SOIL Discuss., author comment AC5
https://doi.org/10.5194/soil-2021-80-AC5, 2022

Reply on RC1

David G. Rossiter et al.

Author comment on "How well does Digital Soil Mapping represent soil geography? An investigation from the USA" by David G. Rossiter et al., SOIL Discuss., https://doi.org/10.5194/soil-2021-80-AC5, 2022

RC1: 'Comment on soil-2021-80', Anonymous Referee #1, 02 Oct 2021

We appreciate the reviewer's careful summary and appreciation of our objective.

1. Reply to summary and general comment:

1.1 "The examples given in the paper do not fully demonstrate that the proposed metrics are necessary, relevant and sufficient to provide a comprehensive evaluation of the different DSM products that can help an unexperienced end-user in choosing between the different available DSM products he/she can get in a given local territory."

We could not come to any conclusion on this -- indeed as the reviewer points out, "No clear and convincing hierarchy across the three DSM products is revealed by the analysis of the result."

The reviewer presents the following three questions to be addressed in the Discussion:

Q: "Which metric best account for the visual differences of soil patterns?"
Q: "Which metrics are redundant to each others?"
Q: "Which metrics best discriminate the different DSM products?"

In the revision we will directly address these, as suggested.

1.2. "The authors used a fairly unusual way to present their data (section 2) i.e., instead of presenting separately each DSM and soil survey products, they chose to deliver the information progressively and “in parallel” along a set of sections that are not always straightforward for an external reader. Furthermore, this induces some redundancies and contradictions."

The "redundancies and contradictions" are dealt with in the response to detailed comments (see below). Other than that, we are not clear on why the presentation is "not always straightforward."

1.3 "The study areas selected as examples for regional and local spatial patterns are, if not presented (the local), not presented with the necessary details (regional) for allowing the interpretation of the results".

These are discussed in detail in the companion Case Studies document, as the first of the four case studies. We omitted them from this paper in order to focus on the methodology and to keep the approach generic. Our aim is to show how an assessment is performed, not primarily to assess this specific case study. However we agree this would aid understanding, so we will repeat the context information from the Case Study document in the main paper, as requested.

1.4 "The presentation of the soil data considered in this paper (actually in sections 2 [Data sources] and 4 [Example area and soil property]) needs to be deeply reworked to improve the understanding by an external reader not familiar with US context."

As explained above, this presentation is already in the Case Studies, but since many readers of this paper will not go to the Case Studies, we will repeat the soil data and context explanations here, and also refer to this specific case in the Conclusions.

2. Reply to comments along the text: (1) Those of form/consistency will be dealt with by appropriate edits; (2) Those of substance are discussed now.

2.1 "Title and line 1: Why do the authors rename "predictive soil mapping" what is currently known as "Digital Soil Mapping"? Indeed, why? The first author has preferred the idea of computer-generated maps as "predictive", and has followed the terminology of Scull (1993). However as pointed out in the Community Comment from the University of Sydney, all maps are in some sense predictive, and DSM is by far the most-used term. The revised paper, including the title, will be adjusted accordingly.

2.2 "Line 114: "data sources" is only the first part of the section (before 2.1.) isn't it? If yes you should replace "data source" by "soil data" and add a subsection "data source" immediately after."

Indeed this is not an appropriate heading. We will change it to "Products to be compared"; this then segues into the next primary section "Evaluation methods".

2.3 "Line 138: Contrary to what is suggested here "Polaris soil properties" is not further systematically replaced by "PSP". A lot of "POLARIS" remains in the text and in the figure. This should be corrected." We are sorry for our lack of care, and will make the text consistent. We use PSP for the soil properties product of the POLARIS project, which also includes the soil class predictions of the original POLARIS (since updated).

2.4 "Line 152: gSSURGO has not been presented before"

The explanation of gSSURGO and its distinction from gNATSGO will be added to the paragraph beginning at L122.

2.5 "why mixing in a same section "environmental Covariates" and "geographic scope"? The relation between them is weak. The statements that deal with the latter (Lines 191-193 and 195) would be better located in the next "mapping methods" section (see my next comment)"

We combined these because the geographic scope of the covariates used differs among the DSM products. We wanted to emphasize that SPCG only uses covariates covering CONUS, and that PSP uses only covariates within each of its tiles, while SG2 uses global covariates.

2.6 "perhaps more comfortable for the reader to know first the set of covariates used by SG2 and also, as I understood, by the two other DSM products. Then cite the specific covariates that have been added to PSP and SPCG."

2.7 "Lines 197-205 : this paragraph is largely redundant with lines 151-161."

2.8 "Lines 218-224 : insert here the geographical scope and the number of samples used for training the models."

2.9 "Line 231: I understand here that gNATSGO is a generalization of gSSURGO. However it was stated before (line 125) that "gNATSGO is a composite of [....] SSURGO and [....] STATGEO and [....] RSS. Please, explain more clearly what are the differences between gNATSGO and gSSURGO."

