Do we really need large spectral libraries for local scale SOC assessment with NIR spectroscopy?

César Guerrero[a],*, Johanna Wetterlind[b], Bo Stenberg[b], Abdul M. Mouazen[c], Miguel A. Gabarrón–Galeote[d], José D. Ruiz–Sinoga[d], Raul Zornoza[e], Raphael A. Viscarra Rossel[f]

[a] Departamento de Agroquímica y Medio Ambiente. Universidad Miguel Hernández de Elche.
[b] Department of Soil and Environment, Swedish University of Agricultural Sciences, Skara, Sweden.
[c] Cranfield Soil and AgriFood Institute, Cranfield University, Bedfordshire, UK.
[d] Departamento de Geografía, Universidad de Málaga, Málaga, Spain.
[e] Departamento de Ciencia y Tecnología Agraria, Universidad Politécnica de Cartagena, Cartagena, Spain.
[f] CSIRO Land and Water, Canberra, Australia
* Corresponding author: César Guerrero. Avenida de la Universidad s/n, E-03202, Elche, Spain. Tel.: +34 966658333. E–mail: cesar.guerrero@umh.es

Abstract
Near infrared (NIR) spectroscopy was used to predict the soil organic carbon (SOC) contents at local scale in eleven target sites. For that, eight spectral libraries of different sizes (ranging from 3482 to 36 samples) were used to construct national, provincial and local scale models. Inaccurate predictions were obtained except when the largest national library was used to construct the model. We also obtained SOC predictions once the models were adapted to target sites characteristics. For the models' adaptation, we used a two–step approach consisting on spiking (as first step) and extra–weighting (as second step). The effect of spiking was small in larger–sized models and high in smaller–sized models, whereas the effect of extra–weighting was small in smaller–sized models and large in larger–sized models. The very high accuracy obtained after models' adaptation (R^2>0.95; RPIQ>5.48), regardless of the size of the spectral library, suggests that large spectral libraries are not needed for local scale SOC assessment. These results have important implications regarding the way that NIR spectroscopy can result highly effective for land management and how users can focus and organize the analytical efforts.

Keywords: near infrared spectroscopy, soil organic carbon, spiking, extra–weighting, local–scale, diffuse reflectance spectroscopy

Soil and Tillage Research

http://dx.doi.org/10.1016/j.still.2015.07.008

1. Introduction

The soil organic carbon (SOC) content is the most important factor affecting the presence, activity and characteristics of the soil microbial communities, plays a central role in soil quality, and due to global warming issues, there is an increasing interest in monitoring changes in SOC stocks. Due to the high spatial variability, a precise quantification of changes in SOC stocks requires large numbers of samples, otherwise the statistical robustness of comparisons would be constrained by large uncertainties, prone to lead to misleading conclusions (Mäkipää et al., 2008; Muukkonen et al., 2009). Therefore, fast, cheap and accurate methods should be preferentially used for SOC analysis, allowing the analysis of thousands of samples with limited budgets and efforts. The near infrared (NIR) reflectance spectroscopy is a technique which meets these characteristics. The NIR spectrum of a soil sample contains abundant information about physical, chemical and biological properties. However, due to the characteristics of the NIR spectra, multivariate models relating NIR spectra with the soil properties are needed to extract that information. Several reports show the capacity of the combined use of NIR spectroscopy and chemometrics to predict several soil properties, including SOC (eg. Stenberg et al., 2010). The models should contain spectra similar to those to be predicted, otherwise, the predictions cannot be expected to be accurate (Stenberg et al., 2010). Since the soils can be extremely different in composition and properties, there is a trend to develop large spectral libraries, in order to ensure that derived models contain similar samples to those to predict (Shepherd and Walsh, 2002; Nocita et al., 2014). However, the adequateness of large–scale (regional, national, continental or global) models cannot be guaranteed when they are applied at local scale (Brown et al., 2005; Wetterlind and Stenberg, 2010; Guerrero et al., 2010; Kuang and Mouazen, 2013), especially for underrepresented soil types (Guerrero et al., 2014). In addition, the development of large spectral libraries requires huge efforts, being expensive, and the spectral measurements are dependent on the instrument and laboratory conditions (Ge et al., 2011). Moreover, the physical accessibility to the spectral libraries samples is sometimes restricted, making the development of calibration transfer difficult, and sometimes impractical or even impossible. These are serious limitations to the big potential of the technique, and frequently, new users decide to develop their own libraries and models, being a step which partially diminishes the NIR advantages.

Recently, Guerrero et al. (2014) designed an approach to ensure the adequateness of models to new target sites. The approach was focused on the adaptation of models to the characteristics of the target site's samples. This targeting approach was based on spiking, which implies an analytical effort, since a spiking subset must be analyzed with the reference method. However, if the approach is used in combination with small–sized spectral libraries, then the overall analytical effort could be drastically reduced, since the development of large spectral libraries might be avoided. Thus, we question the actual necessity of that important effort. Therefore, we found it interesting to compare the accuracy obtained with that approach when applied to a wide range of spectral libraries of different sizes. In the present study, eight different–sized spectral libraries (ranging from 3482 to 36 samples) have been considered, and the SOC contents have been predicted in 886 samples from eleven local target sites. The model's adaptation approach shown in Guerrero et al. (2014) essentially consists of two main steps: 1) spiking with representative samples, and 2) apply an extra–weight to the spiking subset. In order to show clearly the effects of both steps on prediction accuracy, the SOC contents were predicted using (i) unspiked models (UM), which were constructed only with the samples from the spectra libraries, (ii) spiked models (SM), and (iii) spiked models where the spiking subset was extra–weighted (EW).

2. Material and Methods

2.1. Spectral libraries

Eight spectral libraries were selected (Table 1). Three spectral libraries (SL1, SL2 and SL3) contain soil samples collected at national scale (Spain), three spectral libraries (SL4, SL6 and SL8) contain soil samples collected at provincial scale (Alicante Province, in Spain), and two spectral libraries (SL5 and SL7) contain soil samples collected at local scale. The SL5 contains samples from an agricultural area (78 ha) located in Sjöstorp (Southern Sweden). The SL7 contains soil samples from a *Pinus sylvestris* L. forest (ca. 350 ha) located in Peñalén (Guadalajara, Spain). The SL1 was built with soil samples collected for different purposes during independent samplings, but also with soil samples which were exclusively collected for the development of a large and diverse soil spectral library. Therefore, a portion of the samples were randomly collected, while others were collected on the basis of spatial criterion and also to cover the most important (i.e., most abundant) lithologies, climatic conditions, vegetation types and land uses. The same procedures were followed for the development of the SL4. However, only the spatial criterion was used for the collection of samples comprising the local–scale libraries SL5 and SL7.

