

# APPLICATION DE LA SPECTROMETRIE PROCHE-INFRAROUGE A L'ANALYSE DE TERRE

Genot Valérie<sup>1</sup>, Colinet Gilles<sup>1</sup>, Bock Laurent<sup>1</sup>, Vanvyve Dominique<sup>2</sup> et Dardenne Pierre<sup>3</sup>

1. Université de Liège – Gembloux Agro-Bio Tech – Unité de Science du Sol – 2, Passage des Déportés – B-5030 Gembloux – [vgenot@ulg.ac.be](mailto:vgenot@ulg.ac.be)

2. ASBL CPL-PROMOGEST - Laboratoires de la Province de Liège – Rue de Dinant, 110 – B 4557 Tinlot-Scry – [spaa@provincedeliege.be](mailto:spaa@provincedeliege.be)

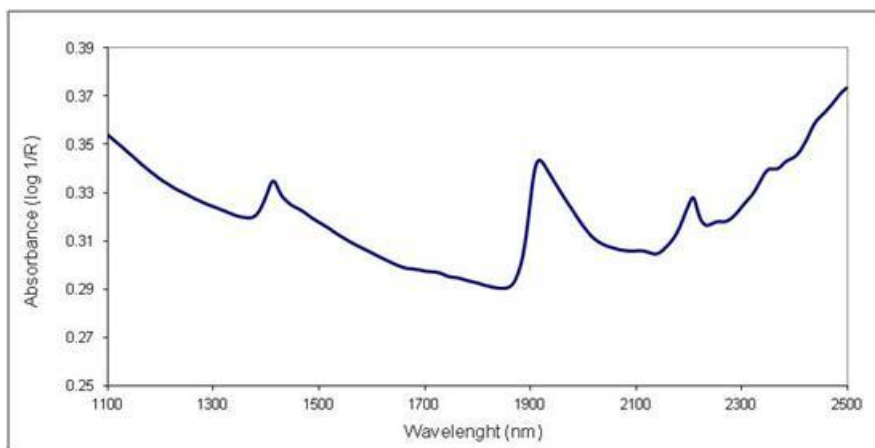
3. Centre Wallon de Recherches agronomiques – Département Valorisation des Productions agricoles – 24, Chaussée de Namur – B-5030 Gembloux – [dardenne@cra.wallonie.be](mailto:dardenne@cra.wallonie.be)

## Contexte

Cette étude s'intègre dans le développement d'un outil d'aide à la décision pour le diagnostic agronomique et le conseil de fumure personnalisé en Wallonie. Il s'agit en effet de répondre aux besoins des laboratoires d'analyses de terre membres de l'ASBL REQUASUD ([www.requasud.be](http://www.requasud.be)) qui souhaitent proposer aux agriculteurs un conseil de fumure précis leur permettant de mieux maîtriser les intrants agricoles dans le cadre d'une gestion durable des sols. Ces laboratoires sont encadrés par une institution scientifique, l'Unité de Science du Sol de Gembloux Agro-Bio Tech (Université de Liège) qui a pour objectifs de (i) veiller à une harmonisation des protocoles de l'échantillonnage jusqu'au conseil, (ii) de suivre la qualité des analyses et de (iii) développer de nouveaux outils permettant d'améliorer les différentes étapes menant à l'élaboration du diagnostic agronomique et du conseil de fumure.

