



Organisation des Nations Unies  
pour l'alimentation  
et l'agriculture

1

MATÉRIEL DE  
FORMATION



# Spectroscopie du sol

## MATÉRIEL DE FORMATION

**1** | Une **introduction** à l'analyse des sols par spectroscopie dans le visible et le proche infrarouge (vis-NIR) ainsi que dans le moyen infrarouge (MIR)





# Spectroscopie du sol

## MATÉRIEL DE FORMATION

**1** | Une **introduction** à l'analyse des sols par spectroscopie dans le visible et le proche infrarouge (vis-NIR) ainsi que dans le moyen infrarouge (MIR)

Par

**Yufeng Ge**

Université du Nebraska

**Alexandre Wadoux**

Université de Sydney

**Yi Peng**

Partenariat mondial sur les sols, FAO



Citation requise:

Ge, Y., Wadoux, A., Peng, Y. 2024. *Une introduction à l'analyse des sols par spectroscopie dans le visible et le proche infrarouge (vis-NIR) ainsi que dans le moyen infrarouge (MIR)*. Rome, FAO. <https://doi.org/10.4060/cb9005fr>

Les appellations employées dans ce produit d'information et la présentation des données qui y figurent n'impliquent de la part de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) aucune prise de position quant au statut juridique ou au stade de développement des pays, territoires, villes ou zones ou de leurs autorités, ni quant au tracé de leurs frontières ou limites. Le fait qu'une société ou qu'un produit manufacturé, breveté ou non, soit mentionné ne signifie pas que la FAO approuve ou recommande ladite société ou ledit produit de préférence à d'autres sociétés ou produits analogues qui ne sont pas cités.

Les opinions exprimées dans ce produit d'information sont celles du/des auteur(s) et ne reflètent pas nécessairement les vues ou les politiques de la FAO.

© FAO, 2024



Certains droits réservés. Ce travail est mis à la disposition du public selon les termes de la Licence Creative Commons - Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 3.0 Organisations Internationales (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo/deed.fr>).

Selon les termes de cette licence, ce travail peut être copié, diffusé et adapté à des fins non commerciales, sous réserve de mention appropriée de la source. Lors de l'utilisation de ce travail, aucune indication relative à l'approbation de la part de la FAO d'une organisation, de produits ou de services spécifiques ne doit apparaître. L'utilisation du logo de la FAO n'est pas autorisée. Si le travail est adapté, il doit donc être sous la même licence Creative Commons ou sous une licence équivalente. Si ce document fait l'objet d'une traduction, il est obligatoire d'intégrer la clause de non responsabilité suivante accompagnée de la citation indiquée ci-dessous: «Cette traduction n'a pas été réalisée par l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO). La FAO n'est pas responsable du contenu ou de l'exactitude de cette traduction. L'édition originale [langue] doit être l'édition qui fait autorité.»

Tout litige relatif à la licence ne pouvant être réglé à l'amiable sera soumis à une procédure de médiation et d'arbitrage au sens de l'Article 8 de la licence, sauf indication contraire aux présentes. Les règles de médiation applicables seront celles de l'Organisation mondiale de la propriété intellectuelle (<http://www.wipo.int/amc/fr/mediation/rules>) et tout arbitrage sera mené conformément au Règlement d'arbitrage de la Commission des Nations Unies pour le droit commercial international (CNUDCI).

**Documents de tierce partie.** Les utilisateurs qui souhaitent réutiliser des matériels provenant de ce travail et qui sont attribués à un tiers, tels que des tableaux, des figures ou des images, ont la responsabilité de déterminer si l'autorisation est requise pour la réutilisation et d'obtenir la permission du détenteur des droits d'auteur. Le risque de demandes résultant de la violation d'un composant du travail détenu par une tierce partie incombe exclusivement à l'utilisateur.

**Ventes, droits et licences.** Les produits d'information de la FAO sont disponibles sur le site web de la FAO ([www.fao.org/publications](http://www.fao.org/publications)) et peuvent être acquis par le biais du courriel suivant: [publications-sales@fao.org](mailto:publications-sales@fao.org). Les demandes pour usage commercial doivent être soumises à: [www.fao.org/contact-us/licence-request](http://www.fao.org/contact-us/licence-request). Les demandes relatives aux droits et aux licences doivent être adressées à: [copyright@fao.org](mailto:copyright@fao.org).

# Contenu

Collaborateurs	V
Résumé exécutif	V
1   Contexte	1
2   Notions fondamentales de vis-NIR et MIR pour l'analyse des sols	1
2.1 Le spectre EM, la longueur d'onde et le nombre d'ondes	1
2.2 Spectres de réflectance et d'absorbance	1
2.3 Absorptions fondamentales dans le MIR, harmoniques et combinaisons dans le NIR	2
2.4 Propriétés du sol qui peuvent être estimées directement ou indirectement par les spectres vis-NIR et MIR	2
2.5 Avantages de l'analyse des sols par vis-NIR et MIR	3
3   Procédures d'analyse des sols par vis-NIR et MIR	4
3.1 Préparation de l'échantillon	5
3.2 Acquisition spectrale	5
3.3 Prétraitement spectral	6
3.4 Développement et test des modèles.	7
3.5 Méthodes de développement des modèles	9
3.6 Évaluation du modèle	10
4   Les bibliothèques spectrales vis-NIR et MIR du sol à l'échelle régionale, continentale et mondiale: motivations, avantages et réserves	13
5   Instruments communs pour l'acquisition spectrale par Vis-NIR et MIR du sol	17
6   Remarques finales	17
Bibliographie	18

# Figures

Figure 1. Spectre vis-NIR d'un sol en réflectance (à gauche) et en absorbance (à droite)	2
Figure 2. Le processus de travail de la spectroscopie vis-NIR et MIR pour l'analyse des sols	4
Figure 3. Spectres de réflectance Vis-NIR de l'ensemble de données de l'exemple (n=201) avec différentes méthodes de prétraitement spectral: (A) spectres de réflectance originaux; (B) spectres d'absorbance; (C) spectres après « Standard Normal Variate » (SNV); (D) spectres après correction multiplicative du signal (MSC) ; (E) spectres après suppression du continuum; et (F) spectres de la première dérivée après le filtrage Savitzky-Golay. Les lignes bleues représentent les spectres moyens calculés sur tous les échantillons	7
Figure 4. Illustration de l'ajustement excessif du modèle avec la régression des moindres carrés partiels	8
Figure 5. Illustration de l'ACP sur les données spectrales MIR du sol: (A) les données MIR du sol dans l'espace spectral; (B) le diagramme de dispersion des 10 premières composantes principales et le pourcentage de variance que chaque PC représente; et (C) le nuage de points dans l'espace défini par chaque paire des 3 premières composantes principales	10
Figure 6. Données simulées pour montrer la comparaison des propriétés du sol prédites par le vis-NIR ou le MIR à celles mesurées en laboratoire dans des diagrammes de dispersion. Différents niveaux de performance du modèle sont donnés et leurs critères d'évaluation sont fournis	11
Figure 7. Diagrammes de dispersion des propriétés du sol prédites par le système vis-NIR et mesurées en laboratoire pour le jeu de données test (n = 60) à l'aide de la régression partielle par les moindres carrés. La ligne grise en pointillés est une ligne d'égalité entre le mesuré au laboratoire et le prédit, tandis que la ligne bleue est la ligne ajustée aux mesures	12
Figure 8. Diagrammes de dispersion des propriétés du sol prédites par le MIR et mesurées en laboratoire pour le jeu de données test (n = 108) en utilisant la régression partielle par les moindres carrés	12

# Tableaux

Tableau 1. Le résultat du test de prédiction des trois propriétés du sol avec la modélisation PLSR et SVR, pour vis-NIR et MIR	13
Tableau 2. Résumé des bibliothèques spectrales vis-NIR et MIR du sol publiées aux niveaux national, continental et mondial	15

# Collaborateurs

Tous les noms énumérés ici sont présentés par ordre alphabétique.

