Response to the reviewers: "Estimation of soil properties with mid-infrared soil spectroscopy across yam production landscapes in West Africa" by Baumann et al.

https://soil.copernicus.org/preprints/soil-2020-100/

1 Letter of Response

We thank the editor and the reviewer for carefully reading and commenting on our manuscript, which helps us to improve the manuscript and to deliver our findings more effectively to the reader. Our responses (author comments = AC) are shown in blue text to each of the reviewer's comments (RC) and editor's comments (EC), who also serves as a second referee, shown in black text. We make suggestions about how and where we plan to implement the appropriate changes based on the valuable feedback received. The section References includes a list of papers that we additionally mention in the response.

Best regards,
Philipp Baumann, on behalf of all authors

2 Synthesis

EC/RC1: The manuscript tackles an important issue: yam production is marred by declining yields and a lack of positive fertilizer responses, the absence of current fertilizer recommendations and generally old and traditional cropping systems and technologies.
AC We thank the associate editor for the comment and for acknowledging that our manuscript addresses a key issue that is relevant to yam production systems: a lack of soil data and a tool to identify factors limiting yam yields across landscapes using the limited data. Below we respond to each of the comments made.

3 General comments

EC/RC2: The title somehow promises to be yam specific, yet the manuscript does not really produce data or results that would be specific to any crop. Although this is not a short fall, I would recommend that the authors reflect upon the need to have yam as a target crop – sure the samples were taken in yam fields and this could be taken as a reason, yet not a strong one.
AC: Thank you for this perspective. We will further address the comment made by more clearly stating this in the introduction so that our intentions and premises of the study are also as upfront as possible there. We collected and analyzed the samples with the intention of developing mid-IR soil spectroscopy as a diagnostic tool for production landscapes where yam is grown as a high-value crop. The premise of the study was clearly on the fields managed by smallholder farmers that they considered suitable for yam growth, relative to land resources and cropping history. We stated this point throughout the manuscript (e.g., page 1 line 1; page 2, line 22/24; "production of yam and other crops").

We acknowledge that results on fertility diagnostics are applicable to other crops in rotation within the selected landscapes. We agree with the reviewer that this would be probably the case for other staple crops that have similar nutrient demands and are grown in the regions. We have already accounted for this point in the discussion and conclusion (e.g., p.17, l.345–346).

Moreover, we will thoroughly go through the manuscript again. We will make sure that sufficient reference is made to applicability of models respective to soil status, spectroscopic predictability of properties, and nutritional demands of crops with comparable nutritional demand to yam grown in the regions. We hope that by adding a short paragraph to the introduction we can convey clearly to the reader that the soil diagnostic implications from using spectroscopy can be similar for other crops.

EC/RC2: The set of chemical variables (elements) measured is rather wide and gives the impression it was done because it was possible. The manuscript does not give any indication of how important these elements are for yam growth and yield, except

for N and K in a by-sentence. Although the focus of the manuscript is clearly on the IR predictions and models, the manuscript would gain quality if nutrient requirements of yam and the problems of yam plant nutrition were considered to a larger extend.
AC: Plant-available macro- and micronutrients are critical for cash crop production. We presume that the status of many of these properties are sensitive to nutrient inputs through fertilization but also inherent soil fertility, whose aspects we will strengthen in the revised manuscript. However, for the yam production regions selected for our study, such external inputs are variable. As a result, yam productivity tends to be also related to some of the total elements and texture that are a longer-term indicators for soil quality, which we also for in those study. We determined available nutrients that have been documented to correlate to crop response, such as Zn-DPTA and resin-P. In addition, the links to organic carbon contents and mineralogical constraints that impair or facilitate their availability are complex. Here, the advantage of the spectroscopy is that it scales to many properties simultaneously with one measurement, so that only fraction of newly collected soils is required to undergo classical wet chemistry. We for example mentioned the ratio of N, P, K (page 2, line 34), the importance of soil organic matter to retain and release nutrients (e.g., page 2, line 26/27 "Thus, maintaining or increasing SOM and available nutrient levels is of utmost importance for sustainable production of yam and other crops in West Africa (Carsky et al., 2010).").

