

Interactive comment on “Error propagation in spectrometric functions of soil organic carbon” by Ellinger et al.: Response to RC3

General comments

Most comments on technical and minor topical suggestions were addressed in the new manuscript.

- 5 Authors’ reply: *“The PLSR components vary largely in dependence on the pre-processing method. The information on the number of selected components, therefore, did not result very informative. Furthermore, it distracts the reader from the main message. We, therefore, refrain from including it in this publication. We have adapted the results section concerning the uncertainty scenarios. We have also extended the introduction section to elaborate on the various sources of uncertainty.”.*

- 10 I insist on giving finally selected tuning parameters for all the modeled scenarios. The reader requires such information to judge the uncertainty in relation to model complexity within different error propagation scenarios. Hyperparameter optimization through Model tuning is a key aspect of spectral modeling, and all scientifically-sound publications in this field report these results. Parameter selection results can easily be included within a model assessment table. This does not distract the reader. This information will reveal also potential over-fitting effects due to replicate spectra in the inner tuning loop in presence of multiple spectral replicates.

- 15 The main concern from the last review round has been mentioned in the conclusion, but is still not resolved. Specifically, the group stratification for replicates of the same sample was not performed for the tuning procedure. The author’s specifically mention this now in the conclusion. However, the author’s refuse to report the finally selected PLSR parameters, but at the same time admit that suboptimal model parameters might be selected. Based on the results shown, no conclusive statement on this issue can be drawn.

- 20 In order to get this manuscript eligible for final publication, the issues arising from multiple replicates in tuning during cross-validation needs to be addressed. At least, the authors should report a majority consensus value of the number of PLSR components selected in the final model. This will at least indicate whether the model errors are biased by resampling artifacts.

Specific comments

Abstract

- 25 The statement in l. 24–26 about precise protocol and measurement protocol is out of scope and needs to be removed. It is not the main topic of this study. Such protocols and procedures are mostly well documented in soil spectroscopy literature and there are recommendations on this, consider for example Wetterlind et al., 2013.

Introduction

- 30 The introduction needs a major rewrite, there are many grammatical, topical, and stylistic errors. See the technical corrections for some examples and suggestions.

Material and Methods

Authors’ reply: *“Reference to soil treatment and scattering effects was made in the introduction. We refrain from referring to soil texture as we are at within-field scale and do not have a pronounced textural variability in our dataset. A reference to sample origin is included in the discussion section.”*

Although the present study covers within-field variability, general conclusions regarding spectral error propagation need to consider soil texture as it affects scattering and averaging effects.

Results and Discussion

5 Some paragraphs contain statements that should be in the Material and Methods section, see e.g. lines 269–271 (see also technical corrections). Table 3 and the corresponding text needs to be moved to the Material and Methods section.

Many sections contain present tense where past tense is needed.

The section 3.1 on soil organic carbon reference values is too long, and needs some general revision. This section needs to be further summarized and presented in a more concise manner.

10 L. 283–284: “The plots for “A” and all samples show higher and lower SOC values than the archive data due to the fact that those data are obtained from compound samples for one plot.”: It is not clear what the authors mean by compound plots..

Authors’ reply: We are not aware of any study that actually quantified the effect of spectral pre-processing on model performance and, therefore, refrain from deleting it from our study.

This is simply not correct. There are many soil spectral studies addressing pre-processing with regard to model performance. Please consider e.g. Stevens et al., 2013 (see References at the end).

15 The discussion around the pre-processing is still way too long.

Table 4 on R^2 values is not informative. First, it only reports R^2 , which is strictly not a measure of performance. Second, it is relative to the range of the measured property, which is not given. The only reference to this table is that the error conditional on the input data is different, and this information is missing in the table. Therefore, this table and discussion around it should be removed.

20 In general, the authors should stay focused on the key topics under investigation. Many sections are too long and therefore distract the reader. The Results and Discussion requires some more work to offer the audience a better flow.

Conclusion

The conclusion should fit on half a page. Focus on the key findings and topics that the study addressed.

Technical corrections

25 L. 12: “...the exact monitoring of...” -> “...precise monitoring of...”

L. 14 “...to enhance conventional SOC analysis and has often been used to predict SOC” -> “...to complement conventional SOC analysis.”

L. 24–26: “We emphasize...and allow for a comparison between publications.”

L. 35: “production of energy”: Energy cannot be produced, rather use “production of energy crops”

30 L. 36: remove Stenberg et al., 2010: this does not fit the context.

L. 36–37: “quality of soil” needs to be described, too generic -> soil properties, soil type...

L. 38–40: “...SOC is also interesting when it comes to the global warning issue...” -> “SOC is also relevant for the global warning issue...”

L. 66: “However, the application of”. -> Delete “However” because there is no reference sentence.

L. 67: “...standard lab analysis” -> “...standard laboratory analysis”

5 L. 68: “...on the other hand side” -> either “hand” or “side”

... (see general comments for abstract)

L. 100: Missing dot after “(Merbach and Schultz, 2013)”

L. 131: “Elementaranalysator” -> “elemental analyser”

L. 155–160: Give original publications for all pre-processing techniques; only give these and remove the other references.

10 L. 160–182: All references that are not original method publications for pre-processing need to be removed.

L. 276: “the SOC distribution of “A” and “B” samples differ” -> “the SOC distribution of “A” and “B” samples differed”

L. 307–308: “The model results are now compared based on their mean $RMSE_{MV}$ and their interquartile range”: this should be in Material and Methods...

L. 308: “It is not surprising that the dataset...”: style -> use “We expected that/It was expected that...” or similar

15 L. 312: “It seems that the within sample variation concerning soil spectra con somehow compendate the within sample variability concerning SOC within the model building process, although replicate measurements do not match” -> consistently use past tense.

L. 387: “cross validation” -> “cross-validation”

All figures need to be in vector graphics format or need a better resolution.

20 Figure 9: The text annotation for “RMSE” must be changed from “RMSE <number>” to “RMSE = <number>”

References

Stevens, A., Nocita, M., Tóth, G., Montanarella, L., van Wesemael, B. (2013). Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. PLoS ONE, 8(6), e66409. <https://doi.org/10.1371/journal.pone.0066409>

25 Wetterlind, J., Stenberg, B., Rossel, R. A. V. (2013). Soil Analysis Using Visible and Near Infrared Spectroscopy. In F. J. M. Maathuis (Ed.), Plant Mineral Nutrients (pp. 95–107). https://doi.org/10.1007/978-1-62703-152-3_6

References

REFERENCE 1

REFERENCE 2