Indeed there is confusion here. This will be explained in the paragraph starting at L122 (as explained in comment 2.4).

2.10 "Lines 255-257: Even before looking at your further results, we could expect that soil survey products and DSM products do not converge toward similar uncertainties assessments. At best, we could expect that the level of uncertainties mapped by these two products could be ranked similarly, independently from their absolute values. You should select a metric for representing this."

We used the IQR (5-95%) as per GlobalSoilMap specifications to characterize SG2 and PSP; these are compatible. The inclusion of the low-high estimate range from gNATSGO is purely informative, we make it clear that this is based on expert opinion backed up (in some cases) by pedon data, but is not meant as specific points on a probability distribution.

2.11 "Line 272: "RMSD adjusted for MD". Not clear. A mathematical formula would clarify."

We think text alone is sufficient, if we rephrase this as "RMSD adjusted for MD, i.e., the RMSD after subtracting the bias, i.e., the MD, from each prediction."

2.12 "Line 301: "region" is qualitative. How can we calculate the variance of a qualitative variable?"

The regions are not qualitative, they are geographic extents of the histogram-equalized class map (note that 'zones' is used for the same concept on the reference map). The variance is ...

2.13 "Lines 359-375: The sections "Regional patterns" and "local patterns" are redundant with the introduction of section 3 (lines 259-265). This should be re-organized."

2.14 "Lines 377-349 (section 3.6.1. "visual method") Does this section refer only to "local patterns" (section 3.6.1.)? I don't think so since you provided further visual comparisons for both regional and local patterns. This should be clarified. Furthermore, this section looks redundant with section 3.1. ("qualitative methods")"

2.15 "Figure 1: This study area looks different from the ones considered further (Figures 4 to 12). This could explain why there is a so great discrepancy between the visual inspection and the quantitative results obtained further. Similar problem occurs with figure 13 (what is this study area?) . Please give the length and width of the rectangle for a better appreciation of the scale."

2.16 "Lines 393-402 (section 4): It would be useful to have more information about the study area that is finally selected as example of regional spatial pattern comparisons (size of the rectangle, scale of the soil survey product gNATSGO at this location, average size of polygons, pedology, landscape drivers of soil variability etc...). All these data would be very useful for interpreting the results. This comment applies also to the description of the study area selected further as example of local spatial patterns comparisons. I did not find any information on this area."

This is explained above (1.3). Indeed this information greatly helps the interpretation of the relative success of the DSM products.

2.17 "Line 412: provide the significance of the sizes of the circles in figure 3. Indicate what you mean by "well-correlated" (threshold?)
Line 464: There is an apparent contradiction with the concluding statement here "overall the agreement is fairly good" and what is written just before (line 459):
"gNATSGO is considerably different from all other products""

2.18 "Lines 539-541: I suppose you should replace "...and covariates limited in geographic scope to the USA" by "...and input soil data limited in geographic scope to the USA". To my opinion, this is the most surprising result of the paper. Have you any explanation ? to my opinion, the similarity of machine learning algorithms cannot be a convincing explanation."

It is so that the covariate coverage was limited to the USA, this was not made clear. The covariates are (mostly) the same but in this case the model does not try to use covariates outside the USA -- although that wouldn't make a difference anyway, because no points for model training would be located there. We have clarified the statement as suggested.

We were also quite surprised by this result. We've thought some more about this and have added some text with tentative explanations. We do still think that the machine learning algorithm is part of the explanation.

2.19 "Line 543-544. you cannot conclude that because the specifications for defining these confidence intervals are different between the DSM and the soil survey products. Furthermore, there is not any ground truth to identify what is the most realistic CI."

Correct, there is no ground truth for the CI. However the CI is applied point-wise (actually, grid-cell-wise)

2.10 "Lines 556-560. I disagree with your diagnosis on PSP. Figure 14 clearly shows that PSP does not bring more knowledge on soil variations than the initial soil survey product from which it was derived. Furthermore, you cannot say that PSP can be useful for unsurveyed areas since PSP requires a soil map as input"

We think the reviewer means that what we say in general about PSM in the indicated lines does not apply to the PSM method of (what we will, in the revision, call) DSM.

(1) We can not find where we claimed that PSP (or any DSM method) brings more knowledge on soil variations than the initial soil survey product; certaintly we don't say that in the indicated lines. We do say that some covariates used in DSM show variation not considered by the original mappers. An example is \gamma-ray surveys. Therefore the output of DSM should be compared with the field survey to see if it might reveal otherwise unrecognized patterns.

Figure 14 shows that PSP does disaggregate polygons (although in this case incorrectly, we think). It's true that the knowledge for the disaggregation comes from the initial soil survey product, so in that sense the reviewer is correct, there is no more knowledge.

(2) Indeed PSP can be used to extrapolate: it takes the constructed model and applies it to the covariate stack in the extrapolation area. It has been calibrated on a map of soil series (and from these, their soil properties) in a mapped area, but once the model is built it can be applied anywhere -- a set of soil series will be predicted, and from that, soil properties. Of course this may not be wise.