Recommendations on yam given by agricultural extension services and research institutions differ between and among countries. Since we did not conduct any field experiment that accompanied soil sampling done in this study, we only briefly touch upon the nutritional roles and requirements of soil elements and instead focus more on addressing the soil's capacity to retain and release nutrients (clay and organic matter) using mid-IR spectra (e.g., page 16 line 291–294).

We agree with the editor that the message of the manuscript can profit from mentioning and critically discussing some important macro-nutrients and losses that can occur besides the modification of soil quality. We will modify the manuscript accordingly and thank for the critical input. Linking yam response to fertilization management would overload the scope of this study, however this is an important next step after establishing the basis for spectroscopic estimations in the regions. Instead, at this stage we focus on these properties for screening purposes. With our data set and models derived, we laid the foundations of spectroscopic model diagnostics and conclude on which diagnostic measures to use. Applications, possibly documented in scientific articles, can henceforth focus on how to monitor soil health whilst attempting to improve region- and site-specific management and yields of yam.

EC/RC2: I would appreciate it if the authors added a short section on which of the accurately predictable chemical or physical properties and elements would be of importance to monitor soil quality for yam. In addition a critical assessment of the differences between total element and available element data – which one matters more?
AC: Thank you for the comment. As suggested, we plan to add the missing content of the comparison. Please also note that we already have discussed the relevance of properties that had predictability in terms of quantitative diagnostics for crop nutrition and soil health in section 3.2 (page 16, line 275–295). Although they are no indicator of plant-available nutrients, total elements such as P can give an indication about mineralogical status (e.g., natural weathering and fertilization) and variability in major soil components. In the revised version, we will further strengthen this aspect in the introduction and discussion. Specifically, we made correlative associations between the direct and indirect properties, backed up with comparable studies aiming for spectral diagnostics in different soil ecoregions and cropping systems (e.g., page 16/17 lines 303–324).

RC1: 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 ($R2>0.75$) while the predictions of other soil properties were less accurate.
AC: This is an important piece of information to make clear which soil properties could be reliably estimated with the method or not in these geographical settings.

RC1: 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.
AC: We disagree with this comment. Mid-IR based methodology has been studied over the past decades. Nevertheless, we keep on making methodological advances in improving modeling accuracy and the data-model integration with new data representing the areas that were not well bio-physically characterized previously. Our study aims to demonstrate some capabilities

and limitations of mid-IR spectroscopy to diagnose soil properties that account for landscape-scale variation in soil fertility. Please note that there is still no such soil library representing the studied regions. The spectroscopic assessment was specifically framed within the gradient of soil-ecological landscapes used among other for yam production. We then have clarified the links between soil measures as well as the spectral regions and molecular fingerprints of soil constituents that are involved in model-based predictions of properties. As we state in the conclusion, any follow-up studies can benefit from the data and derived models of this library.

We would like to remind the reviewer that our study is a contextual study. Since our study is a method comparison paper, we are not convinced that the discussion in the direction that the reviewer recommends would be on-topic. In fact, pros and cons of different spectral technologies and their trade-offs are well established in fundamental spectroscopy and soil spectroscopy literature. For brevity, therefore, we refer to the following publications that can provide more meaningful details on this evaluation (see section References).

RC1: 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.
AC: Mid-IR spectroscopy is an established tool that makes the measurement of soil properties less complicated (consumables, time spent, costs). In a single day, one person using a single instrument can easily manage to measure 100 to 200 samples depending on the spectrometer configuration and replicates (high-throughput module). The advantage of the used ALPHA spectrometer is that it is small and relatively affordable for poorer countries. Spectra allow to extract multiple properties and to perform iterative validation with only a fraction of the new soils. Milling is still necessary, which has do be done for many conventional analyses, too. Certainly, 5–15 ,% or even less of the samples to be characterized are needed for model adaption and validation purposes. We refer to existing original articles and reviews about costs, efficacy and opportunities of different soil spectroscopic modeling approaches and measurement technologies. See e.g.: chapter "3.3 Cost/Benefic Analysis of Soil Spectroscopy" in Nocita et al., 2015.