2.2. Target sites

A total of 974 topsoil samples were collected in eleven local target sites. Each target site is a relatively small area, ranging from several hectares to a few square kilometres. The target sites represent examples where a dense sampling was needed at local–scale, although for different purposes. Therefore, there was not a common sampling scheme. In some target sites, the samples were randomly collected, while in other sites the sampling was stratified, and in others regularly distributed in space (grid). Ten sites were located in Spain and one in

the UK. These sites were located, at least, at a distance larger than 20 kilometres from samples included in any of the spectral libraries. The pH of the soils included in the target sites ranged from acid to basic, and were developed over different parent materials, mostly sedimentary (including evaporitic) but also metamorphic rocks, and covering contrasted climatic conditions. Moreover, different land uses can be found in the target sites. The most important characteristics of these target sites are shown in Table 2.

Table 1. Main characteristics of the spectral libraries (SL)

			Soil organic carbon (g C kg-1) statistics					
	Scale / Country	n	Minimum	Maximum	Mean	SD[c]	Skewness	Kurtosis
SL1	National (Spain)	3482	0.19	144.9	18.9	23.8	2.55	6.42
SL2[a]	National (Spain)	1096	1.22	125.5	10.0	9.28	5.14	39.1
SL3[a]	National (Spain)	362	1.54	125.5	10.8	12.2	4.92	31.2
SL4	Provincial (Alicante, Spain)	147	4.02	142.4	58.5	35.5	0.07	–1.08
SL5	Local (Sjöstorp, Sweden)	123	12.0	33.8	18.0	4.66	1.74	2.70
SL6[b]	Provincial (Alicante, Spain)	73	4.48	142.4	60.9	35.8	0.12	–0.98
SL7	Local (Peñalén, Spain)	40	13.6	102.1	41.1	16.6	1.02	2.45
SL8[b]	Provincial (Alicante, Spain)	36	8.33	142.4	60.6	35.5	0.36	–0.81

[a] Randomly selected subsets of SL1
[b] Randomly selected subsets of SL4
[c] SD: standard deviation

2.3. Models types

In all cases, the models relating SOC with NIR spectra were constructed with partial least squares regression as multivariate method, using the OPUS software (version 6.5; BrukerOptik GmbH, Ettlingen, Germany). The three types of models were constructed as follows:

i. *Unspiked model* (UM): one model was obtained from each of the eight spectral libraries. Thus, eight UM were obtained. These models were constructed using only samples from the specific spectral library (Appendix A). The spectra were preprocessed by the first derivative, and cross validation (leave–one–out) was used to select the number of PLS–vectors (or rank), based on the root mean square error of cross validation (RMSECV) changes, observed in a scree-plot.

ii. *Spiked model* (SM): these models were constructed with samples from the specific spectral library and a spiking subset, which consisted of eight local samples from the target site. The spiked models were unique for each target site (Appendix A). Since we have eight spectral libraries and eleven target sites, then, a total of 88 SM were calibrated. For the selection of the spiking subset of a target site, a principal component analysis (PCA) was performed with the NIR spectra of the target site samples. Then, the Kennard–Stone algorithm (Kennard and Stone, 1969) was used to select eight local samples whose scores were evenly distributed across the space defined by the first three PCs (Guerrero et al., 2014). A different PCA was performed separately for each target site. In this case the model's rank was set to the same number as in the corresponding UM (i.e., all the spiked models derived from the same SL had the same rank).

iii. *Spiked model where the spiking subset was extra–weighted* (EW): in each of the 88 above–described SM, the spiking subset was extra–weighted; next, the model was recalibrated. For extra–weighting, the 'statistical' weight of the spiking subset was increased 'n' times, being 'n' the ratio between the size of the spectral library and the spiking subset. The extra–weight was achieved adding 'n' copies of the spiking subset (Guerrero et al., 2014). The rank of these models was set to the same as in the corresponding SM. Thus, all the models derived from the same SL (regardless their type) had the same rank.

2.4. Predictions

In each of the local target sites, the SOC contents were predicted with the different models (UM, SM and EW) derived from SL1. Then, the predictions of the eleven target sites obtained with the same model type (UM, SM or EW) were pooled together, and the determination coefficient (R^2), RMSEP (RMSE of prediction), standard error of prediction (SEP), bias, ratio of performance to deviation (RPD) and ratio of performance to interquartile range (RPIQ) (Bellón-Maurel et al., 2010) were computed. Prior the computation of the above mentioned prediction performance parameters, the 88 samples used as spiking subsets were excluded. So, the prediction performance parameters were computed using 886 samples (974–88=886) from the eleven target sites. The same steps were followed with predictions obtained with UM, SM and EW models derived from the SL2, SL3, etc. A full description of models used in each target site is shown in the supplementary material (Appendix A).

Table 2. Main characteristics of each target site (TS)

Code	Site (Province/Country)	UTM [a]	Land use[b]	Parent material	Elevation (m a.s.l.)	Temp. (ºC)	Rainfall (mm)	Soil organic carbon (g C kg^{-1}) statistics[c]						
								n	Min	Max	Mean	SD	Skew	Kurt
TS1	Anna (Valencia/Spain)	30SYJ0121	A, F	Limestone	190	16.9	425	121	6.9	22.7	12.8	2.30	0.79	2.59
TS2	Sª Maigmó (Alicante/Spain)	30SYH0664	A, F	Marls	975	15.9	341	130	4.8	144.9	56.6	39.4	0.14	–1.06
TS3	Ricote (Murcia/Spain)	30SXH4022	S, F	Limestone & sandstone	310	16.5	330	155	28.8	141.5	65.3	25.5	0.53	–0.40
TS4	Sax (Alicante/Spain)	30SXH9067	F	Gypsum	470	16.1	315	95	4.6	40.3	18.0	7.1	0.89	1.11
TS5	Gérgal (Almería/Spain)	30SWG4210	S	Mica schist	1130	13.9	240	60	0.7	67.0	12.3	10.5	2.85	10.6
TS6	San Clemente (Cuenca/Spain)	30SWJ4864	A, S, F	Dolomite	695	14.3	450	55	5.7	72.8	17.5	11.3	2.39	8.30
TS7	Amusco (Palencia/Spain)	30TUM8171	A	Silt/clay & marls	710	12.5	500	56	3.5	43.0	12.0	7.78	1.97	4.07
TS8	Migueláñez (Segovia/Spain)	30TUL8649	A	Schists	900	11.2	550	55	2.8	27.4	8.1	4.22	1.94	5.81
TS9	Riello (León/Spain)	30TTN5541	F	Quartzite & gneiss	1140	9.8	1100	55	21.9	79.6	49.4	11.7	–0.20	0.12
TS10	Sª Orihuela (Alicante/Spain)	30SXH7920	A, F	Calcareous	200	18.1	295	88	3.2	78.1	28.6	18.0	0.28	–0.71
TS11	Silsoe (Bedfordshire/UK)	30UXC7765	A	Mudstone	64	9.9	594	104	12.1	34.1	22.0	5.9	0.14	–1.35
							ALL	974	0.7	144.9	3.41	31.2	1.31	0.94
							Without Spiking Subsets	886	0.7	141.5	3.47	31.5	1.25	0.72