Dans ce réseau, le constat a été fait que certaines propriétés de sol soit sont rarement analysées en routine, malgré leur importance pour un conseil pertinent, soit consomment des extractifs nocifs. Des solutions alternatives ont donc été cherchées pour remplacer ces méthodes analytiques coûteuses et donc peu demandées par les agriculteurs. Ainsi, un paramètre aussi important que la capacité d'échange cationique (CEC) n'est demandé qu'une fois sur 1 000 en moyenne par les agriculteurs. Au vu de ces considérations, des recherches ont été menées pour étudier la faisabilité d'une application en routine de la spectroscopie dans le domaine du proche infrarouge pour déterminer ces paramètres. La spectroscopie proche infrarouge (SPIR ou NIRS pour near infrared spectroscopy) offre des perspectives intéressantes pour répondre au besoin d'une meilleure connaissance du potentiel de fertilité des sols. En effet, cette technique ne nécessite qu'une préparation réduite de l'échantillon et offre une grande rapidité de mesure, un faible besoin en matériel, une formation minimale des manipulateurs et permet la détermination simultanée de plusieurs composants (Russel et al., 2002). De plus, aucun extractif chimique n'est utilisé. Enfin, la mesure de certaines propriétés chimiques ou biologiques des sols non analysées actuellement car trop coûteuses malgré leur intérêt dans l'évaluation de la fertilité du sol est envisageable en spectroscopie proche infrarouge (Stenberg et al., 2010).

La SPIR est une technique analytique basée sur le principe d'absorption des rayonnements infrarouge par la matière. Cette absorption étant liée à la composition chimique de la matière, il est possible d'estimer celle-ci en mesurant simplement l'absorption du rayonnement par l'échantillon (figure 1). Cette technique a tardé à être utilisée en analyse de sol car les liens fondamentaux entre les mesures chimiques et les propriétés du sol sont complexes. Il est donc difficile de mettre directement en relation une caractéristique physique, chimique ou physico-chimique du sol avec une empreinte spectrale ou une longueur d'onde particulière (Shepherd et Walsh, 2002). C'est ainsi que le spectre est le plus souvent traité dans son intégralité en utilisant des méthodes chimiométriques adaptés comme, par exemple, la régression selon les moindres carrés partiels (PLS).



**Figure 1** : Exemple d'un spectre infrarouge d'un échantillon de terre séchée, émottée et tamisée à 2 mm

L'objectif de l'étude est donc de mettre en place une méthodologie pour la détermination par spectroscopie proche infrarouge du taux de carbone organique total (COT), d'azote total (NT), d'argile et de la capacité d'échange cationique (CEC) d'échantillons de terre rencontrés en Wallonie. La stratégie a consisté à (i) comparer différentes méthodes de prétraitements et de calibrage basés sur la régression PLS global et local, (ii) évaluer la répétabilité et la reproductibilité de la méthode, (iii) transférer la base de données spectrales créée dans les laboratoires membres du réseau, (iv) utiliser la technique dans le cadre d'un travail de routine, tout en améliorant en continu la qualité des prédictions par l'ajout de nouveaux spectres à la base de données de référence.

Le détail de la stratégie mise en place ainsi que les résultats obtenus au niveau des trois premiers axes ont été présentés par Genot et al. (2011).

## Méthodologie

Les différentes étapes de la méthodologie sont :

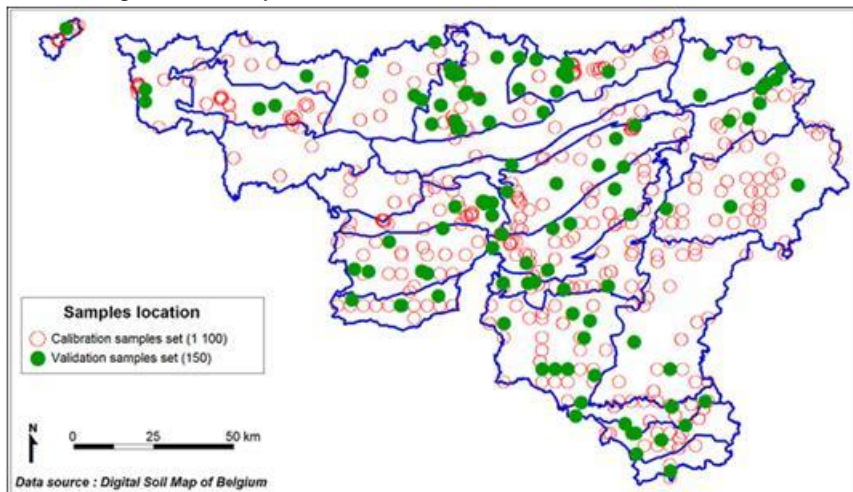
1. Evaluer la diversité de la population pour les quatre paramètres étudiés ;
2. Sélectionner un set d'échantillons de terre représentatifs de cette diversité ;
3. Analyser les échantillons selon les procédures standards ;
4. Définir une procédure de préparation des échantillons à appliquer pour la phase de calibrage et pour le travail de routine ;
5. Scanner en double l'ensemble des échantillons selon cette procédure ;
6. Elaborer des modèles prédictifs pour chaque paramètre sur base d'une comparaison entre différents prétraitements des spectres et des régressions PLS ;
7. Evaluer la répétabilité et la fidélité intermédiaire de la méthode ;
8. Transférer la base de données et les modèles élaborés dans les différents laboratoires ;
9. Prédire en routine les échantillons de terre, implémenter la base de données spectrales et améliorer en continu la qualité des prédictions.

