Using NIRS for analysis of soil clay content

Mariage[1] C., Genot[2] V., Baeten[3] V., Dardenne[3] P., Colinet[1] G.

[1] **Liege University – Gembloux Agro-Bio Tech – Soil-Plant Exchanges Axis (Belgium)**

clemence.mariage@uliege.be

[2] **Analytical Laboratory SPAA, Tinlot-Scry (Belgium)**

[3] **Walloon Agricultural Research Centre (CRA-W) - Valorisation of Agricultural Products Department, Quality and authentication of agricultural products Unit (Belgium)**

Avec le soutien de la Wallonie

Context

REQUASUD network (based in Wallonia, Belgium) consists of analytical laboratories working directly with farmers and citizens, giving them soil fertility diagnosis and advice for a good management of soils and cultures (Figure 1). The network is supervised by a research laboratory allowing research development benefit for the network.

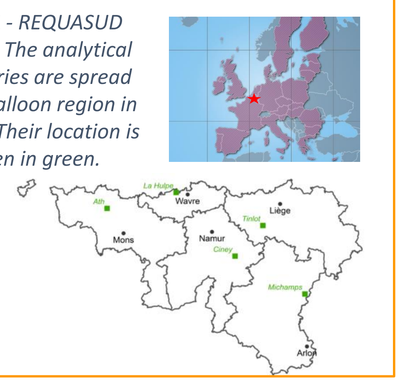


Figure 1 - REQUASUD network. The analytical laboratories are spread on the Walloon region in Belgium. Their location is written in green.

To correctly interpret the analytical results and give relevant fertility advice, the laboratories need, among others, information about the texture of the soil. Analyzing clay content of the soil with reference methods (like ISO 11277) is time-consuming. To overpass this problem, the near-infrared spectroscopy (NIRS) has been developed and is now used routinely in the laboratories of REQUASUD network. The main advantages of this method are:

- rapid & low-cost
- reproducible
- easy to proceed
- non-destructive

The NIRS principle

NIRS is an analytical technique that characterizes materials according to their reflectance in the wavelengths ranging between 800 and 2500 nm. A light source emit a radiance that is received by the sample. This light react with the sample and is then (i) transmitted through the sample, (ii) absorbed by the sample and (iii) reflected by the sample. The way the light is divided into these 3 parts depends on the molecules that constitute the sample. The reflected light is caught by a light receptor. The reflectance (R) or absorbance (log 1/R) is put in graph as a function of wavelengths. This graph is the spectral signature of the sample (Figure 2).

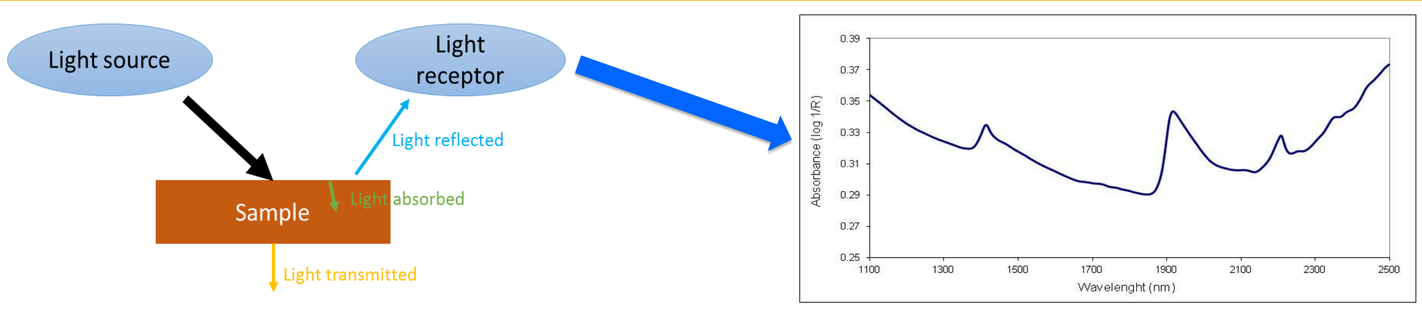


Figure 2 - The NIRS principle and the spectral signature of a sample

What you need to use the NIRS method routinely in a lab

1. An analytical protocol and a spectrometer

The spectral signature is influenced, among others, by soil humidity, particles size and obviously by the laboratory or the analyst.

Different tests have been performed to find the better analytical protocol, easy to proceed and repeatable (Genot et al., 2008).

We work with 2 mm sieved, dried samples (following ISO 11464) placed in a quarter cup. Each sample is scanned in duplicate. The 2 spectra are compared calculating the RMS (Root Mean Square):

- RMS ≤ 10,000: the measurement is accepted and the spectra are averaged;
- RMS > 10,000: the measurement needs to be done again.