[a] UTM coordinates (at 1 km^2 resolution)
[b] Land use: Agriculture (A), Forest (F), Shrubland (S)
[c] Min: minimum; Max: maximum; SD: standard deviation; Skew: Skewness; Kurt: Kurtosis

2.5. Laboratory analysis

All the analyses were carried out on air–dried and sieved (<2mm) samples. The SOC content (g C kg^{-1}) was analyzed with the Walkley–Black method, except the samples from Sweden and UK, where the loss on ignition (LOI) method was used. The diffuse reflectance NIR spectra of the Swedish samples were obtained with an ASD FieldSpec Pro Fr (ASD, Boulder, CO, USA) between 350 nm and 2500 nm. The samples from Spain and UK were scanned in a FT–NIR spectrophotometer (MPA, Bruker Optik GmbH, Ettlingen, Germany). The x–scale of the FT–NIR spectra was transformed from wavenumbers to wavelengths (834 nm to 2632 nm) and all the spectra were resampled to 1 nm. Both instruments share the spectral range comprised between 834 to 2500 nm, which was the range used when data from both instruments were needed.

3. Results and discussion

3.1. Predictions obtained with unspiked models

Predictions obtained with the unspiked models are shown in Figure 1 (left panels). In general, large errors were observed, except when the model was constructed with the spectral library SL1 (Fig. 1a), which is the largest–sized library. The RMSEP obtained with this model was 9.9 g C kg^{-1}, which corresponds with a RPD and a RPIQ of 3.17 and 4.08 respectively (Fig. 2). This RMSEP is slightly lower than the expected by the relationship observed between RMSEP and SD (Stenberg et al., 2010; Araújo et al., 2014). The unspiked models constructed with the other national–scale libraries (SL2 and SL3) were unable to predict the SOC with acceptable accuracy (Figs. 1d and 1g), being the RPD values below 1.59, mostly due to large bias (Fig 2), despite of high R^2 values (>0.81). The SOC predictions obtained with models constructed with the local–scale spectral libraries SL5 and SL7 were inaccurate (Figs. 1m and 1s). This result was expected because these small local–scale libraries did not contain samples with similar characteristics to those found in the target sites (Brown et al., 2005; Waiser et al., 2007). The spectral library SL5 only contains samples from an agricultural area located in Sjöstorp (Sweden), and the spectral library SL7 only contains soil samples from a forest located in Peñalén (Guadalajara, Spain), whereas the target sites are located in other provinces of Spain and UK, where the lithologies (parent material), climate, vegetation type, land use and soil type were different. The use of a different instrument can also contribute to explain the large errors obtained with the model derived from the Swedish library (SL5). The predictions obtained with the models derived from the provincial libraries SL4, SL6 and SL8 (Figs. 1j, 1p, 1v) were not accurate. This was also expected because most of the target sites were not located in that province. However, large errors were also obtained in soil samples from a target site which was located within that province. This is the case of the target site TS4, which is an area with gypsiferous soils. This soil type is underrepresented in that provincial spectral library, since gypsiferous soils only cover 4.9% of the total area in that province. The result indicates that location within the library's geographical limits *per se* is not a guarantee to obtain accurate predictions when downscaling (Sudduth and Hummel, 1996; Gogé et al. 2012; Kuang and Mouazen, 2013). The less abundant soil types would be underrepresented in the spectral library, especially when systematic sampling is used to construct the spectral library. The underrepresentation could be smaller if a stratified sampling is used to generate the library. However, it is almost unknown how to define or identify the strata to be sampled (Gogé et al. 2012; Grinand et al., 2012; Peng et al., 2013; Araújo et al., 2014; Shi et al., 2014). In general, these results were in agreement with those from Shepherd and Walsh (2002), who suggested the need of large spectral libraries for the construction of accurate models.

3.2. Predictions obtained with spiked models

As expected, the predictions obtained with the spiked models (Fig. 1, central panels) were more accurate than those obtained with the unspiked models (Fig. 1, left panels). Spiking ensures that the model contains samples with similar characteristics to those from the target site (McCarty and Reeves, 2001; Guerrero et al., 2010; Wetterlind and Stenberg, 2010). However, the magnitude of the improvement was inversely related with the size of the spectral library (Figs. 2 and 3). The higher the model's size was, the lower the impact of the spiking subset (~dilution). Indeed, the SOC predictions obtained before (UM) and after spiking (SM) the biggest–sized model were highly correlated (Pearson's r = 0.998; slope = 1.005; data not shown), indicating a very small effect due to spiking. On the other hand, the spiking effect on models derived from small–sized libraries was very high, since the eight samples added (i.e., spiking subset) represented a considerable proportion in the spiked model. As consequence, the predictions were drastically improved (Figs. 1 and 2). For example, when the model constructed with the Swedish local–scale library (SL5) was spiked, the prediction R^2 raised from 0.04 (Fig. 1m) to 0.94 (Fig. 1n), and the RPIQ increased in more than 4.8 units, reaching a value of 5.3 (Fig. 2). Similar improvements were observed on the other local–scale library (SL7; Fig. 1t), allowing to RPIQ to reach a value of 5.7 (Fig. 2). It is worth highlighting that, after spiking, the most accurate predictions were obtained with the models derived from the smaller–size libraries (Fig. 2).