4 Specific comments

4.1 Abstract

4.2 Introduction

EC/RC2: Page 2 line 20: population growth does not cause soil degradation it is the increased land use frequency and intensity that causes soil degradation. Combine with previous sentence to create a correct statement.
AC: Thank you for highlighting direct causation. We will reword this statement.

RC1: Line 67: The map of soil sampling sites is missing. This is a really helpful message to readers.
AC: Thank you for this suggestion. We will add a map showing the four regions and the locations where the soils were sampled from. We did not include it because we were unsure about the added value in terms of spatially explicit information (in addition to outputs and discussion), but as you say it can be convincing for the reader and can have illustrative purposes.

EC/RC2: Page 2 line 25 – A particularly strong positive .....
AC: Thank you. We will rephrase as suggested.

EC: Page 2 line 31: into the physical, chemical and biological major components
AC: Thank you for the correction. We will follow the suggestion.

EC/RC2: Page 2 line 36: not only mineral fertilizer but as well the soils' inherent fertility or nutrient status. Additional factors are the tillage regime, the planting date, staking, and stake height see: Enesi, O.R., Hauser, S., Lopez-Montez, A., Osunobi, O.

(2018) Yam Tuber and Maize Grain Yield Response to Cropping System Intensification in South-West Nigeria. Archives of Agronomy and Soil Science, 64:7, 953-966.
AC: Thank you for that suggestion. We will add inherent nutrient status to the example factors influencing tuber yields and make a reference to Enesi et al. (2018).

EC/RC2: Page 2 line 48 is it really the case that the IR approaches complement the wet chemistry or is it more that the IR approach requires the wet chemistry to make sense of the spectra?
AC: To avoid any confusion, we have changed the sentence to "...assess soil properties in a complementary manner to conventional laboratory analytical methods". The wet chemistry is fundamental to calibrate the attribute of interest from the spectra (see l51–54, p2). Once wet chemistry is done for a representative collection of soils, it needs be re-done to a considerably smaller extent because the aligned properties can be extracted from spectra and established libraries.

EC/RC2: Page 4 line 89 YAMSYS in full please
AC: Thanks for spotting this. We will correct it.

EC/RC2: 16 line 281 - Proposed sentence: To give a specific example, yam requires relatively large quantities of N and K (e.g., O'Sullivan, 2010); on light-textured soils yam can attain high tuber yields but at a high risk of losing large proportions of applied N and K, to the environment (e.g., Diby et al., 2011).
To be considered here is that if a large portion of N and K are lost yields are unlikely to be high
AC: We fully agree, thank you for that proposition. We will particularly incorporate that applying larger amounts of N and K at once would not improve yield potential under such situations.

EC/RC2: Page 2 line 25 – A particularly strong positive .....
AC: This is true. We will change this.

EC/RC2: Page 2 line 31: into the physical, chemical and biological major components
AC: Thank you for the correction, we will modify as suggested.

RC1: L110 Change to modelling approaches "for" SC
AC: This will be changed as suggested.

4.3 Methods

RC1: Line 144: 128 measurements? I am confused here as there are only 94 milled soil samples.
AC: These were internal measurements, repeated scans that were collected and returned as an average spectrum by the spectrometer and samples. This was done for improving the signal-to-noise ratio. We have added "scans", accordingly.

RC1: 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?
AC: The efficiency in spectroscopic prediction of soil properties stems from the simultaneous assessment of multiple soil properties. Once a calibration and libraries that fits the scope and purpose of newly collected soils, only a subset needs do be analyzed using expensive and time-consuming conventional analysis, for validation and possibly is made for refining the model. In general, both sample preparation and measurement take considerably longer for ED-XRF compared to mid-IR. It is intuitively clear that no modeled data can be more accurate than direct measurements (i.e., wet chemistry). It can be of importance, however, that spectroscopic prediction can on average produce more consistent results if there is large inherent unsystematic variation in a method, which is the case for some extraction-based techniques but much less for ED-XRF.

RC1: Line 153: I recommend adding the use of a non-linear model (e.g., Cubist, RandomForest) to better demonstrate the predictive ability of MIR.