## Coordination générale

Ronald Vargas Rojas (FAO-GSP)

## Auteurs

Alexandre Wadoux (Université de Sydney)

Yi Peng (FAO-GSP)

Yufeng Ge (Université du Nébraska)

## Réviseurs

Budiman Minasny (Université de Sydney)

Jose Alexandre Melo Dematte (Université de Sao Paulo, Brésil)

## Édition et publication

Filippo Benedetti (FAO-GSP)

Matteo Sala (FAO-GSP)

Tasneem Alsiddig (FAO-GSP)

## Édition française

Sougueh Cheik (laboratoire agronomie et écologie,  
Institut des sciences de la vie, CERD - Djibouti)

Elh Moudi Moustapha Abdourahaman  
(IRAN/centre régional de la recherche agronomique - Niger)

Mariage Clemence (axe échanges eau-sol-plantes, GxABT,  
Université de Liège - Belgium)

Alexandre Wadoux (Université de Sydney)

Isabelle Verbeke (FAO-GSP)

# Remerciements

"Une introduction à l'analyse des sols par spectroscopie dans le visible et le proche infrarouge (vis-NIR) ainsi que dans le moyen infrarouge (MIR)" est le premier matériel de formation sur le thème de la spectroscopie des sols pour les niveaux débutants, grâce à l'Initiative du Réseau mondial des laboratoires des sols sur la spectroscopie des sols (GLOSOLAN-Spec) du Partenariat mondial sur les sols, FAO. Il est le résultat de la collaboration d'experts en spectroscopie des sols de différents instituts à travers le monde. Ce document est une introduction à l'utilisation de la spectroscopie des sols pour l'analyse des sols; il permet aux lecteurs de comprendre les procédures fondamentales et de base de l'utilisation de cette technologie pour l'analyse des sols.

L'initiative GLOSOLAN-Spec et les auteurs sont particulièrement reconnaissants envers le projet de la Banque mondiale "Leveraging technology for Uzbekistan's agriculture modernization" financé par le Korean Green Growth Trust Fund pour le soutien financier de la préparation de ce document et l'apport de technologies modernes d'analyse des sols à tous les pays.

Les auteurs tiennent également à remercier le Kellogg Soil Survey Laboratory de l'USDA-NRCS pour avoir fourni les jeux de données vis-NIR et MIR des sols permettant d'illustrer ce document d'introduction.

Enfin, des remerciements particuliers sont adressés à Sougueh Cheik, Elh Moudi Moustapha Abdourahaman, Mariage Clemence et Alexandre Wadoux pour leur collaboration à la préparation de la version française du manuel paru initialement en anglais.

# Résumé exécutif

La spectroscopie de réflectance dans le visible et le proche infrarouge (vis-NIR) et dans le moyen infrarouge (MIR) a émergé et s'est développée comme une méthode importante pour l'analyse quantitative des sols, avec un potentiel pour devenir une alternative à l'analyse chimique conventionnelle en laboratoire pour plusieurs propriétés des sols. La spectroscopie vis-NIR et MIR est plus souhaitable en raison de sa rapidité, de son faible coût et de son caractère non-destructif dans l'analyse, mais elle nécessite un nouvel ensemble de compétences au sein du personnel de laboratoire et des praticiens. Ce document d'introduction est destiné aux débutants qui veulent utiliser la spectroscopie vis-NIR et MIR dans l'analyse des sols. Le manuel de formation couvre les sujets suivants: (1) les principes fondamentaux de la spectroscopie vis-NIR et MIR et les interactions avec le sol (2) les procédures courantes de laboratoire pour l'analyse des sols par vis-NIR et MIR, en mettant l'accent sur l'acquisition spectrale, le prétraitement spectral, le développement et le test des modèles, la régression partielle des moindres carrés et l'évaluation des performances des modèles, et (3) les bibliothèques spectrales des sols par vis-NIR et MIR à l'échelle régionale, nationale et mondiale. Ce document est le premier d'une série de trois supports de formation couvrant les sujets de base, intermédiaires et avancés de la spectroscopie vis-NIR et MIR du sol.





# 1 | Contexte

La spectroscopie de réflectance dans le visible et l'infrarouge proche (vis-NIR) et dans l'infrarouge moyen (MIR) a émergé et s'est développée au cours des trois dernières décennies comme une méthode importante pour l'analyse quantitative des sols en laboratoire (Baumgardner *et al.*, 1986; Chang *et al.*, 2001; Reeves, 2010; Viscarra Rossel *et al.*, 2006). De nombreux chercheurs pensent que le vis-NIR et le MIR peuvent devenir une alternative aux méthodes conventionnelles de chimie en laboratoire pour l'analyse des sols (Janik, Merry et Skjemstad, 1998; Nocita *et al.*, 2015). Diverses applications modernes nécessitent de grandes quantités de données quantitatives sur les sols à haute résolution (à la fois dans l'espace et dans le temps). L'agriculture de précision en est un exemple: des échantillons de sol sont régulièrement prélevés sur le terrain (par exemple selon un quadrillage) et analysés en laboratoire pour générer des cartes des propriétés du sol. Ces cartes des propriétés du sol deviennent ensuite des cartes de référence pour générer des zones de gestion ou pour guider les applications à taux variable d'engrais, d'eau et de chaux (Nawar *et al.*, 2017). Un autre exemple est la séquestration du carbone dans le sol (en lien avec le droit d'émission de gaz carbonique), où le même champ pourrait être échantillonné et mesuré à plusieurs reprises pour déterminer les changements dans le stock de carbone organique du sol (Smith *et al.*, 2020). Le nombre d'échantillons de sol à analyser nécessite des méthodes rapides et peu coûteuses comme le vis-NIR et le MIR pour rendre ces applications viables d'un point de vue économique. C'est pour ces raisons qu'il y a un grand intérêt à rendre opérationnel le vis-NIR et le MIR pour l'analyse de routine des sols.