Sur base de la Carte Numérique des Sols de Wallonie et des bases de données analytiques existantes, une étude de la diversité des sols rencontrés sur le territoire wallon et de leurs caractéristiques analytiques a été réalisée (Genot, 2006). Sur cette base, mille-trois cent échantillons de terre, représentatifs de la diversité géopédologique, ont été prélevés en Wallonie, séchés, émottés et tamisés à 2 mm, analysés selon les méthodes de référence (Metson pour la capacité d'échange cationique, Springler-Klee pour le carbone organique total, Kjeldhal pour l'azote total et De Leenheer pour le pourcentage d'argile).

Ensuite, la procédure pour scanner les échantillons a été définie. Les objectifs étaient de mettre au point une méthodologie robuste, indépendante de la nature de l'échantillon (par exemple de sa granulométrie), de l'opérateur et reproductible dans le temps. Pour rester au plus près du travail de routine et après étude bibliographique, il a été décidé de travailler avec les échantillons séchés,

émottés et tamisés à 2 mm, de manière à ne pas ajouter une étape de broyage en laboratoire. Ensuite, plusieurs techniques de remplissage des coupelles ont été testées (cf. Genot et al., 2011). Une fois la méthode de remplissage des coupelles définie, l'ensemble des échantillons ont été scannés en double sur un FOSS NIR SYSTEM 5000 couvrant une gamme de longueur d'onde allant de 1 100 nm à 2 500 nm.

La base de données spectrales enrichies des résultats analytiques selon les méthodes de référence a ensuite été divisée en trois lots : un premier lot servant à paramétrer l'étape de calibrage en testant les différents modèles ; un second lot permettant de valider le modèle retenu et un troisième lot constituant la base de données de calibrage. La figure 2 localise les échantillons faisant partie du lot de calibrage et ceux ayant servi à la validation des modèles.



**Figure 2** : Localisation des échantillons constitutifs des sets de calibrage et de validation (source : Genot et al., 2011).

Des modèles de prédiction ont été comparés, basés sur l'application de divers prétraitements des spectres et sur l'utilisation de la régression selon les moindres carrés partiels. Ensuite, une approche en mode local, utilisant également le principe de la régression PLS a été testée et comparée aux modèles obtenus selon la régression classique. La régression locale utilise le critère de similarité entre les spectres (basé sur leur corrélation) pour sélectionner les spectres les plus proches de l'échantillon à prédire et établir une équation de régression uniquement sur base de ces spectres sélectionnés. Contrairement à la régression PLS classique ou global, il n'y a pas une équation obtenue au terme de la phase de calibrage mais une base de données spectrale avec des critères bien définis pour sélectionner les spectres et développer une nouvelle équation pour chaque nouvel échantillon à prédire. Cette technique permet notamment de tenir compte de la non-linéarité de la relation entre la réflectance aux différentes longueurs d'onde et la propriété du sol à prédire.