AC: This is a good comment but we think adding such non-linear models might not be necessary given the soil and data context of our study. The reason is that the number of samples is relatively low compared to the number of spectral variables. This is where chemometric approaches such as PLSR excel and can provide sufficiently accurate outputs. In our study, PLSR on the data set and mixing the four data sets yielded accurate estimates for the properties that are well known to be spectrally predictable, with small errors (RMSEs) even when compared to farm or field scale libraries with much higher local sampling density (e.g., RMSE(total C) = 1.6 g C/kg). There are soil properties that have a variable spectral response rely on indirect estimation via correlated major chemical components (e.g. soil organic matter, clays), such as available P and a subset of total and available micronutrients. Due to the high complexity of soil chemistry more data does not necessarily mean higher predictability. Since such spectroscopic patterns and relationships might be only learned in a very local context, this might, if possible, only be addressed with higher sampling densities within specific regions or farms.

It is true that machine learning can be appropriate for deriving global/general rules and property estimates in larger spectral libraries, but much of the recent literature shows adaptation approaches (transfer) currently seem more appropriate and accurate to account for the characteristics of the target soils of a region or farm. Nevertheless, memory-based learning approaches or other transfer learning methods can be more appropriate for local adaption. However, this requires bigger libraries and harmonized protocols in order to make the workflow effective, and soil tests that are indicative of inherent soil fertility and crop yield response.

RC1: Line 172: The definition of Sy is missing in equation 2.
AC: Thank you for spotting missing notion of the standard deviation. We have defined it.

RC1: Lines 174-155: Until the end of this manuscript, I did not see any results relevant to the uncertainty analysis mentioned here.
AC: Perhaps, the reviewer seems to overlook it. Please note that we reported both uncertainty and goodness-of-fit estimates for the model data and cross-validated uncertainties on a point basis (for all samples). See e.g. Figure 3 and section 3.1 of the results.

RC1: 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.
AC: Thank you for making this consideration. Indeed, given that there is a major difference in the range and distribution of soil properties across the four regions, holding out single regions can lead to extrapolation. This is of course dependent on which region is left out. Cross-validation is appropriate enough to address the objectives of this work and we do not see immediate value of such an analysis. To apply calibrations from this small-sized library to newly sampled field in the regions, a spiking or data-driven selection approach might be necessary. For this purpose, we refer to Lobsey et. al (2017) and Ramirez-Lopez et. al, 2013.

4.4 Discussion

EC/RC2: Page 16 line 273 I would challenge the importance of SOC to store water – it is mainly the soil texture that determines water holding capacity, whereby SOC may play a positive role in aggregation and thus improve water holding capacity yet it is not primarily the "body" holding water. Please rephrase or remove the water
AC: Thank you. We will specify that the principal effect of organic matter regarding water holding properties is aggregation.

EC/RC2: Page 16 line 275 – land use pressure is a bad term what we look at is a higher land use frequency and shorter fallow – please rephrase accordingly.
AC: We agree with that suggestion and will modify accordingly.

EC/RC2: Page 16 line 276 – the sentence is too long and combines too many factors. It is correct that fallow and slash and burn go together, with the burning at least partially destroying the C input to the soil – please phrase to show the consequences of the cropping system in brief – monitoring the soil quality is a different aspect and would need being discussed in connection with the farm households ability to engage with such activities.

AC: Thank you for this valuable input. We will modify the sentence so that it links the topics in more straightforward manner. The primary goal of this study was to develop validated, interpretable, and publicly available spectroscopic models. We made validated and interpretable statements about soil quality variables specific to these four climatically different ecoregions. The feasibility to implement such quantitative monitoring methods goes way beyond the technical and scientific one, as you mention. For the moment we see rather immediate potential for collaborative farm trials that test sustainable soil and crop management options, involving researchers and farmers (YAMSYS project as an example). We have added that to the manuscript.