## 2 | Notions fondamentales de vis-NIR et MIR pour l'analyse des sols

### 2.1 Le spectre EM, la longueur d'onde et le nombre d'ondes

Le spectre électromagnétique (EM) est composé de rayons gamma, de rayons X, d'ultraviolets, de visible, d'infrarouge (proche, moyen et lointain), de micro-ondes et d'ondes radio, couvrant plusieurs ordres de grandeur en longueur d'onde  $\lambda$  (ou inversement, en fréquence  $\nu$ ). Les parties du spectre EM les plus utilisées pour l'analyse des sols sont le visible et l'infrarouge proche (vis-NIR) et l'infrarouge moyen (MIR). Le Vis-NIR est conventionnellement spécifié en longueur d'onde en nm (nanomètre,  $10^{-9}$  m) ou  $\mu\text{m}$  (micromètre,  $10^{-6}$  m). Le Vis-NIR combine les zones du visible et du proche infrarouge et se réfère généralement à une gamme de longueurs d'onde comprise entre 350 et 2500 nm (0,35 à 2,5  $\mu\text{m}$ ). Le MIR est conventionnellement spécifié en nombre d'onde ( $\text{cm}^{-1}$ ), ce qui signifie littéralement le nombre d'ondes EM qui tiennent dans une longueur d'un cm (centimètre). Le MIR commence généralement à 4 000  $\text{cm}^{-1}$  et se termine à 600 (ou 400)  $\text{cm}^{-1}$ , selon les instruments utilisés. La longueur d'onde  $\lambda$  en nm et le nombre d'onde en  $\text{cm}^{-1}$  sont inversement liés par:  $\text{nombre d'onde} = \frac{10^7}{\lambda}$ . Prenons l'exemple d'une longueur d'onde de 2 500 nm; son  $\lambda$  nombre d'onde équivalent est le suivant  $\frac{10^7}{2500} = 4000 \text{cm}^{-1}$  et on peut voir dans cet exemple que la fin du vis-NIR correspond au début du MIR dans le spectre EM.

### 2.2 Spectres de réflectance et d'absorbance

Lorsque l'énergie EM est dirigée vers la surface d'un sol, elle peut être absorbée, transmise et réfléchi. L'énergie absorbée peut parfois être réémise sous forme de fluorescence. En spectroscopie des sols, c'est l'énergie réfléchi par la surface du sol qui est la plus intéressante. Il existe deux types de réflexion: la réflexion spéculaire (comme les miroirs) et la réflexion diffuse. Le mode de réflexion diffuse est celui que les spectromètres utilisent dans l'analyse des sols. Ce mode est souhaitable car l'énergie EM en réflexion diffuse pénètre et interagit suffisamment avec la matrice du sol, et contient donc plus d'informations sur les constituants du sol. C'est pourquoi des termes comme DRS (Diffuse Reflectance Spectroscopy) ou DRIFTS (Diffuse Reflectance Infrared Fourier Transform Spectroscopy) sont souvent utilisés dans la littérature sur la spectroscopie des sols. En outre, la réflectance totale atténuée (ATR) est un autre mode de mesure appliqué à l'analyse des sols. Les spectres du sol peuvent être représentés en réflectance ou en absorbance, par conséquent, les praticiens doivent soigneusement déterminer s'ils travaillent avec des spectres de réflectance ou d'absorbance. La réflectance (R) et l'absorbance (A) peuvent être converties réciproquement par les équations:

$$A = \log_{10}\left(\frac{1}{R}\right) \text{ et } R = 10^{-A}.$$

La **figure 1** montre un exemple de spectres de réflectance et d'absorbance vis-NIR d'un échantillon de sol.

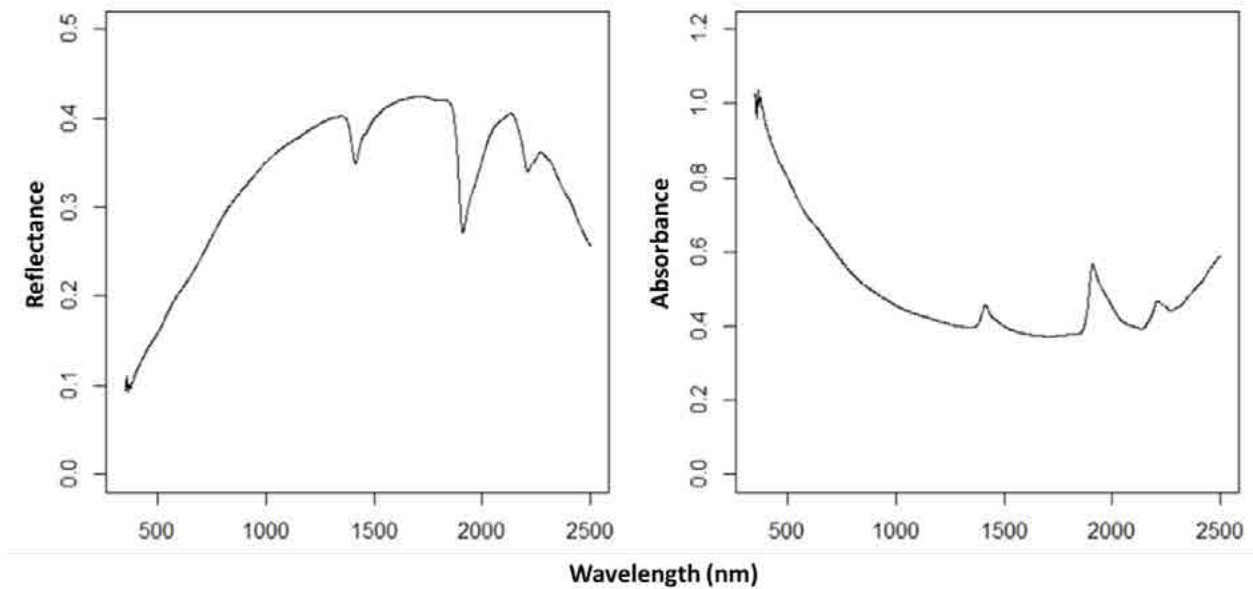


Figure 1. Spectre vis-NIR d'un sol en réflectance (à gauche) et en absorbance (à droite). Source : Auteurs du présent document

### 2.3 Absorptions fondamentales dans le MIR, harmoniques et combinaisons dans le NIR

Une autre distinction majeure qui doit être faite entre le vis-NIR et le MIR est le mode d'interaction entre l'énergie EM et le sol. Les énergies MIR provoquent des vibrations moléculaires des liaisons chimiques généralement présentes à la fois dans les composés organiques et minéraux des sols. Dans les composés organiques (c'est-à-dire la matière organique), les liaisons chimiques pertinentes sont O-H, N-H, C-H, C-C, C=C, C-N, C=O, etc... Dans les composés minéraux, les liaisons chimiques pertinentes de Al-OH et Si-OH sont présentes dans les argiles. Ces bandes d'absorption vibratoires fondamentales sont généralement fortes et bien définies; raison pour laquelle les modèles MIR sont supérieurs aux modèles vis-NIR pour prédire les propriétés du sol telles que la matière organique et l'argile. Les harmoniques et les combinaisons de ces bandes fondamentales apparaissent dans la région NIR. Par exemple, les caractéristiques d'absorption des spectres NIR du sol aux alentours de 1450 et 1920 nm sont associées aux harmoniques des vibrations d'étirement O-H et H-O-H de l'eau libre. Les bandes d'absorption dans le NIR sont plus faibles, se chevauchent davantage et sont moins faciles à distinguer que les bandes fondamentales du MIR. En outre, dans le domaine visible, la présence d'oxydes de fer ferreux et ferriques provoque des caractéristiques d'absorption dues aux transitions électroniques des cations de fer. Enfin, une teneur plus élevée en matière organique dans le sol tend à diminuer la réflectance dans le domaine visible, ce qui lui donne un aspect plus sombre.

