



EGUsphere, referee comment RC2  
<https://doi.org/10.5194/egusphere-2022-246-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## **Comment on egusphere-2022-246**

Anonymous Referee #2

---

Referee comment on "Accuracy of regional-to-global soil maps for on-farm decision-making: are soil maps "good enough"?" by Jonathan J. Maynard et al., EGU sphere, <https://doi.org/10.5194/egusphere-2022-246-RC2>, 2022

---

### **Summary**

This manuscript compares soil maps (conventional and digital soil mapping) in terms of their prediction accuracies and their accuracies for soil suitability assessment for maize farms across Ghana. It aims to assess whether these soil maps could provide useful information for guiding farm management. This is an important question, and the manuscript is well written and provides a useful analysis, finding limitations in the ability of the soil maps to predict constrained soils. I have a few suggestions and queries below, but overall would think it could be published with some revisions.

### **Details**

Section 2.5 describes field-support validation of predictions and I like the fact that this is being done, but am not convinced by the details of how it has been applied. The described

process seems to be (for each field) first inferring a field boundary, second extracting from the soil map the predictions from all pixels within the field boundary, and third averaging these predictions; the result is compared with the average of the three sampled soil data from the field. But given your field-support validation data here are the average of three known points, shouldn't your associated prediction be the average of the predictions (as extracted from the soil maps) at these three same points? This could still be referred to as 'field-support validation', but would be comparing like with like, rather than two slightly different definitions of 'field support'. (While I don't expect this to make a huge difference to your results, I think it would be a more justifiable approach.)

Line 237: Is the Balanced error rate the (unweighted) average of the Producer's accuracies for the classes? I'm not convinced/sure about what this measure is saying – if there is a class (e.g. silty clay/sandy clay) with very few data (presumably because it is quite rare in the region of interest), won't the balanced error rate place too much importance on predictions of this class? Perhaps add a sentence after line 237 to clarify why use the balanced error rate. (I could see why a BER as a weighted average of PA values could be useful, for instance where you know that for some reason some classes were over-represented in your data, and had some knowledge of what the real proportions of the classes across the region of interest should be – but to use equal weights in the averaging seems to me to be saying that your initial expected proportions were equal for all classes, which seems unlikely to me.)

I wonder if another test might be helpful to provide further context about the quality of predictions from the maps. With the dataset that has 3 points sampled in each field, you could test how accurate predictions would be if you used just one of the sampled points from the field as 'representative' of the field, and evaluated how accurate this was at predicting the two other points in the field (to evaluate how good management would be based on a single 'representative' sample from the field). If there is a lot of within-field variation, even this might give poor prediction accuracies. This might add further insight into what is written on line 562 – "...fail to meet the accuracy requirements..." – would the use of a sample from the field also fail?

Line 431: I can't see the numbers you are referring to here in Table 2, they seem to me to be a lot higher than this – from my interpretation of the numbers in Table 2, it seems that OA-adj was over 80% in many cases (eg iSDA, OA-adj, soil texture: 0.9). Also check the numbers on line 593, and check the numbers throughout.

Regarding everything being predicted as 'no constraint' (eg by iSDA, Fig 8) – I think that predicting the more extreme values of soil constraints is always going to be difficult, and the 'expected value' as extracted from the soil map will rarely give the extreme values – I guess a bit like predicting rare events. A possible point for discussion is that you could use the map of predictions + its uncertainty to give (for each pixel/prediction location) a probability of constraint, and this could be a more appropriate way of informing management decisions. (Although this would very rarely be done in practice, could note that tools to implement this type of analysis of the soil map + its uncertainty map could be something worth looking at in future?)