The conventional methods do simply not offer enough quantitative measurements of properties to derive region-specific and farm-adapted nutrient management strategies (combined management factors; variable inputs, fertilizers (organic and inorganic), different inherent soil fertility and variable risk for nutrient leaking etc.). Based on your comment, we will make that more clear in the manuscript what we mean by monitoring. We agree, science certainly can't resolve all issues, but our hope is at least that the availability of cheaper diagnostics will hopefully deliver better recommendations for understudied crops such as yam. Making spectral soil diagnostics available as a service so that it becomes more cost-effective and directly applicable to farmers and extension services offers a huge potential for more sustainable crop and soil management, but that goes way beyond what a single study can deliver.

EC/RC2: Page 16 line 278 – this sounds like closing the yield gap is a process that simply runs parallel to improving soil quality, however it would be the consequence and not something that happens automatically. Please break down the long sentence and separate cause and effect here. It would be of particular importance to clearly single out the monitoring of soil properties from activities improving soil properties – make sure it is clear that there are tools to improve soil quality and tools to monitor soil quality.

AC: Thank you for that suggestion. We will make the distinction between the crop and farm management activities more evident. This manuscript mainly tackles tools, data, and models to monitor soil quality, which we hope we have now highlighted more distinctly. Inevitably, cheap and simple tools to monitor soil quality are favored when evaluating the short and long-term effectiveness of tools to improve soil quality.

RC1: 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 Viscarra Rossel. Please also check the relevant typos (e.g., line 299).Sanderman, J., Savage, K., & Dangal, S. R. (2020). Mid-nfrared spectroscopy for prediction of soil health indicators in the United States. Soil Science Society of America Journal, 84(1), 251-261.

AC: Thank you for highlighting, we might consider adding this study to the introduction. It is interesting that in the suggested paper the authors refrain from making a final statement about the predictability of available nutrients. According to the authors there might not be enough data and because soils in this library were not collected with the objective of deriving available nutrient diagnostics. We see here a data gap that our study addresses. Mid-IR overall produces more accurate results when there is more soils sampled from smaller areas and at higher sampling density (local (chemometrics-based) models with locally-linear relationships), for example at field or farm scale is where most accurate predictions can be expected. We agree with the sampling density, but we are unsure about the relationship between R^2 and the number of samples in the library (Sanderman et al., 2020). The value and potential of larger and harmonized libraries is undisputed in the soil science community, since there is a trade-off between accuracy the extent of new conventional analyses needed for different applications that deal with a particular soil variability. Our main focus was to test the applicability of models for soil quality and nutrient diagnostics, specific to these landscapes in Burkina Faso and Ivory Coast.

5 References

England, J. R., & Viscarra Rossel, R. A. (2018). Proximal sensing for soil carbon accounting. SOIL, 4(2), 101–122. https://doi.org/10.5194/s 4-101-2018

Guerrero, C., Wetterlind, J., Stenberg, B., Mouazen, A. M., Gabarrón-Galeote, M. A., Ruiz-Sinoga, J. D., Zornoza, R., & Viscarra Rossel, R. A. (2016). Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy? Soil and Tillage Research, 155, 501–509. https://doi.org/10.1016/j.still.2015.07.008

Lobsey, C. R., Viscarra Rossel, R. A., Roudier, P., & Hedley, C. B. (2017). rs-local data-mines information from spectral libraries to improve local calibrations: rs-local improves local spectroscopic calibrations. European Journal of Soil Science. https://doi.org/10.1111/ejss.12490

Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Ben Dor, E., Brown, D. J., Clairotte, M., Csorba, A., Dardenne, P., Demattê, J. A. M., Genot, V., Guerrero, C., Knadel, M., Montanarella, L., Noon, C., Ramirez-Lopez, L., Robertson, J., … Wetterlind, J. (2015). Chapter Four - Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. In D. L. Sparks (Ed.), Advances in Agronomy (Vol. 132, pp. 139–159). Academic Press. http://www.sciencedirect.com/science/

Ramirez-Lopez, L., Behrens, T., Schmidt, K., Stevens, A., Demattê, J. A. M., & Scholten, T. (2013). The spectrum-based learner: A new local approach for modeling soil vis–NIR spectra of complex datasets. Geoderma, 195–196, 268–279. https://doi.org/10.101