### 2.4 Propriétés du sol qui peuvent être estimées directement ou indirectement par les spectres vis-NIR et MIR

Le sol est un mélange complexe d'une vaste gamme de constituants chimiques, et présente différents états physiques en termes de taille des particules, d'agrégation, de rugosité de surface et de teneur en eau. Certains de ces constituants physiques et chimiques interagissent avec l'énergie vis-NIR et MIR et produisent des caractéristiques d'absorption dans les spectres, ce qui est essentiellement la base de la spectroscopie des sols avec vis-NIR et MIR. Ces propriétés "primaires" du sol, notamment la matière organique (ou carbone organique), le carbonate (ou carbone inorganique), l'azote total, les minéraux argileux, la teneur en fer, les fractions granulométriques de l'argile, du limon et du sable, et la teneur en eau, peuvent généralement être calibrées à partir des spectres vis-NIR et MIR du sol, car les bandes d'absorption dans les spectres correspondent à ces composants minéraux et organiques du sol. Il est important de noter que les spectres vis-NIR et MIR peuvent également estimer les propriétés du sol qui n'interagissent pas directement avec l'énergie vis-NIR ou MIR. Par exemple, les cations du sol comme le Mg ou le Ca ne provoquent pas d'absorption spectrale active, mais ils peuvent souvent être estimés de manière raisonnable avec le vis-NIR et le MIR, très probablement par une corrélation secondaire avec les minéraux argileux et les carbonates du sol. De la même manière, les propriétés du sol telles que le pH, la CEC (capacité d'échange cationique), la salinité et les teneurs en nutriments (par exemple, le

phosphore et le potassium total) peuvent également être estimées grâce à leurs corrélations avec une ou plusieurs propriétés "primaires" spectralement actives du sol. Il existe des manuels et des références dans la littérature qui résument les absorptions MIR fondamentales des constituants du sol, leurs harmoniques et bandes combinatoires dans le NIR, et les transitions électroniques dans le visible. Les lecteurs intéressés peuvent se référer à Viscarra Rossel *et al.*, 2016 et Soriano-Disla *et al.*, 2014 pour un résumé. Il est important de souligner que, même si de nombreux constituants et groupes chimiques du sol peuvent être attribués à des caractéristiques d'absorption dans les spectres vis-NIR et MIR, ces bandes d'absorption sont rarement utilisées seules pour estimer les propriétés du sol. Les méthodes de modélisation multi-variée et d'apprentissage automatique (« machine learning ») impliquant toutes les bandes d'ondes sont plutôt les approches courantes de la spectroscopie vis-NIR et MIR du sol. En plus des problèmes de corrélation "primaire" et "secondaire" mentionnés ci-dessus, il est également important de réaliser que les approches de modélisation basées sur les spectres vis-NIR et MIR sont empiriques et fortement "orientées données". Ces approches (qui seront examinées plus en détail dans la section 3.5) conduisent souvent à l'obtention de modèles numériques complexes qui ne sont pas faciles à interpréter. Ces modèles sont susceptibles d'être surajustés et il est toujours possible d'obtenir des résultats précis sur la base de corrélations parasites trouvées dans les spectres. Par conséquent, les praticiens sont mis en garde contre la tentation d'utiliser les modèles qui semblent fonctionner le mieux pour les données disponibles, tout en ignorant d'autres connaissances du sol qui pourraient être pertinentes dans l'analyse, comme le fait de considérer si la propriété du sol à estimer est en effet directement ou indirectement active spectralement dans la gamme vis-NIR et MIR.

Bien que l'analyse et la modélisation des spectres vis-NIR ou MIR du sol seuls soient plus courantes, la fusion des spectres vis-NIR et MIR pour estimer les propriétés du sol peut tirer parti des avantages des deux régions spectrales et améliorer la précision de l'estimation. Pour ce faire, un laboratoire doit être équipé d'un instrument vis-NIR et MIR, ce qui représente souvent un investissement financier initial important pour le laboratoire.

## **2.5 Avantages de l'analyse des sols par vis-NIR et MIR**

L'analyse des sols par vis-NIR et MIR présente plusieurs avantages évidents par rapport aux méthodes conventionnelles d'analyse des sols basées sur la chimie. Premièrement, l'obtention de données spectrales vis-NIR et MIR est rapide et non destructive; la prise d'un seul spectre ne prend que quelques secondes et aucun matériau du sol n'est consommé pendant le processus de mesure. Deuxièmement, le spectre d'une seule mesure permet l'estimation simultanée de plusieurs propriétés du sol sur le même volume de matériau. Ceci est différent des méthodes traditionnelles de laboratoire où chaque propriété du sol est analysée indépendamment sur différents sous-échantillons, et permet ainsi d'éviter la variation potentielle introduite par la micro-hétérogénéité entre les sous-échantillons. Troisièmement, après le coût initial d'acquisition du spectromètre, il s'agit d'une technologie peu coûteuse et respectueuse de l'environnement car elle nécessite une préparation minimale des échantillons (principalement le séchage et le broyage) et aucun réactif chimique n'est nécessaire. Bien que les spectromètres vis-NIR et MIR puissent être coûteux (par exemple, le coût d'un spectroradiomètre ASD Labspec pour l'analyse vis-NIR des sols est compris entre 50 000 et 65 000 USD; et un instrument Bruker FTIR avec un accessoire d'échantillonnage à haut débit pour l'analyse MIR coûte plus de 100 000 USD; ces prix peuvent également varier considérablement d'un pays à l'autre et d'une région à l'autre), le retour sur investissement de l'équipement initial est rapide car des dizaines de milliers d'échantillons de sol peuvent être scannés et analysés, avec des gains importants par rapport à l'investissement initial. Ces avantages combinés conduisent à une analyse quantitative des sols avec un débit beaucoup plus élevé et un coût plus faible que lors de l'utilisation des méthodes conventionnelles de chimie en laboratoire. Il convient de noter que des spectromètres NIR portables moins coûteux avec des gammes de longueurs d'onde limitées sont disponibles sur le marché depuis peu. Cependant, tous les spectromètres ne peuvent pas produire une estimation précise des propriétés du sol, cela dépend généralement de la couverture en longueur d'onde. Plusieurs études ont examiné ces spectromètres NIR limités pour l'analyse des sols (Sharififar *et al.*, 2019; Tang, Jones et Minasny, 2020).