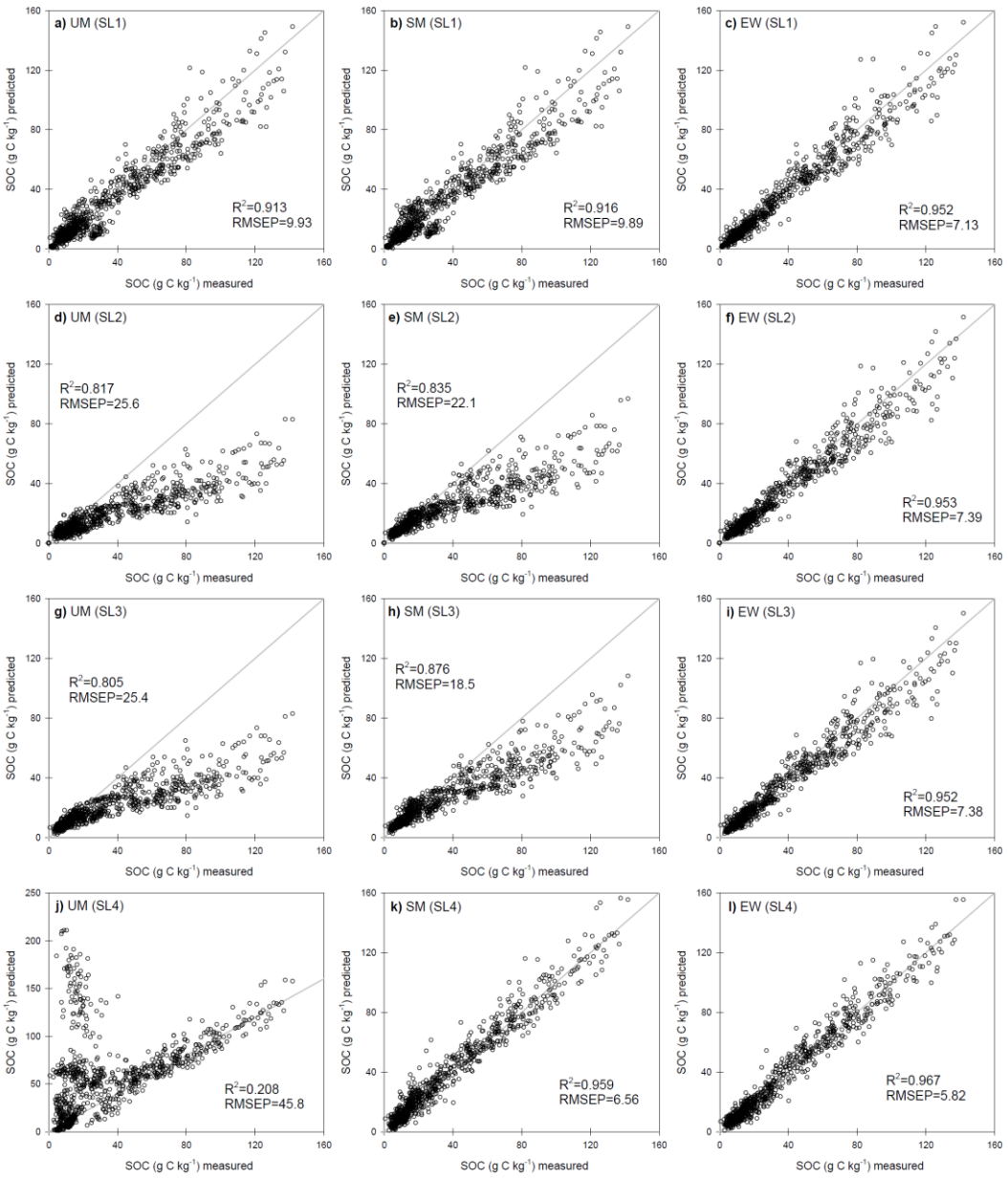

Figure 1. Scatter plots of predicted vs. measured soil organic carbon (SOC) contents in the 886 samples collected in eleven target sites. Panels on left: predictions obtained with unspiked models (UM). Central panels: predictions obtained with spiked models (SM). Panels on right: predictions obtained with spiked models where the spiking subset was extra–weighted (EW). Predictions obtained with models derived from the same spectral library (SL) are shown in the same row. The identification of the spectral library used is shown in brackets in each panel. The root mean square error of prediction (RMSEP) is given as g C kg^{-1}.

3.3. Predictions obtained with spiked models where the spiking subset was extra–weighted

The accuracy of the SOC predictions was additionally improved when the spiking subset was extra–weighted (Fig. 1, panels on right). All the predictions obtained were very accurate, regardless of the size of the spectral library (Fig. 2). The results clearly indicated the need to extra–weight the spiking subset when medium– and large–sized models are spiked (Fig. 3). The impact of the spiking subset is 'diluted' on large–sized models, but extra–weighting counterbalances the dilution effect. When the spiking subset is extra–weighted, the model is forced to fit preferentially to these samples. The spiking subset is composed by representative samples of the target site; consequently, the targeted or adapted model can provide accurate predictions for the rest of samples from that target site. Despite of the small effect of extra–weighting on small–sized models (Fig. 3), it is always recommended because it caused an additional improvement in the accuracy of the SOC predictions, and it does not imply any additional analytical effort respect to spiking. Indeed, the most accurate predictions were obtained when the adaptation approach was applied to the smallest–sized library (Figs. 2 and 3).