REQUASUD network work with Foss XDS instruments.

2. An initial spectral library

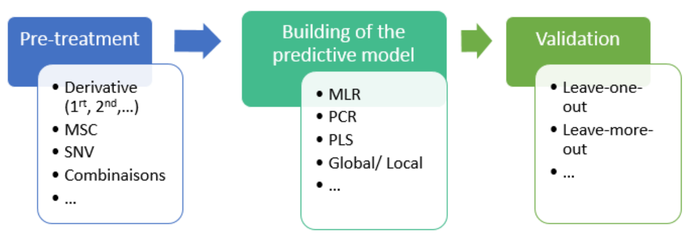
As many other analytical methods, NIRS don't allow to obtain directly the wanted value (clay content of the soil); the spectral signature needs to be interpreted.

To build the interpretation model, a spectral library is needed. It contains clay content determined with the reference method (ISO 11277) *and* the associated spectrum (couple REF value – NIR spectrum). The main characteristics of the library are (for the studied territory – Wallonia):

- Representativeness of the clay content;
- Representativeness of the land uses (cultures, grasslands, forests, etc);
- Representativeness of the spectral diversity.

3. Model to predict the value from the spectrum

To build the interpretation model, 3 steps can be distinguished: pre-treatment, building of the predictive model and validation of the model. For each step, different methods exist:



These different methods have been tested to implement the NIRS method in REQUASUD network (Genot et al., 2008). The choices made for clay content are:

Pre-treatment: 1rt derivative. The pre-treatment allows to get a better view on the spectrum, get rid of the background noise, etc

Predictive model: local PLS. The PLS method consists of replacing p correlated variables by k non-correlated new variables (the PLS factors). The PLS factors are then combined to write a predictive equation. The "local" approach means that, for each unknown sample to be predicted, the closest samples (from spectroscopy point of view) are selected in the library, used to define the PLS factors and write a predictive equation. As the closest samples are different for each unknown sample, the predictive equation is also different.

4. Grow the spectral library

Growing the initial spectral library and its representativeness's (clay content, land uses and spectra), leads to strengthen the predictive model, a better accuracy of the predicted values and a higher amount of correctly predicted samples.

The contributors are:

- the analytical laboratories of REQUASUD network: spread on the Walloon region, they are the most able to feed the library on an efficient way;
- the European LUCAS spectral library, for croplands and grasslands uses (survey implemented in 2009).

These contributors work with the same reference method (ISO 11277) and the same equipment. That way, the so built spectral library is homogeneous and usable by every lab of the network, using the same predictive model.

Nowadays, the spectral library is so big that it has been divided into 3 parts according to land use: croplands, grasslands et vegetable gardens.

Here are the current basic statistics of the REQUASUD spectral library for clay content:

	Croplands	Grasslands	Vegetable gardens
Amount of samples	9,603	4,653	273
Min value (%)	1.00	0.20	4.04
Max value (%)	79.00	49.00	39.03
Mean value (%)	21.94	18.84	17.02
Standard deviation (%)	12.45	10.30	5.67

Routine use

The analytical laboratories use the defined analytical protocol, a Foss XDS and the ISIScan software associated with RINA or Mosaïc platform. The lab choose the land use (croplands, grasslands or vegetable gardens) and the software gives the predicted value and 2 statistical parameters indicating the reliability of the predicted value:

1) GH (global H or Mahalanobis distance) is the distance between the unknown sample (to be predicted) and the gravity centre of the closest samples (from spectroscopy point of view) selected in the library;
2) NH (neighbor distance) is the distance between the unknown sample (to be predicted) and the closest of the closest samples (from spectroscopy point of view) selected in the library.

GH ≤ 3 and NH ≤ 1	→ the predicted value is validated as reliable
GH > 3 and/or NH > 1	→ the predicted value is not reliable enough; analyze the sample with REF method, add the couple REF value – NIR spectrum to the spectral library

A "Repeatability file" is used when doing the predictions. It contains the spectra of same samples obtained on the equipment of the different labs of the network. This file strengthen the predictions regarding modifications of temperature, differences between the equipment in the different labs, etc.

Performances

Here are the current values for the statistical parameters indicating the performances of a predictive model, for the REQUASUD spectral library and clay content:

	Croplands	Grasslands	Vegetable gardens	
% GH ≤ 3	96.22	95.79	89.74	
% NH ≤ 1	85.93	84.05	81.68	
R^2	0.80	0.67	0.42	determination coefficient between the REF values and the predicted values
SEP (standard error of prediction) (%)	5.62	5.99	4.34	$\sqrt{\frac{\sum(\text{predicted value}-\text{REF value})^2}{\text{Total amount of samples}}}$
RPD	2.22	1.72	1.31	$\frac{\text{standard deviation (REF method)}}{SEP}$

The amount of correctly predicted samples is high (more than 80 %). RPD values show a performant model for croplands and a medium model for grasslands. The performances of the model can be improved for vegetable gardens. This can happen by adding couples REF value – NIR spectrum in the spectral library.