### 3 | Procédures d'analyse des sols par vis-NIR et MIR

Le processus de travail de base de l'analyse des sols par vis-NIR et MIR est résumé dans la **figure 2**. Les méthodes vis-NIR et MIR se distinguent des autres techniques conventionnelles d'analyse des sols en ce que les instruments vis-NIR ou MIR ne rapportent pas directement les résultats quantitatifs des propriétés des sols. Plutôt, un jeu de données d'essai (également appelé jeu de données de calibrage) est utilisé pour former (ou calibrer) des modèles empiriques qui relient les données spectrales aux propriétés du sol. Par conséquent, les échantillons de sol du jeu de données d'essai doivent également être analysés par une méthode analytique de référence (analyse chimique traditionnelle), en laboratoire (en tant que norme acceptée). Les modèles empiriques doivent être testés par rapport à un ensemble indépendant d'échantillons, qui n'a pas été utilisé pour calibrer le modèle (appelé jeu de données test, qui est également mesuré par la même méthode de référence) afin d'évaluer la performance du modèle. Après une calibration complète et un test, les modèles peuvent alors être utilisés pour estimer les propriétés du sol de nouveaux échantillons inconnus à partir de leurs spectres vis-NIR ou MIR. Le développement et le test des modèles sont au cœur de l'analyse vis-NIR et MIR, et nécessitent un ensemble différent de connaissances et de compétences (par exemple, le traitement d'un grand volume de données spectrales, la modélisation statistique, la programmation dans un langage informatique et l'évaluation des performances du modèle) par rapport à l'analyse des sols conventionnelle en laboratoire. Les étapes de l'analyse des données spectrales vis-NIR et MIR sont décrites dans les sous-sections suivantes.

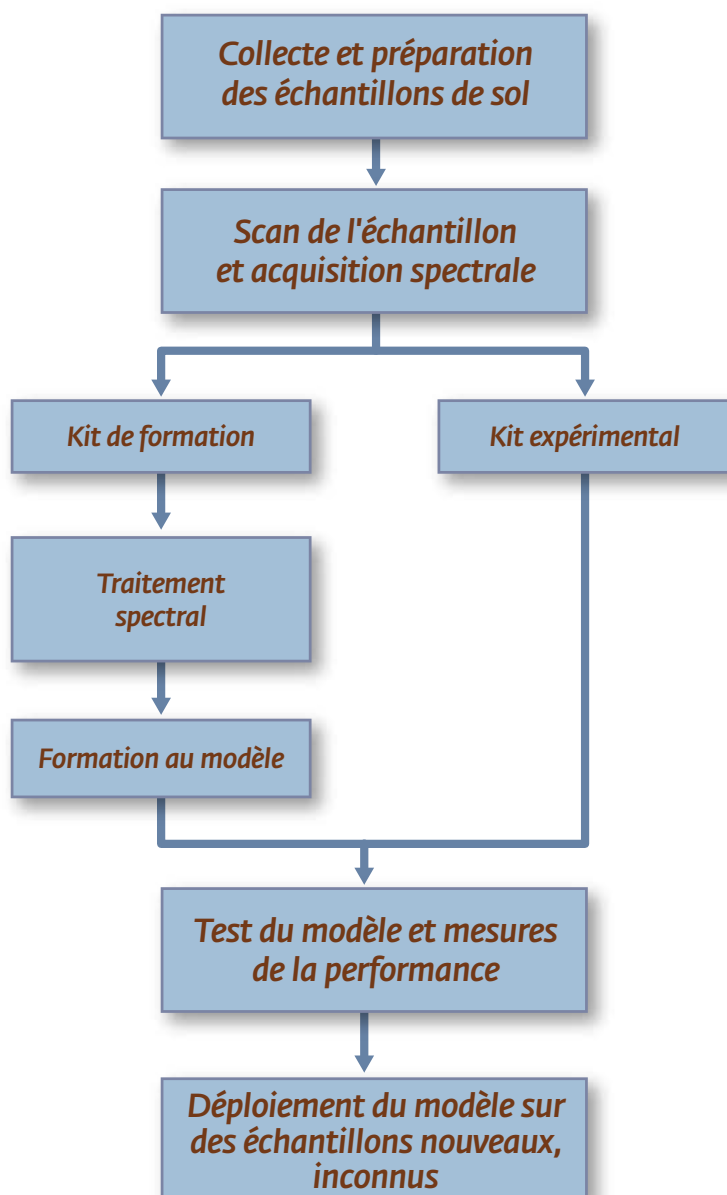


Figure 2. Le processus de travail de la spectroscopie vis-NIR et MIR pour l'analyse des sols. Source : Auteurs du présent document

### 3.1 Préparation de l'échantillon

Lorsque les échantillons de sol sont collectés dans les champs, les éléments grossiers et les racines des plantes doivent être retirés. Les échantillons de sol sont séchés à l'air (35 - 40°C maximum) jusqu'à obtention d'un poids constant et sont broyés pour passer à travers un tamis de 2 mm. C'est la seule préparation d'échantillon nécessaire avant l'acquisition spectrale vis-NIR pour s'assurer que les échantillons sont dans des conditions similaires à celles des échantillons utilisés dans l'analyse chimique classique en laboratoire. Le broyage des échantillons en particules plus fines n'est pas courant pour l'analyse vis-NIR, et comme la taille des particules de l'échantillon affecte la pénétration et la diffusion de l'énergie vis-NIR, un broyage fin rendra les résultats moins comparables à ceux d'autres études (Nduwamungu *et al.*, 2009). Pour la méthode MIR, les échantillons doivent être en outre finement broyés jusqu'à une taille de particule inférieure à 100 µm. Certaines études de la littérature montrent que le broyage fin améliore considérablement les performances de l'analyse MIR (Wijewardane *et al.*, 2021). Le broyage fin est une étape supplémentaire de préparation de l'échantillon et augmente le temps et le coût de l'analyse MIR (mais il est jugé nécessaire pour assurer la meilleure performance de la méthode). Ainsi, un laboratoire devrait investir dans un broyeur mécanique pour le broyage fin des échantillons de sol. Le processus de broyage fin lui-même prend beaucoup plus de temps, y compris le temps nécessaire pour charger et décharger chaque échantillon dans le broyeur, et pour nettoyer et sécher soigneusement les porte-échantillons afin d'éviter la contamination croisée des échantillons de sol; le broyage fin est donc considéré comme un facteur limitant pour le MIR. Cette vidéo détaille la procédure de broyage fin du sol par l'US-DA-NRCS (United States Department of Agriculture - Natural Resources Conservation Service) (<https://www.youtube.com/watch?v=6RpK3zUWMaQ>).