Les critères retenus pour évaluer la qualité des modèles établis sont notamment : l'erreur standard de validation croisée (SECV), l'erreur standard de prédiction (SEP), le coefficient de corrélation ( $r^2$ ), le RPD qui représente le rapport entre l'écart-type de la population et le SEP ainsi que le coefficient de variation (CV). Le RPD permet de qualifier les modèles obtenus en considérant trois catégories principales (Chang et al., 2001) : (1)  $RPD \geq 2$  signifiant que le paramètre étudié est bien prédit par la technique, (2)  $1,4 \leq RPD < 2$  signifiant qu'il est possible de prédire cette propriété du sol par SPIR moyennant une amélioration des modèles et (3)  $RPD < 1,4$  signifiant que le modèle prédit mal le paramètre.

Une fois les modèles obtenus, la répétabilité et la fidélité intermédiaire de cette technique analytique ont été évalués (Genot et al., 2011).

Ensuite, la base de données spectrales a été transférée pour être compatible avec les FOSS XDS utilisés par les laboratoires du réseau. Cette base de données est centralisée au niveau du Centre Wallon de Recherche agronomique et les laboratoires, travaillant en mode local, y sont connectés en permanence. Ils peuvent ainsi, en routine, prédire les quatre paramètres étudiés sur les échantillons de terre provenant des sols agricoles wallons.

Avant de passer à une utilisation en routine de la technique, les laboratoires ont systématiquement scanné et analysé leurs échantillons selon les méthodes analytiques classiques. Les échantillons présentant les spectres les plus différents de la base de données spectrales sont sélectionnés et analysés pour tous les paramètres selon les procédures de référence (pouvant différer des procédures utilisées en routine dans les laboratoires). La base de données est ainsi enrichie par de nouvelles données permettant d'améliorer la qualité de la prédiction.

La dernière étape consiste à utiliser la technique en routine, c'est-à-dire, à préalablement scanner tous les échantillons et à ne plus analyser selon les techniques traditionnelles que ceux qui ne sont pas correctement prédit par le modèle.

## Résultats et discussion

### Choix et validation des modèles de prédiction

Le tableau 1 synthétise les résultats pour les trois modèles sélectionnés après application de divers prétraitements et paramétrages des modèles locaux. Au total 15 prétraitements des spectres et 5 300 combinaisons de paramètres en régression local ont été traités pour choisir les modèles qui permettaient les prédictions les plus précises.

Les résultats présentés ici sont ceux obtenus sur base d'une soixantaine d'échantillons présents dans le lot de validation. Ces échantillons ont été retenus car ils sont prédits par les trois modèles et ils permettent donc de comparer la qualité de la prédiction. Le premier modèle local (N fixe) est basé sur un nombre fixe de spectres retenus, c.à.d. que pour tout échantillon à prédire, le modèle sélectionne dans la base les  $x$  spectres les plus proches par ordre décroissant de corrélation entre les spectres. Le second modèle est basé sur un coefficient de corrélation fixe, ce qui signifie que pour prédire un nouvel échantillon, le modèle sélectionne dans la base uniquement les spectres dont le coefficient de corrélation est supérieur à la limite fixée. Si le nombre d'échantillons répondant à ce critère est inférieur à un seuil minimum, l'échantillon n'est pas prédit.

Les résultats, tant en terme de SEP que de RPD, montrent que la qualité de la prédiction est nettement meilleure pour les modèles locaux et que le fait de travailler sur la corrélation entre les spectres améliore également la qualité de prédiction.

**Tableau 1** : Résultats obtenus pour les trois modèles de régression PLS global et locaux retenus au terme de la phase de calibrage sur base d'une sélection d'échantillons du lot de validation.

