Do we really need large spectral libraries for the assessment of soil organic carbon at local scale?

César Guerrero (1), Johanna Wetterlind (2), Bo Stenberg (2), Raphael A. Viscarra Rossel (3), Raúl Zornoza (4), Fernando T. Maestre (5), Abdul M. Mouazen (6), Boyan Kuang (6), José Damián Ruiz-Sinoga (7), and Miguel A. Gabarrón-Galeote (8)

(1) University Miguel Hernández, Agrochemistry and Environment, Elche, Spain (cesar.guerrero@umh.es), (2) Swedish University of Agricultural Sciences (SLU), Department of Soil and Environment, Skara, Sweden, (3) CSIRO Land and Water, Canberra, Australia, (4) Polytechnic University of Cartagena, Department of Agricultural Science and Technology, Cartagena, Spain, (5) University Rey Juan Carlos, Department of Biology and Geology, Móstoles, Spain, (6) Cranfield University, Department of Environmental Science and technology, UK, (7) University of Málaga, Department of Geography, Málaga, Spain, (8) Catholic University of Louvain, Earth and Life Institute, Belgium

Spiking is an approach to improve the accuracy of large-scale spectroscopic models when they are used to predict at local scale. But, if models are to be spiked, do we really need large-sized spectral libraries? Different calibrations relating the SOC and NIR spectra were obtained using PLS as regression method: i) model #1: local-scale model (n=40); ii) model #2: local-scale model (n=88); iii) model #3: provincial-scale model (n=147); iv) model #4: provincial-scale model, constructed with 50% of samples used in model #3 (n=73); v) model #5: provincial-scale model, constructed with 25% of samples used in model #3 (n=36); vi) model #6: national-scale model (n=1096); vii) model #7: national-scale model, constructed with 33% of samples used in model #6 (n=362). Each of these models was used to predict the SOC contents in target site samples. In this work, nine target sites were evaluated. Each target site is a relatively small area (from several hectares to a few square kilometers), where a dense sampling was made. The coefficient of the determination (R^2), root mean square error of prediction (RMSEP), bias, standard error of prediction (SEP) and the ratio of performance to deviance (RPD) were calculated pooling the predictions of the nine target sites. In overall, more than 900 local samples were predicted. The highest R^2 values were obtained with the national-scale models ($R^2 > 0.85$), and the lowest R^2 values were obtained with the models of small size. In general, the RMSEP tended to decrease with the increase of the models size. However, the predictions obtained with the large-sized models were clearly biased, and despite the high R^2 values, the RPD values were below 1.2. We also obtained predictions when these models were spiked with eight local samples (i.e., from the target site). After spiking, the predictions obtained with the small-sized models were substantially improved. As example of the changes due to spiking, the predictions obtained with the smallest-sized model changed the R^2 from 0.03 to 0.96, the RMSEP from 8.02% to 0.61% SOC, and the RPD from 0.39 to 5.20. The spiking effects on the large-sized models were clearly smaller than in small-sized models. The added samples (i.e., the spiking subset) were more influential on the small-sized than on the larger-sized models. We also obtained predictions when the spiking subset was extra-weighted. The addition of several copies of the spiking subset increases the statistical weight of these samples in the model, becoming more important than the other samples. Thus, the calibrations are forced to fit preferentially to the extra-weighted samples. If the extra-weighted samples are representative of the target site, then, an improvement of the predictions is expected. Indeed, the predictions were better than those obtained with the spiked model. In this case, the most important improvements of the prediction quality were observed in large-sized models, but the best results were obtained using small-sized models. When both approaches are used (spiking with extra-weight) the results were very accurate ($R^2 > 0.96$; RMSEP<0.71% SOC; bias<0.21% SOC; RPD>4.4) regardless of the number of samples used to construct the model. This is an interesting result, since it suggests that large libraries are not needed to use the NIR spectroscopy as technique to predict at local scale. In contrast, a few local samples (spiking subset) must be analyzed with the reference method. These results suggest that the analytical efforts shouldn't be focused exclusively in the development of large-sized calibrations.

Acknowledgements