### 3.2 Acquisition spectrale

Pour obtenir la réflectance diffuse des échantillons de sol en vis-NIR, on utilise un panneau standard certifié présentant une réflectance supérieure à 99 % à toutes les longueurs d'onde (considéré comme un réflecteur diffus parfait à 100 %). Ce processus est appelé « référencement blanc ». Dans le MIR, une surface métallique rugueuse (telle que l'aluminium) est suffisante pour servir d'étalon de réflectance. Le spectromètre enregistre trois mesures,  $DN_{\text{Blanc}}$  comme spectre brut de l'étalon de réflectance (DN signifie « numéro numérique »),  $DN_{\text{Noir}}$  comme spectre brut du courant d'obscurité, et  $DN_{\text{Sol}}$  comme spectre brut de l'échantillon de sol. La réflectance du sol est calculée comme suit  $= \frac{DN_{\text{Sol}} - DN_{\text{Noir}}}{DN_{\text{Blanc}} - DN_{\text{Noir}}}$ ; et ce calcul est généralement effectué automatiquement par l'instrument. Le référencement blanc et la mesure du courant d'obscurité doivent être effectués périodiquement au cours d'une session pour s'assurer que l'instrument est bien calibré. Dans le laboratoire, où les facteurs environnementaux tels que la température et l'humidité sont stables dans le temps, l'étalonnage de l'instrument peut être effectué toutes les 15 minutes. Cependant, dans les applications les plus rigoureuses et exigeantes, cette étape d'étalonnage est mise en œuvre entre chaque échantillon.

Une autre étape importante pour garantir la qualité des spectres du sol consiste à allumer l'instrument suffisamment longtemps (selon le type et la marque de l'instrument) avant d'effectuer toute mesure. La sortie spectrale des sources lumineuses vis-NIR et MIR, ainsi que la réactivité des détecteurs spectraux, dépendent toutes de la température. En fait, de nombreux instruments vis-NIR et MIR utilisent certains types de refroidissement (par exemple, par des moyens thermoélectriques ou de l'azote liquide) pour maintenir un faible bruit thermique dans les détecteurs. Cette étape permet d'éviter l'influence négative du réchauffement initial de l'instrument sur les données spectrales.

Les protocoles et les accessoires avec lesquels les échantillons de sol sont présentés aux différents instruments varient. Ici, nous fournissons une description de la présentation des échantillons pour les spectromètres ASD, qui sont les spectromètres vis-NIR du sol les plus utilisés. L'ASD possède deux types d'accessoires, la lampe (« muglight ») et la sonde de contact. Les échantillons de sol (séchés à l'air, <2mm) peuvent être chargés dans des rondelles et ensuite placés sur la lampe pour être scannés. Des boîtes de Pétri en borosilicate (telles que Duroplan) sont souvent utilisées aussi pour contenir les échantillons de sol et les scanner avec la lampe. La sonde de contact est portable, plus flexible que la lampe, et utilisée plus souvent pour scanner des échantillons dans un sac, ou du sol à l'état naturel. Il est important de nettoyer soigneusement les porte-échantillons (rondelles ou boîtes de Pétri) et la fenêtre en quartz de la lampe/sonde de contact entre les échantillons pour éviter toute contamination croisée.

Les données spectrales du sol sont d'abord stockées dans l'instrument dans des formats propriétaires. Ces instruments fournissent également des logiciels pour le traitement de base des données, le pré-traitement des spectres et la modélisation spectrale,

tels que le logiciel Indigo Pro d'ASD et le logiciel OPUS de Bruker. Cependant, la communauté des sols se tourne vers des logiciels libres comme le langage de programmation R (<https://www.r-project.org>) ou Python pour le traitement et la modélisation des données spectrales du sol. Ces logiciels fournissent des fonctions et bibliothèques spectrales puissantes. Ils sont particulièrement efficaces pour traiter de grands ensembles de données (contenant par exemple des milliers d'échantillons) ou pour utiliser des algorithmes d'apprentissage automatique avancés pour la modélisation spectrale. Les données spectrales brutes sont converties en fichiers TXT (délimités par des tabulations) ou CSV (variables séparées par des virgules) qui contiennent la matrice des données spectrales avant d'être lues dans R ou Python. Dans ces fichiers, les échantillons de sol sont organisés en lignes, et les longueurs d'onde (ou nombre d'onde) en colonnes.

Les sections suivantes présentent les étapes du prétraitement spectral, du développement et du test des modèles, des méthodes de modélisation spectrale et de l'évaluation des performances des modèles. Des exemples de jeux de données en vis-NIR et MIR sont préparés pour démontrer les résultats de ces étapes. Le jeu de données vis-NIR contient les données spectrales de 201 échantillons de sol mesurés par un spectroradiomètre ASD LabSpec. Le jeu de données MIR contient les données spectrales de 540 échantillons mesurés par un spectromètre FT-IR Bruker Vertex 70. Les données de laboratoire incluses pour démontrer la modélisation spectrale sont le carbone organique (mesuré par combustion sèche), l'argile (mesurée par la méthode de la pipette) et le pH (extraction à l'eau 1:1). Les deux jeux de données sont extraits des bibliothèques spectrales vis-NIR et MIR ouvertes maintenues au Kellogg Soil Survey Lab de l'USDA-NRCS. En outre, des codes R sont fournis pour illustrer les concepts, montrer les étapes de modélisation et reproduire les résultats.

### 3.3 Prétraitement spectral

L'étape de «prétraitement spectral» est généralement appliquée aux données spectrales brutes vis-NIR et MIR. Le prétraitement spectral peut réduire le bruit aléatoire dans les spectres bruts, améliorer le rapport signal/bruit, minimiser l'effet étranger de la diffusion de la lumière, réduire la dimensionnalité des données spectrales pour un calcul efficace, et/ou améliorer les caractéristiques d'absorption. La méthode de prétraitement la plus simple est le calcul de la moyenne mobile, qui consiste simplement à calculer la moyenne des données spectrales le long de la longueur d'onde ou du nombre d'onde (par exemple, tous les 10 ou 20 points de données) afin de réduire simultanément le bruit et la dimensionnalité. D'autres méthodes de prétraitement couramment utilisées dans l'analyse des sols sont : (1) la « Standard Normal Variate » (SNV), (2) la correction multiplicative du signal (MSC), (3) le filtrage Savitzky et Golay (SG), et (4) la suppression du continuum. R possède un paquet dédié appelé `prospectr` (<https://CRAN.R-project.org/package=prospectr>) pour le prétraitement spectral.

La méthode SNV normalise chaque ligne de la matrice spectrale en soustrayant chaque ligne de sa moyenne et en la divisant par son écart-type. SNV est une méthode simple de normalisation des données spectrales qui vise à corriger la diffusion de la lumière.

$$SNV_i = \frac{x_i - \bar{x}_i}{s_i}$$

où  $x$  est le  $i$ ème spectre brut,  $\bar{x}$  est sa moyenne et  $S$  est l'écart-type du  $i$ ème spectre.