Paramètres	N = 60	Modèle global	Modèle local	
			N fixe	r <sup>2</sup> fixe à 0.99
COT (g 100 g <sup>-1</sup> )	SEP	0.50	0.20	0.13
	RPD	1.58	3.95	6.08
NT (g kg <sup>-1</sup> )	SEP	0.24	0.19	0.08
	RPD	0.83	1.05	2.5
Argile (%)	SEP	4.55	3.8	1.82
	RPD	1.04	1.24	2.60
CEC (cmol <sup>+</sup> kg <sup>-1</sup> )	SEP	3.03	2.59	1.09
	RPD	0.78	0.92	2.18

La figure 3 illustre les relations obtenues pour les trois modèles retenus sur base du lot de validation qui compte 150 échantillons. Les deux premiers modèles permettent de prédire un plus grand nombre d'échantillons mais avec une précision moindre ; le modèle local donnant de meilleurs résultats que l'approche globale. Le troisième modèle permet de prédire moins d'échantillons, une part importante des terres de culture sont prédites mais les sols sous prairies et sous forêt le sont plus rarement. Par contre, la qualité de la prédiction est très nettement améliorée.

La moins bonne prédiction des terres sous prairie et sous forêt est essentiellement liée au fait que la base de données spectrales ne contient pas suffisamment de données représentatives de ces milieux.

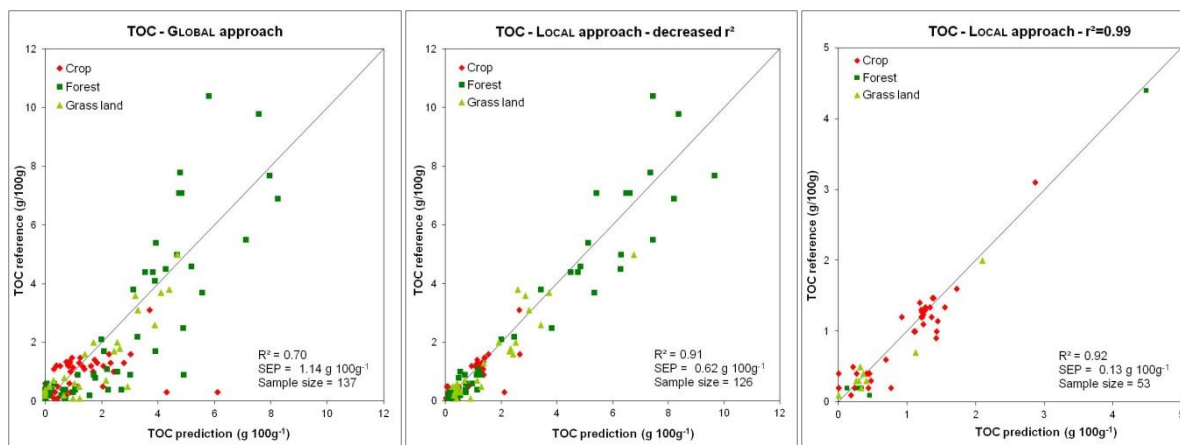


Figure 3 : Prédiction du set de validation par les trois modèles retenus (extrait de Genot et al., 2011).

### Etude de la répétabilité et de la fidélité intermédiaire de la méthode

Cette étude a été réalisée sur base de quatre échantillons couvrant une gamme intéressante de taux d'argile, de carbone ainsi qu'une large gamme de capacité d'échange cationique. Sur un même appareil, les échantillons ont été scannés par deux manipulateurs sur trois jours, chaque manipulateur ayant réalisé quatre remplissages des coupelles et chaque coupelle ayant été scannée trois fois. La figure 4 présente les spectres obtenus lors de cette étude.

