SOIL Discuss.,
https://doi.org/10.5194/soil-2019-48-RC1, 2019

Interactive comment on "Estimation of effective calibration sample size using visible near infrared spectroscopy: deep learning vs machine learning" *by* Wartini Ng et al.

Anonymous Referee #1

Received and published: 19 October 2019

The paper presents a case study of prediction performance comparisons between a couple of standard machine learning methods commonly used in soil spectroscopy (PLS and Cubist) and a deep learning algorithm (convolutional neural networks, CNN). These algorithms are tested in a large soil spectral library. The paper clearly shows that CNN outperforms PLS and Cubist when the number of calibration observations is large. All the algorithms tend to perform poorly when they are used in small (calibration) sample sets. In my opinion, the manuscript does not present a clear contribution to soil science. I just have some comments that I hope help the authors to improve their manuscript.

Printer-friendly version

Discussion paper

General comments:

- The method used by the authors to estimate the effective calibration sample size is entirely based on prediction performance indicators (e.g. root mean square error) which requires prior knowledge of the response variables of the samples used as candidates for calibration. Therefore, the method is rather unrealistic/impractical.

- Since the main objective of the paper is related to calibration sampling for soil spectroscopy, I encourage the authors to review the literature available on this topic. This might help to clearly identify research needs and also to identify already available methods to optimize the amount of samples used in calibration (see Esbensen et al., 2014; Ramirez-Lopez et al., 2014; De Gruijter et al., 2006 ; Petersen et al., 2005; Minkkinen, 2004).

- The effective size of the calibration set for a given spectral dataset largely depends on the variability or complexity embedded in such dataset. For example, a small area where a large number of soil spectra is available (as in the case of on-the-go soil spectroscopy), the optimal size of the calibration set would be rather small. Furthermore, in such non-complex scenario, the use of CNN would be arguable, as the conventional methods would be expected to perform well (as it has been proven). In this respect, the authors seem to focus only on the size of the calibration sets disregarding very important aspects of the theory of sampling (see Minkkinen, 2004) and draw general conclusions from a single experimental dataset.

- The conclusions are not clear and despite their original research question (how many samples are required to get CNN performing better than PLS and Cubist) is answered for their particular dataset, there is no useful procedure or method presented by the authors to reproduce or extrapolate this to other cases in a useful way.

Specific comments:

- Section 3 (chemometrics model): the authors need to provide information on model

optimization and references. For example, why do they choose a learning rate of 0.001 and adam optmizer, what does it mean? Is there any reference the readers can be referred to? How many PLS components and committees were tested in PLS and cubist respectively? Optimization of the algorithms play a key role in their performance.

- Section 4.4 (sensitivity analysis): this whole section does not seem to bring any significant contribution to the objectives of the paper.

- Section 4.4 (sensitivity analysis): The estimations of the importance of variables for modeling for different modeling algorithms are based on different methods, therefore the comparisons between the results carried out in the paper are not appropriate.

- Section 4.4 (sensitivity analysis): the authors need to be more clear with their statement: “the wavelengths used Cubist were derived based on model usage”.

References

De Gruijter, J., Brus, D. J., Bierkens, M. F., & Knotters, M. (2006). Sampling for natural resource monitoring. Springer Science & Business Media.

Esbensen, K. H., & Wagner, C. (2014). Theory of sampling (TOS) versus measurement uncertainty (MU)–A call for integration. TrAC Trends in Analytical Chemistry, 57, 93-106.

Minkkinen, P. (2004). Practical applications of sampling theory. Chemometrics and intelligent laboratory systems, 74(1), 85-94.

Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., & McBratney, A. B. (2019). Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. Geoderma, 352, 251-267.

Padarian, J., Minasny, B., & McBratney, A. B. (2019). Using deep learning to predict soil properties from regional spectral data. Geoderma Regional, 16, e00198.

Petersen, L., Minkkinen, P., & Esbensen, K. H. (2005). Representative sampling for reliable data analysis: theory of sampling. Chemometrics and intelligent laboratory systems, 77(1-2), 261-277.

Ramirez-Lopez, L., Schmidt, K., Behrens, T., Van Wesemael, B., Demattê, J. A., & Scholten, T. (2014). Sampling optimal calibration sets in soil infrared spectroscopy. Geoderma, 226, 140-150.

Interactive comment on SOIL Discuss., https://doi.org/10.5194/soil-2019-48, 2019.

Printer-friendly version

Discussion paper