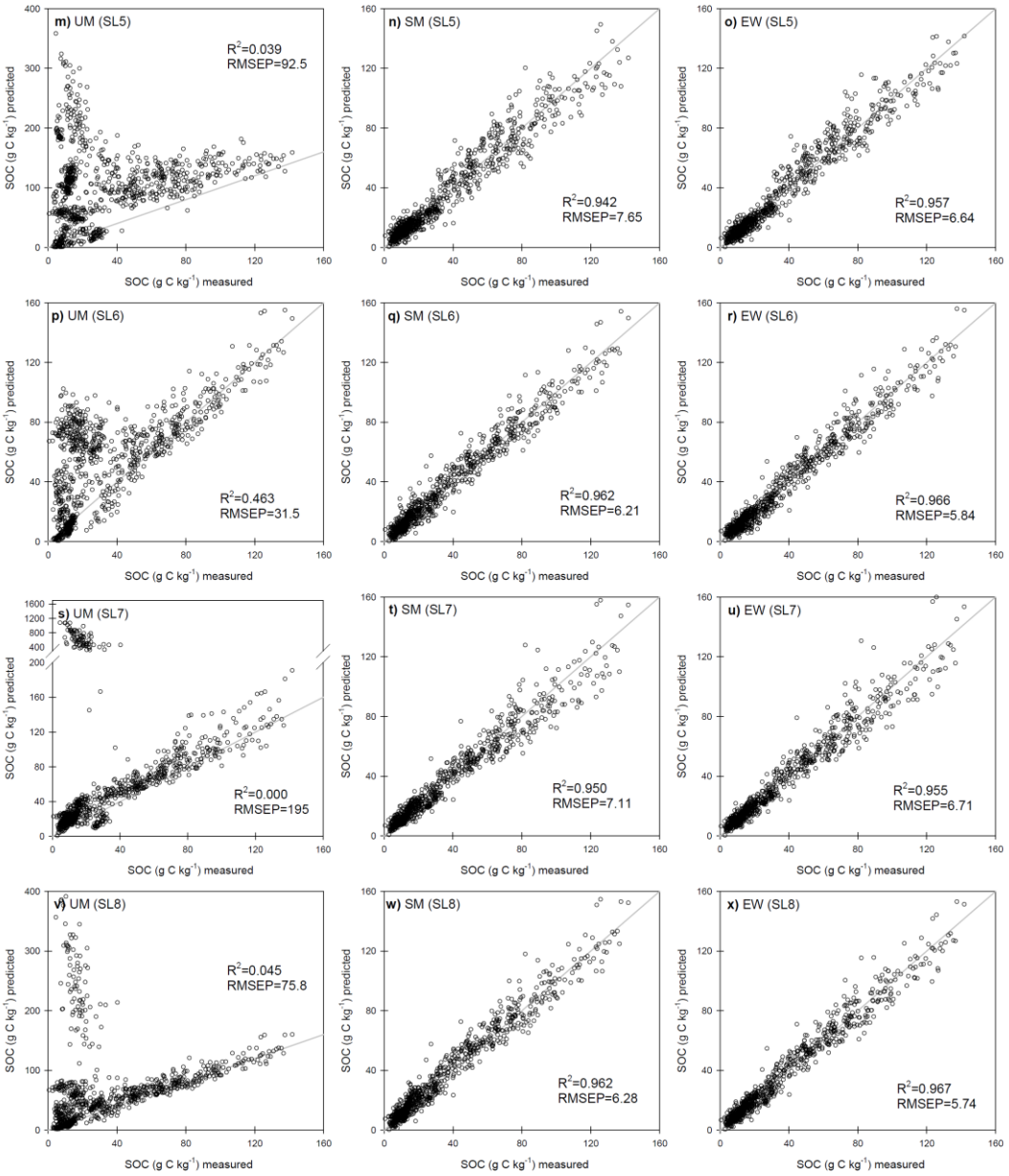

Figure 1. (continued)

3.4. Advantages of the model's adaptation by the extra-weighted spiking approach

The adaptation approach overcomes a number of potential limitations related with the applicability and use of models. First, it allows the use of provincial and national models to predict at local scale (i.e., 'downscaling') without bias, even in those sites with soils underrepresented in the spectral library. Secondly, when the approach was tested on the provincial libraries, the results indicated that the approach increased the model's geographical limits, allowing the successful downscaling to sites located outside the library's geographical limits. Furthermore, the approach also allows the model transfer to other sites, regardless if they were located in different provinces, or even in different countries. This is clearly shown when local–scale models were used (i.e., those derived from SL5 and SL7). The samples used in local–scale models and samples from target sites were differing in a wide range of characteristics: parent material, soil type, land use, climatic conditions and organic carbon contents. In addition, the spectra from the local–scale model derived from library SL5 were obtained with a different instrument than that used to scan the samples from the target sites. Summarizing, the approach allows the model's transferability to a wide range of conditions. Thirdly, the approach requires fewer efforts than the development of large spectral libraries, despite the need to make reference analyses on the spiking subset for each new target site considered. In this study, a total of 88 local samples were needed to spike the models (8 local samples × 11 target sites) with a very high accuracy.