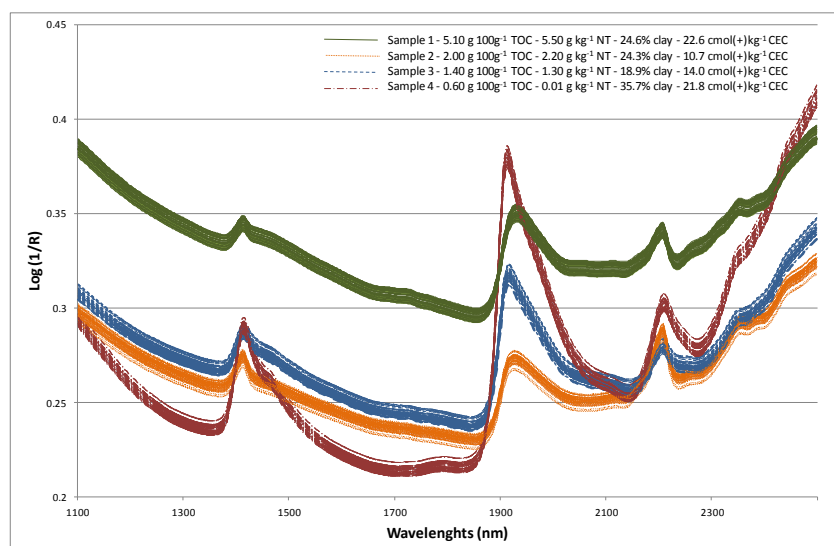


Figure 4 : Spectres obtenus pour les quatre échantillons scannés pour déterminer la répétabilité et la fidélité intermédiaire de la technique SPIR pour les quatre propriétés étudiées (extrait de Genot et al., 2011).

Le mode de calcul des valeurs de répétabilité et de fidélité intermédiaire est décrit dans la norme ISO 5725. Le tableau 2 présente les résultats obtenus par SPIR en terme de répétabilité et de fidélité intermédiaire.

**Tableau 2** : Répétabilité et fidélité intermédiaire obtenues pour l'analyse par SPIR des quatre propriétés de sol étudiés.

Paramètres	sr	r (%)	sR	R (%)
COT (g 100g <sup>-1</sup> )	0.001	4.5 %	0.008	10.6 %
NT (g kg <sup>-1</sup> )	0.002	4.3 %	0.005	7.3 %
Argile (%)	0.83	9.7 %	1.82	21.3 %
CEC (cmol(+) kg <sup>-1</sup> )	0.19	6.9 %	0.81	14.1 %

La norme ISO 14235 pour la détermination du carbone organique total donne une répétabilité estimée à 6.5 %, soit plus élevée que celle obtenue par SPIR. Il en est de même pour l'azote total, où dans la norme ISO 11261, la répétabilité est estimée entre 7 et 17 % selon le taux d'azote total de quatre échantillons (1.9 g kg<sup>-1</sup> à 10.9 g kg<sup>-1</sup>). La répétabilité estimée à l'Unité de Science du Sol de GxABT est estimée à 8 % pour le taux d'argile et la CEC.

On constate que la technique infrarouge est au moins aussi fidèle que les méthodes de référence et que, comparativement à l'erreur standard de prédiction, la part d'erreur liée à la fidélité est très faible.

### ***Intégration de la technique infrarouge en remplacement ou en complément des analyses de routine***

Après le transfert de la base de données spectrales et des modèles de prédiction, les laboratoires sont en mesure d'utiliser la technique en remplacement ou en complément des méthodes d'analyses utilisées traditionnellement dans les laboratoires.

La base de données est centralisée au Centre Wallon de Recherches agronomiques et les XDS communiquent en continu avec cette base de données, de manière à utiliser systématiquement la dernière version de la base. Lorsqu'un échantillon est scanné dans un laboratoire, en fonction du paramétrage défini pour le mode local, une sélection des échantillons les plus proches est effectuée et le modèle prédit les quatre propriétés de sol. Une valeur de GH et de NH est associée à chaque valeur analytique proposée donnant une idée sur la position du spectre par rapport à l'ensemble des spectres de la base de données et donc évaluant la qualité de la prédiction.

Dans un premier temps (années 2009-2010), trois laboratoires ont scanné leurs échantillons sur leur FOSS XDS et ont continué à analyser ces échantillons selon leurs procédures en routine. Comme les quatre propriétés de sol ne sont pas analysées en routine ou sont analysés selon d'autres méthodes, une sélection des spectres les plus différents, sur base de la corrélation entre les spectres, a été réalisée et les échantillons analysés selon les méthodes de référence avant d'être intégrés à la base de données.

Ensuite, à partir de 2011, un laboratoire a commencé à utiliser la technique en routine. Chaque échantillon est d'abord scanné. En fonction des valeurs de GH ou NH associés à la valeur prédite, celle-ci est retenue ou rejetée. Dans ce dernier cas, l'échantillon est analysé selon les voies classiques.