MSC signifie à l'origine «corrélation de dispersion multiplicative», mais la signification de l'abréviation a changé au fil des ans, car elle est également utile pour d'autres types de problèmes multiplicatifs, outre la dispersion. Le concept de base de la MSC est d'éliminer les non-linéarités dans les données causées par la diffusion des particules dans les échantillons. La MSC est mise en œuvre en alignant chaque spectre brut  $x$  avec un spectre de référence idéal  $x$ , en supposant que  $x = m \cdot x + a$ , comme suit:

$$MSC_i = \frac{x_i - a_i}{m_i}$$

où  $a_j$  et  $m_j$  sont les termes additifs et multiplicatifs, respectivement. Dans la mise en œuvre du MSC, le spectre de référence idéal  $x_r$  n'est généralement pas disponible. Le spectre moyen de l'ensemble de données est souvent utilisé à cette fin. Le filtrage SG (Savitzky et Golay, 1964) est largement utilisé pour le prétraitement des spectres vis-NIR et MIR du sol. Cette méthode est polyvalente et peut être utilisée pour le lissage/la réduction du bruit et la différenciation. Il y a trois paramètres dans le filtrage SG. Le premier paramètre est la taille de la fenêtre ( $2g+1$ ). Le deuxième paramètre détermine l'ordre de la différenciation appliquée (par exemple,  $p=0$  signifie lissage;  $p=1$  signifie dérivée première;  $p=2$  signifie dérivée seconde). Le troisième paramètre est l'ordre des polynômes qui détermine le degré de lissage. L'un des inconvénients du filtrage SG est que pour une taille de fenêtre de  $2g+1$ , le spectre résultant perdra les  $g$  points de données de début et de fin. Par exemple, si une fenêtre de 11 points ( $g=5$ ) est utilisée, il

manquera 5 points de données au début et à la fin du spectre résultant. Enfin, la suppression du continuum vise à éliminer la partie continue du spectre et à accentuer l'absorption. Cette méthode est utile pour la visualisation des bandes d'absorption subtiles et superposées dans le vis-NIR. Il n'est cependant pas courant d'utiliser cette méthode pour la modélisation.

La sélection des méthodes de prétraitement est quelque peu subjective et dépend du cas. De nombreux chercheurs optent pour l'utilisation conjointe de quelques méthodes de prétraitement (par exemple, une correction multiplicative du signal suivie d'un filtrage Savitzky-Golay), tandis que d'autres ont constaté que le prétraitement n'améliore pas nécessairement les performances du modèle. Il n'y a pas d'orientation générale sur l'utilisation des méthodes de prétraitement spectral, et il est recommandé aux utilisateurs d'employer une approche heuristique, d'essayer quelques méthodes de prétraitement, et de choisir celle(s) qui améliore(nt) substantiellement la performance du modèle. La **figure 3** montre les résultats de ces méthodes de prétraitement spectral.

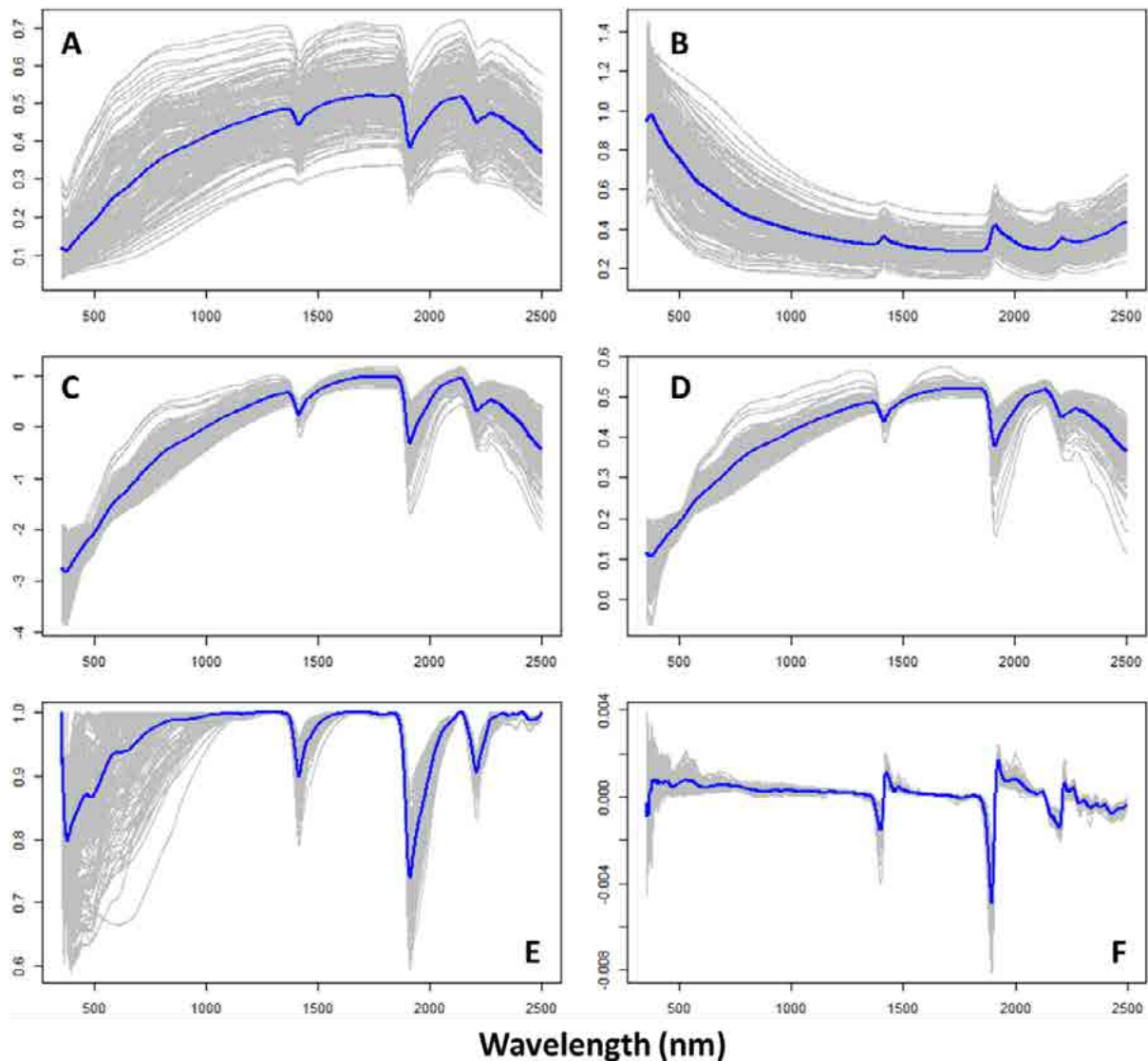


Figure 3. Spectres de réflectance Vis-NIR de l'ensemble de données de l'exemple (n=201) avec différentes méthodes de prétraitement spectral: (A) spectres de réflectance originaux; (B) spectres d'absorbance; (C) spectres après « Standard Normal Variate » (SNV); (D) spectres après correction multiplicative du signal (MSC); (E) spectres après suppression du continuum; et (F) spectres de la première dérivée après le filtrage Savitzky-Golay. Les lignes bleues représentent les spectres moyens calculés sur tous les échantillons  
Source : Auteurs du présent document