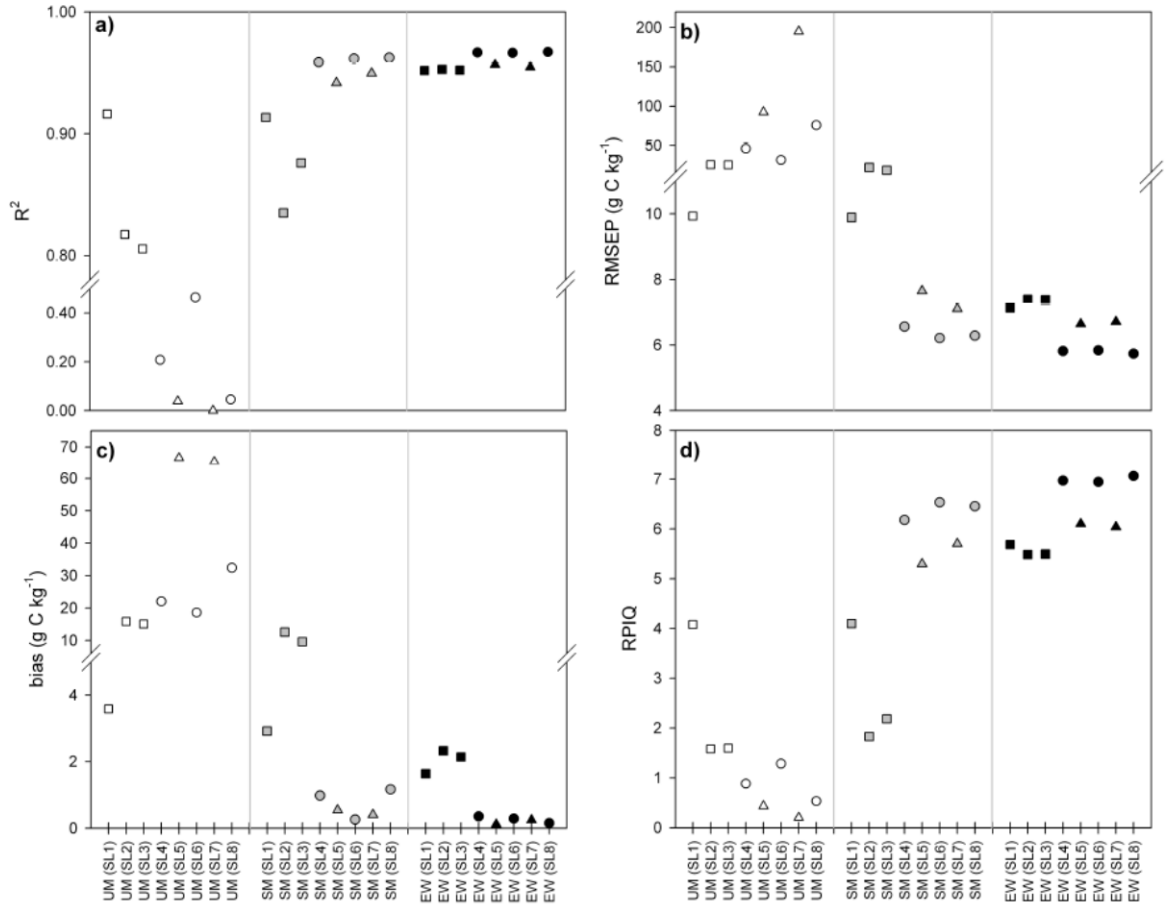

Figure 2. Values of the coefficient of determination (R^2), root mean square error of prediction (RMSEP), bias and ratio of performance to interquartile range (RPIQ) of predictions shown in Figure 1. White, grey and black symbols denote, respectively, predictions obtained with unspiked models (UM), spiked models (SM) and spiked models where the spiking subset was extra–weighted (EW). Squares, circles and triangles were used for national, provincial and local–scale spectral libraries (SL), respectively.

This small analytical effort was enough to predict the SOC in 886 samples (or 974 samples if we consider also the spiking subsets). The reference analysis of 88 samples (as spiking subsets) represents ca. ~10% of the total samples to be measured (974). When the smallest library was used (36 samples), the total number of samples analyzed with the reference method was only 124 samples (36+88). This number represents 20 times less effort than the needed for the development of the largest spectral library, composed by 3482 samples. In addition to the analytical efforts, the predictions obtained with the unspiked model derived from the largest SL (Figure 1a) were less accurate than the obtained with the extra–weighted spiked model derived from the smallest SL (Figure 1x). In comparison with the analytical efforts needed for the development of a national spectral library, composed by 3482 samples, our approach would allow the evaluation of more than 400 target sites, assuming that 8 local samples would be needed as spiking subset in each target site, and using a small–sized model of ~100 samples. For this reason, this approach is adequate for those studies were abundant and repeated measurements of SOC are needed at fine spatial resolution. Some examples of this type of studies are those needed to monitor changes in SOC stocks as consequence of land use conversions, agricultural and forest management, or those related with climatic change, where hundreds of SOC measurements per site are periodically needed. In addition, this type of studies implies replications in different sites. Therefore, an intensive monitoring in more than 400 target sites can provide a realistic view of changes in SOC stocks, especially if these monitoring sites are covering a wide range of representative units (ecosystems, land uses, etc.) across a country. Thus, once a small spectral library is available, this approach requires fewer samples with reference analysis than other approaches (Wetterlind et al., 2008, 2010; Mahmood et al., 2013; Debaene et al., 2014; Gogé et al., 2014; Ramirez–López et al., 2014). In this way, new NIR users would start to estimate SOC at local scale with simple and austere libraries. After the evaluation of several target sites, the spiking subsets can be added to the library for further development. In this way the creation of the spectral library is a dynamic process, but never a limiting issue. Once a library is large enough, other approaches can be followed, such the LOCAL ones (Gogé et al., 2012; Ramirez–López et al., 2013;

Nocita et al., 2014), where a large spectral library is needed. The LOCAL approach frequently outperforms the predictions obtained with a global approach (i.e., using a unique model from the whole spectral library) (Genot et al., 2011; Rabenarivo et al., 2013; Gogé et al., 2014). However, the predictions obtained with EW models are more accurate than those reported by other authors using the LOCAL approach, as Gogé et al. (2012, 2014), or Nocita et al. (2014) who used a very large European spectral library composed by ca. ~20000 samples (LUCAS database).

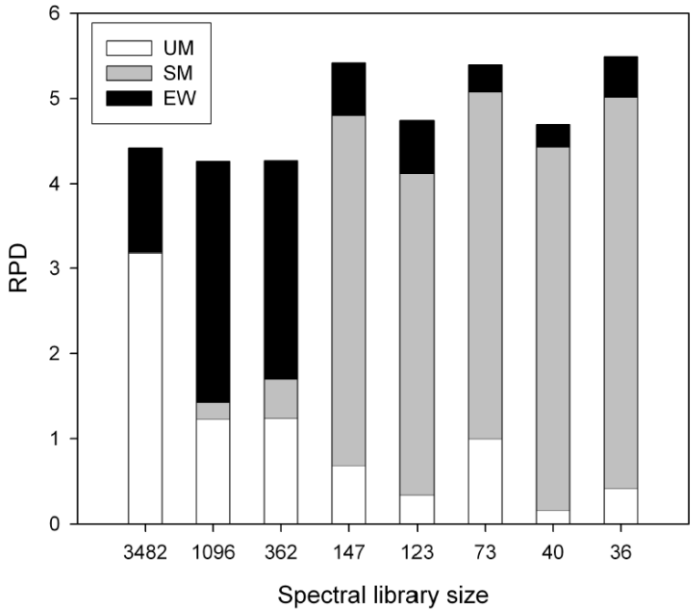

Figure 3. Values of ratio of performance to deviation (RPD) obtained in predictions shown in Figure 1. Data are not stacked.

