remaining region.

Line 297: Please check a highly relevant paper (see below) which indicated a good model performance of MIR spectral models to a list of soil properties at a national scale. In addition, Rossel should be corrected as Visarra Rossel. Please also check the relevant typos (e.g., line 299).

Sanderman, J., Savage, K., & Dangal, S. R. (2020). Midâ□□infrared spectroscopy for prediction of soil health indicators in the United States. *Soil Science Society of America Journal*, 84(1), 251-261.