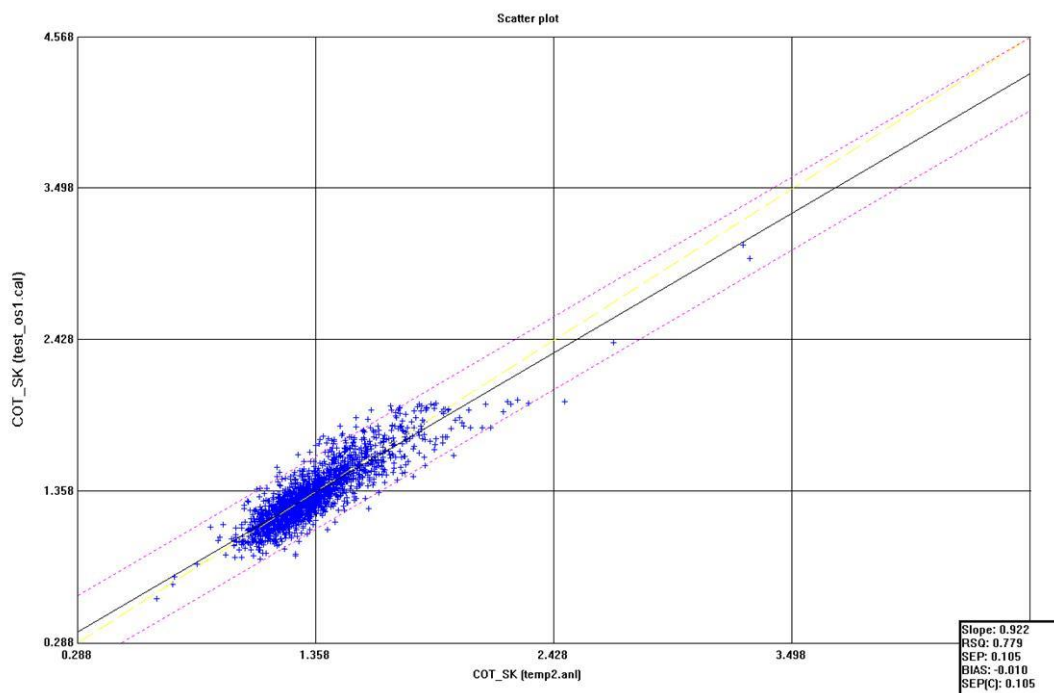
Le tableau 3 illustre la part des échantillons prédits selon les modèles locaux élaborés en les différenciant selon l'occupation du sol. Nous observons qu'environ 70 % des terres sous culture ou sous prairie sont correctement prédites par la technique. Cette proportion est d'environ 50 % pour la CEC. Pour les taux d'azote total et d'argile, le nombre d'échantillons correctement prédits est moins important, essentiellement en relation avec un moindre nombre d'analyses de référence dans la base de données pour ces deux propriétés. Ces statistiques indiquent que d'ores et déjà, la technique infrarouge permet de fortement diminuer le nombre d'analyses en laboratoire.

**Tableau 3 :** Nombre d'échantillons prédits dans un laboratoire du réseau REQUASUD (ASBL CPL PROMOGEST, Tinlot) par la spectroscopie proche infrarouge.

Paramètres	OS	COT (g 100g <sup>-1</sup> )	NT (g kg <sup>-1</sup> )	Argile (%)	CEC (cmol(+) kg <sup>-1</sup> )
Nbr total d'éch.		3249			
Nbr d'éch. Prédits	CULT	2186	989	298	1199
% d'éch. prédits		76 %	20 %	7 %	43 %
Nbr total d'éch.		1545			
Nbr d'éch. Prédits	PR	975	479	331	597
% d'éch. prédits		67 %	29 %	19 %	41 %

Sur base de 3 000 échantillons de terre de culture pour lesquels nous disposons des spectres et des valeurs analytiques de référence, une évaluation de la qualité de la prédiction a pu être réalisée pour le taux de carbone organique total. Deux mille échantillons sont correctement prédits sur base des critères définis (GH et NH). La figure 5 compare la relation entre les valeurs obtenues pour le COT selon les méthodes classiques aux valeurs prédites pour les terres de cultures pour ces 2 000 échantillons. Le SEP obtenu est de 0.11. Par comparaison, au BIPEA, pour la mesure du carbone organique total selon la norme ISO 14235, l'écart-type robuste considéré varie de 0.06 à 0.15 sur base des 8 derniers résultats d'essais interlaboratoires.