Nevertheless, we consider unfair a simple comparison of RMSEP values with those reported in literature, especially with those studies where the ranges of the predicted variable are considerably different. The reason is that RMSEP values are affected by the range included (Bellon–Maurel et al., 2010; Kuang and Mouazen, 2011). This can be easily observed in the right part of Figure 4, where ranges were successively increased in a cumulative way, resulting in RMSEP values affected by the range. Probably the fairest comparison of models' accuracies is by comparing RMSEP values at fixed small intervals (classes) of the predicted variable (i.e., at specific ranges of SOC concentrations). Even a unique number cannot be used to summarize the overall accuracy, this is probably the unique useful way to provide RMSEP for further comparisons with other studies having different ranges of the target variable. These classes should be short (around 10 or 20 g C kg^{-1} SOC classes) otherwise the relative RMSEP will be influenced by the range. Nevertheless, excessively short classes are not advised, otherwise the number of observations per class can be low and scarcely informative. This would help to make fairer comparisons of the accuracy obtained by other researchers using different approaches, other algorithms, different spectral ranges, other instruments, etc. since it is practically improbable that other researchers will be using prediction sets with the same range. Fortunately, Nocita et al. (2014) used a LOCAL approach and provided the RMSEP values at arbitrary 15 g C kg^{-1} SOC classes. Concretely, they expressed these data as relative RMSEP values, just as percentage of the mean SOC of the class. Our relative RMSEP values (at 15 g C kg^{-1} SOC classes) obtained with the EW models derived from the smallest spectral library (SL8) were lower than those provided by Nocita et al. (2014) (Fig. 4). Thus, it is not reasonable to expect that the LOCAL approach will outperform the extra–weighting approach for local–scale SOC predictions. However, this should be confirmed in future studies.

The use of small libraries includes additional advantages, since the quality of the reference data of samples used for model's construction can be increased, because several analytical replicates can be afforded, which might be prohibitive if the size of the spectral library is large. Additionally, it minimizes the use of legacy samples and archived collections, which samples properties might have been affected by the storage and can contain errors in labels (the larger the library the higher the chance to contain errors). Moreover, the analysis of samples comprising large archives might have been carried out by different analysts, and probably also using different reference methods. All these sources of errors associated with large spectral libraries can negatively affect the models' prediction quality. However, this can be easily avoided in small–sized libraries. Furthermore, the use of small libraries may promote the sharing of analyzed samples among different NIR users.

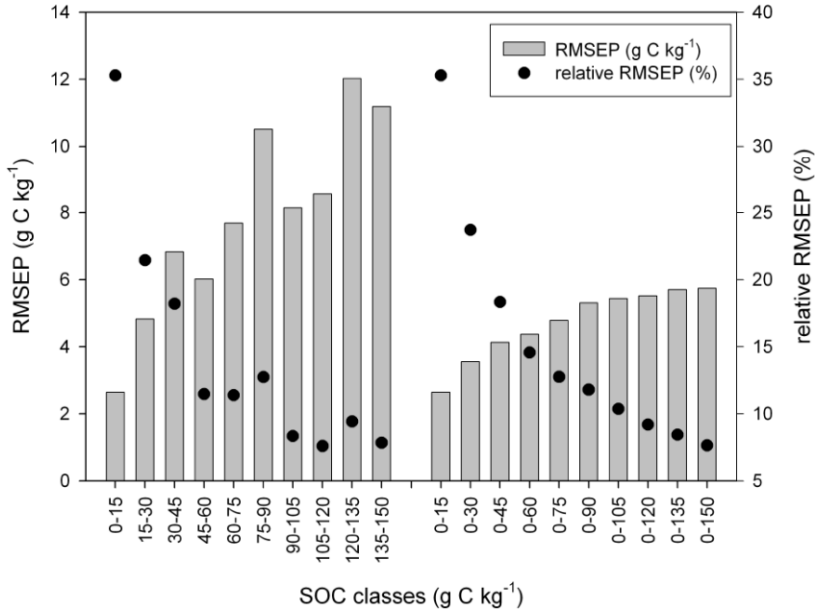

Figure 4. Values of the root mean square error of prediction (RMSEP) calculated at arbitrary classes of soil organic carbon (SOC). Grey bars denote values of RMSEP (g C kg^{-1}) obtained in each class (left axis). Black dots denote values of RMSEP expressed as percentage of the central value of the class (relative RMSEP) (right axis). Results shown in this figure correspond to data shown in Figure 1x.

It is important to mention that we are not in opposition with the development of large–sized spectral libraries, which are undoubtedly useful in soil science, especially for mapping purposes (Viscarra Rossel and Webster, 2012; Viscarra Rossel et al., 2014). However, large–spectral libraries are less effective for highly precise quantitative analysis of soil properties at local–scale, where the use of small–sized libraries adapted (or targeted) to the target site is clearly a more competitive approach. The adaptation approach always guarantees that the models contain relevant information of the target site, since the model is targeted. Thus, there are clear indicators that large spectral libraries are not needed for the assessment of SOC at local scale. However, this approach is not valid for other scales, where a unique sample is collected at each target site or the sampling density is very low, although these scenarios are not the ones where the NIR spectroscopy presents advantages. Large scale maps are useful to establish policies for land use and management. Even so, the implementation of policies and proper land use and management decisions require small scale (fine spatial resolution) maps. The development of smaller scale maps implies soil analysis in large number of samples collected at local scale, where the budget for that purposes is usually low. Therefore, NIR spectroscopy, as fast and cheap method, has an important role at that scale.

4. Conclusions

Models derived from very small–sized spectral libraries can provide accurate SOC predictions once they have been adapted to target site's characteristics, outperforming those predictions obtained with a model derived from a large spectral library. The results suggest that large spectral libraries might not be needed for local scale SOC assessment with NIR spectroscopy. The model adaptation approach does not allow the total replacement of the reference method. Even considering this disadvantage, the adaptation approach requires small efforts. The combined use of the adaptation approach and small–sized spectral libraries can increase the effectiveness of the NIR spectroscopy as rapid and cheap tool for proper land management.

Acknowledgements

This research project (Ref. CGL2011-27001) was financed by the Spanish Government Ministerio de Economía y Competitividad (Plan Nacional I+D+i, 2008–2011), and C. Guerrero gratefully acknowledges this financial support. C. Guerrero also acknowledges the Spanish Government Ministerio de Educación for a travel grant (ref. JC2011–0342). The Swedish part of the study was funded by the Swedish Farmers' Foundation for Agricultural Research and the Swedish Research Council Formas. Authors acknowledge to Fernando Tomás Maestre for provide some soil samples.

References

Araújo, S. R., Wetterlind, J., Demattê, J.A.M., Stenberg, B. 2014. Improving the prediction performance of a large tropical vis–NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. Eur. J. Soil Sci. 65, 718–729.