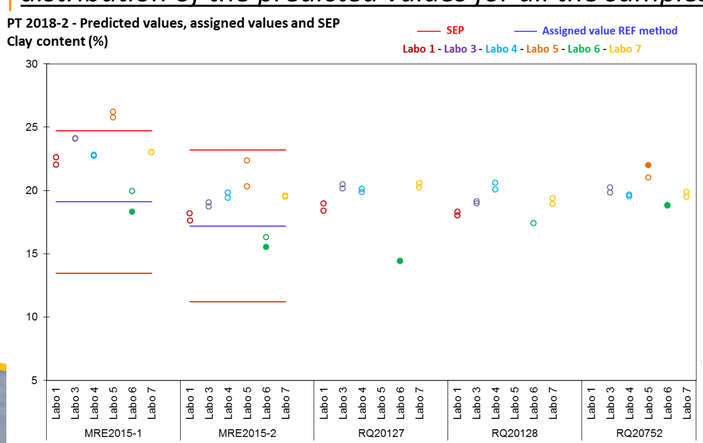
Control performances of the laboratories

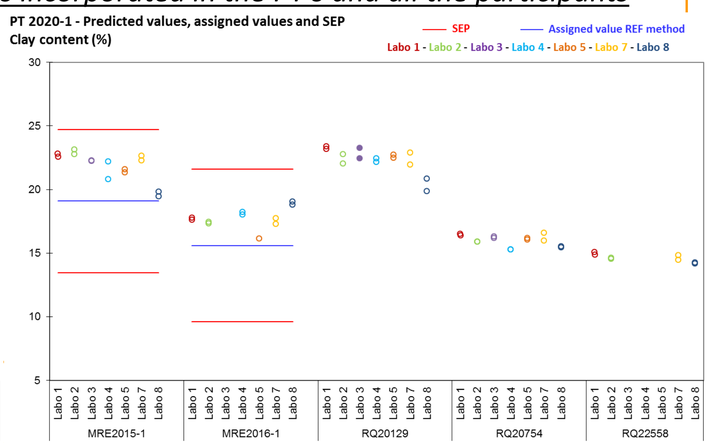
PT's (Performance Testing's) are organized within REQUASUD network (2 per year), and NIRS has been implemented recently. Five samples are sent to the participants, each sample in duplicate so the repeatability can be assessed. Some samples are incorporated into several following PT's, allowing the participants to better follow their performances over time.

Maximum difference between the predicted values of all the participants on identical samples, over time

	PT2018-2		PT2019-1		PT2019-2		PT2020-1	
	Max diff	Nbr of participants	Max diff	Nbr of participants	Max diff	Nbr of participants	Max diff	Nbr of participants
MRE2015-1	7.89	7	3.00	6	1.67	6	3.68	7
MRE2016-1			2.28	6	2.28	6	2.88	6

Difference between the predicted values and the assigned REF value; visual representation of the distribution of the predicted values for all the samples incorporated in the PT's and all the participants





- The first PT (2018-2) allowed the participants to make great progress. The narrower distribution of the predicted values for 2019 and 2020 PT's is also thanks to the "Repeatability file" that has been updated meanwhile. Since 2019, the results are quite stable, but the predicted values become less and less dependent of the lab.
- The laboratories improved their repeatability.
- The difference between the predicted value and the assigned REF value is lower than SEP. The performances of the laboratories are thus rather good.
- The laboratories still need to improve their performances (working for example on analysts' ability).
- The spectral library needs to be fed with more diverse couples REF value – NIR spectrum, especially for other soil covers than croplands (grasslands, vegetable gardens, orchards,…).

N.B. 1) A same laboratory may be represented by a different number over time. These graphs don't show the evolution of a lab's performances in particular.
N.B. 2) A full dot means that GH > 3 and/or NH > 1. The predicted value is not validated and is given for information.
N.B. 3) The assigned values (REF method) is determined by an accredited body (External Reference Material).

References

- *Genot V et al., 2008. Study of the near infrared reflectance spectroscopy performances for the determination of soil parameters useful for the fertility diagnosis. European Geosciences Union. Vienna, Austria, 13 – 18 April 2008.*
- *LUCAS database : https://ec.europa.eu/eurostat/fr/web/lucas/overview*