Le travail important réalisé pour implémenter la base de données spectrales au cours de ces dernières années (8 000 données pour le COT par exemple) a donc permis une amélioration sensible de la qualité des prédictions qui s'approche de ce qu'on peut observer pour les analyses de référence.



**Figure 5 :** Relation entre valeurs de COT prédites et valeurs de référence pour 2 000 échantillons de terre de culture.

## Conclusions

Comme dans la plupart des études basées sur la SPIR, la constitution de la base de données spectrales et l'étape de calibrage sont les phases les plus coûteuses et les plus longues de la méthodologie. Les investigations menées ici ont permis de montrer que l'approche par régression PLS local donnait des résultats très encourageants pour prédire quatre propriétés de sol, le taux de

carbone organique total, d'azote total, d'argile et la capacité d'échange cationique. La base de données spectrales disponible au terme de la phase de calibrage permettait de prédire correctement les sols sous culture. Celle-ci méritait cependant d'être complétée pour les autres occupations de sol. Après transfert de la base de données et des modèles de prédiction vers les laboratoires du réseau, de nouvelles données ont pu être intégrées à la base et les modèles ont pu être testés avant une utilisation en routine.

A l'heure actuelle, environ 120 échantillons peuvent être scannés chaque jour par un technicien sur un FOSS XDS, plus de la moitié de ces échantillons sont correctement prédits par les modèles pour le taux de carbone et la capacité d'échange cationique, ce qui réduit considérablement le temps de travail au laboratoire. La qualité des prédictions s'approchant, au moins pour le taux de carbone organique total, de la justesse observable pour les méthodes de référence.

La technique SPIR est donc potentiellement utilisable en complément ou en remplacement de méthodes analytiques existantes et contribue donc à un conseil de fumure personnalisé.

## Remerciements

Nous remercions l'ASBL CPL PROMOGEST, laboratoire d'analyses de terre à Tinlot-Scry, pour sa contribution très active à chacune des étapes de ce travail.

## Bibliographie

C.W. Chang, D.A. Laird, M.J. Mausbach and C.R. Hurburgh, 2001. "Near-infrared reflectance spectroscopy - Principal component regression analyses of soil properties", *Soil Science Society of America Journal* **65**, 480-490.

V. Genot, 2006, "Appréhension de la variabilité spatiale et vertical des sols wallons et potentialité de la spectroscopie proche infrarouge. Contribution à l'élaboration d'un conseil de fumure personnalisé", DEA, ULg GxABT, Gembloux, p 70.

V. Genot, G. Colinet, L. Bock, D. Vanvyve, Y. Reusen, P. Dardenne, 2011. "Near infrared reflectance spectroscopy for estimating soil characteristics valuable in the diagnosis of soil fertility", *Journal of Near Infrared Spectroscopy*, **19**, 117-138.

C.A. Russell, J.F. Angus, G.D. Batten, B.W. Dunn and R.L. Williams, 2002. "The potential of NIR spectroscopy to predict nitrogen mineralization in rice soils", *Plant and Soil* **247**, 243-252.

K.D. Shepherd and M.G. Walsh, 2002. "Development of reflectance spectral libraries for characterization of soil properties", *Soil Science Society of America Journal* **66**, 988-998.

B. Stenberg, R.A. Viscarra Rossel, A.M. Mouazen and J. Wetterlind, 2010, "Visible and Near Infrared Spectroscopy in Soil Science" *Advances in Agronomy* **107**, 163-215.