Bellon–Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M., McBratney, A. 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. Trac-Trends Anal. Chem.29, 1073–1081

Brown, D.J., Bricklemyer, R.S., Millar, P.R. 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. Geoderma 129, 251–267.
Debaene, G., Niedźwiecki, J., Pecio, A., Zurek, A. 2014. Effect of the number of calibration samples on the prediction of several soil properties at the farm–scale. Geoderma 214–215, 114–125.
Ge, Y., Morgan, C.L.S., Grunwald, S., Brown, D.J., Sarkhot, D.V. 2011. Comparison of soil reflectance spectra and calibration models obtained using multiple spectrometers. Geoderma 161, 202–21.
Genot, V., Colinet, G., Bock, L., Vanvyve, D., Reusen, Y., Dardenne, P. 2011. Near infrared reflectance spectroscopy for estimating soil characteristics valuable in the diagnosis of soil fertility. J. Near Infrared Spectrosc.19, 117–138.
Gogé, F., Gomez, C., Jolivet, C., Joffre, R. 2014. Which strategy is best to predict soil properties of a local site from a national Vis–NIR database? Geoderma 213, 1–9
Gogé, F., Joffre, R., Jolivet, C., Ross, I., Ranjard, L. 2012. Optimization criteria in sample selection step of local regression for quantitative analysis of large soil NIRS database. Chemometrics Intell. Lab. Syst. 110, 168–176.
Grinand, C., Barthès, B.G., Brunet, D., Kouakoua, E., Arrouays, D., Jolivet, C. Caria, G., Bernoux, M. 2012. Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS). Eur. J. Soil Sci. 63, 141–151.
Guerrero, C., Stenberg, B., Wetterlind, J., Viscarra Rossel, R.A., Maestre, F.T., Mouazen, A.M., Zornoza, R., Ruiz–Sinoga, J.D., Kuang, B. 2014. Assessment of soil organic carbon at local scale with spiked NIR calibrations: effects of selection and extra–weighting on the spiking subset. Eur. J. Soil Sci. 65, 248–263.
Guerrero, C., Zornoza, R., Gómez, I., Mataix–Beneyto, J. 2010. Spiking of NIR regional models using samples from target sites: Effect of model size on prediction accuracy. Geoderma 158, 66–77.
Kennard, R.W., Stone, L.A. 1969. Computer aided design of experiments. Technometrics 11, 137–148.
Kuang, B., Mouazen, A.M., 2011. Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. Eur. J. Soil Sci. 62, 629–636.
Kuang, B., Mouazen, A.M. 2013. Effect of spiking strategy and ratio on calibration of on-line visible and near infrared soil sensor for measurement in European farms. Soil Till. Res. 128, 125–136.
Mahmood, H.S., Bartholomeus, H.M., Hoogmoed, W.B., van Henten, E.J. 2013. Evaluation and implementation of vis–NIR spectroscopy models to determine workability. Soil Till. Res. 134, 172–179
Mäkipää, R., Häkkinen, M., Muukkonen, P., Peltoniemi, M. 2008. The costs of monitoring changes in soil carbon stocks. Boreal Environ. Res. 13, 120–130.
McCarty, G.W., Reeves III, J.B. 2001. Development of rapid instrumental methods for measuring soil organic carbon, in: Lal, R., Kimble, J.M., Follett, R.F., Stewart, B.A. (Eds.), Assessment methods for soil carbon, Advances in Soil Science. Lewis Publishers, Boca Raton, FL, pp 371-380.
Muukkonen, P., Häkkinen, M., Mäkipää, R. 2009. Spatial variation in soil carbon in the organic layer of managed boreal forest soil – implications for sampling design. Environ. Monit. Assess. 158, 67–76.
Nocita, M., Stevens, A., Toth, G., Panagos, P., Van Wesemael, B., Montanarella, L. 2014. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. Soil Biol. Biochem. 68, 337–347.
Peng, Y. Knadel, M., Gislum, R., Deng, F., Norgaard, T., Wollesen de Jonge, L., Moldrup, P., Humlekrog Greve, M. 2013. Predicting soil organic carbon at field scale using a national soil spectral library. J. Near Infrared Spectrosc. 21, 213–222. doi: 10.1255/jnirs.1053
Rabenarivo, M., Chapuis-Lardy, L., Brunet, D., Chotte, J.-L, Rabeharisoa, L., Barthès, B. 2013. Comparing near and mid-infrared reflectance spectroscopy for determining properties of Malagasy soils, using global or LOCAL calibration. J. Near Infrared Spectrosc. 21, 495–509
Ramirez–Lopez, L., Behrens, T., Schmidt, K., Stevens, A., Demattê, J.A.M., Scholten, T. 2013. The spectrum–based learner: A new local approach for modeling soil vis–NIR spectra of complex datasets. Geoderma 195–196, 268–279
Ramirez–Lopez, L., Schmidt, K., Behrens, T., van Wesemael, B., Demattêe, J.A.M., Scholten, T. 2014. Sampling optimal calibration sets in soil infrared spectroscopy. Geoderma 226–227, 140–150
Shepherd, K.D., Walsh, M.G. 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Sci. Soc. Am. J. 66, 988–998.
Shi, Z., Wang, Q., Peng, J., Ji, W., Liu, H, Li, X., Viscarra Rossel, R.A. 2014. Development of a national VNIR soil–spectral library for soil classification and prediction of organic matter concentrations. Science China Earth Sciences 57, 1671–1680
Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J. 2010. Visible and near infrared spectroscopy in soil science. Adv. Agron. 107, 163–215.
Sudduth, K.A., Hummel, J.W. 1996. Geographic operating range evaluation of a NIR soil sensor. Trans. ASAE 39, 1599–1604.

Viscarra Rossel, R. A., Webster, R. 2012. Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. Eur. J. Soil Sci. 63, 848–860.
Viscarra Rossel, R.A., Webster, R., Bui, E.N., Baldock, J.A. 2014. Baseline map of organic carbon in Australian soil to support national carbon accounting and monitoring under climate change. Global Change Biol. 20, 2953–2970.
Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T. 2007. In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. Soil Sci. Soc. Am. J. 71, 389-396.
Wetterlind, J., Stenberg, B., Söderström, M. 2008. The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. Precis. Agric. 9, 57–69.
Wetterlind, J., Stenberg, B. 2010. Near–infrared spectroscopy for within–field soil characterization: small local calibrations compared with national libraries spiked with local samples. Eur. J. Soil Sci. 61, 823–843.
Wetterlind, J., Stenberg, B., Söderström, M. 2010. Increased sample point density in farm soil mapping by local calibration of visible and near infrared prediction models. Geoderma 156, 152–160.