### 3.4 Développement et test des modèles.

Comme indiqué précédemment, le jeu des données expérimentales est divisé de manière aléatoire en deux ensembles, l'un pour le développement du modèle et l'autre comme ensemble «indépendant» pour le test du modèle. Les répartitions courantes sont 70 % pour le développement et 30 % pour le test. La division du jeu de données en ensembles de développement et de test est appelée division des données. Idéalement, le processus de fractionnement aléatoire des données est répété plusieurs fois

afin de garantir que le risque d'obtenir une bonne ou une mauvaise prédiction soit faible. D'autres méthodes sont la validation croisée (CV) et le « bootstrapping ». Dans la validation croisée, le jeu de données est divisé en K sous-ensembles disjoints. K-1 sous-ensembles sont utilisés pour la formation du modèle, tandis que la performance est évaluée sur le sous-ensemble restant. Habituellement, K est fixé à 10 (appelé CV à 10 blocs), mais il peut être égal au nombre de spectres, comme dans la « leave-one-out » CV. Dans le « bootstrapping », de multiples sous-échantillons sont sélectionnés avec remplacement pour l'ajustement et l'évaluation du modèle. Ces deux méthodes constituent une meilleure alternative à l'évaluation du modèle que le fractionnement des données, car les statistiques d'évaluation du modèle ne sont pas sensibles au fractionnement aléatoire unique des données que l'on effectue.

Le but du développement du modèle est d'obtenir un modèle empirique multivarié (par exemple, une régression linéaire) qui relie les données spectrales aux propriétés cibles du sol. Comme les données spectrales contiennent généralement des centaines ou des milliers de bandes hautement colinéaires, la clé du développement d'un modèle est de trouver un équilibre entre la « prévisibilité » et la « suradaptation ». L'adaptation excessive désigne une situation dans laquelle le modèle présente d'excellentes performances avec le jeu de données utilisé pour le développement, mais des performances médiocres avec le jeu de données test indépendant; en d'autres termes, le modèle n'a pas la capacité de généraliser. La validation croisée interne avec des données de développement est une méthode utilisée pour trouver le « meilleur modèle », celui qui ne se suradapte pas au jeu de données utilisé pour le développement, en utilisant des statistiques telles que la RMSECV (racine carrée de l'erreur quadratique moyenne de validation croisée, « Root Mean Square Error of CV ») interne, qui peut être comparée à la complexité du modèle pour atteindre un équilibre entre les deux.

Un exemple d'évaluation de suradaptation du modèle obtenu par PLSR (régression selon les moindres carrés partiels, « Partial Least Squares Regression ») est illustré à la **figure 4**, où « développement » fait référence à la RMSE de resubstitution et « CV » à la valeur RMSE obtenue à partir de CV à 10 blocs. Dans la méthode PLSR, le nombre de variables latentes (LV) est un paramètre de modèle indiquant la complexité du modèle. Plus le nombre de LV est élevé dans le modèle, plus la RMSE de resubstitution continue de diminuer. Cependant, pour la CV à 10 blocs, la valeur RMSE atteint le minimum lorsque le nombre de LV est de 13. Par conséquent, nous concluons que les modèles PLSR contenant plus de 13 LV sont surajustés pour cet exemple et les jeux de données utilisés, comme l'indique leur RMSE de développement plus faible mais leur RMSE en validation croisée plus élevée.

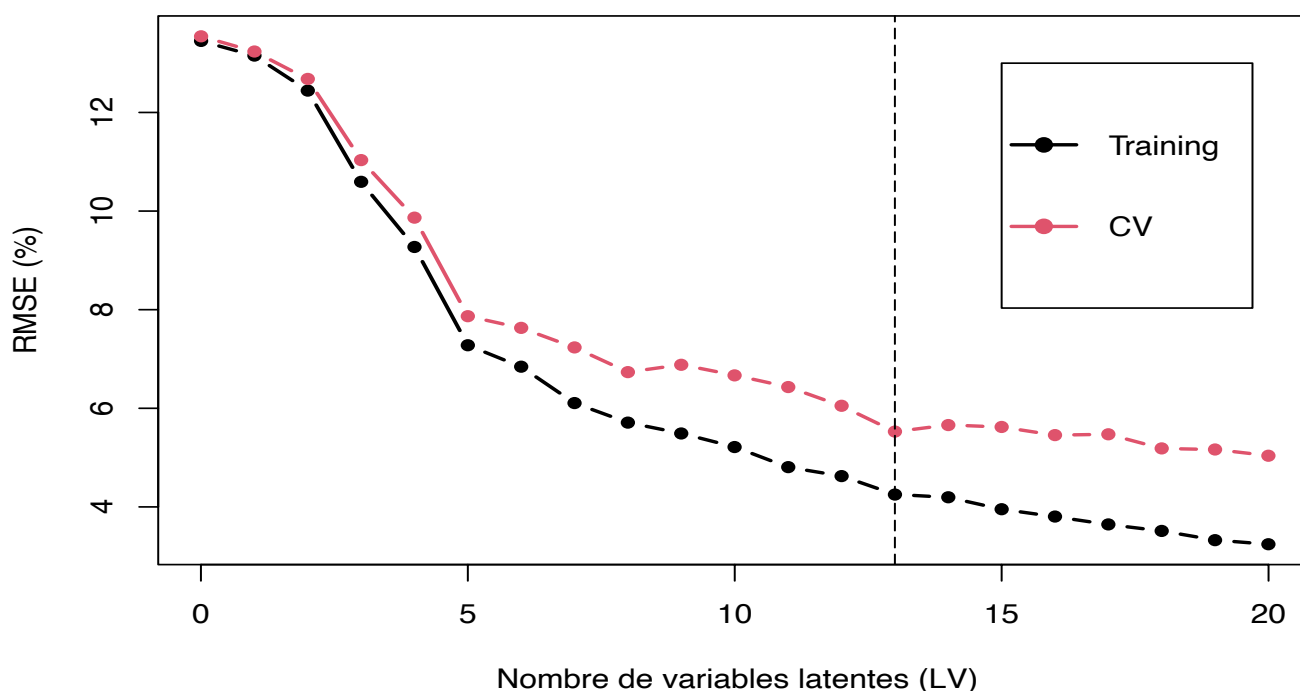


Figure 4. Illustration de l'ajustement excessif du modèle avec la régression des moindres carrés partiels. Source : Auteurs du présent document



### 3.5 Méthodes de développement des modèles

Bien que les bandes d'absorption spectrale associées à certains constituants du sol soient identifiées et résumées dans la littérature, ces bandes seules sont rarement utilisées pour construire des modèles quantitatifs permettant de prédire ces constituants du sol par le biais de la loi de Lambert-Beer bien établie. Au lieu de cela, les approches courantes consistent à former des modèles empiriques multivariés qui associent les données spectrales vis-NIR et MIR aux propriétés cibles du sol. Les premières études ont utilisé la régression linéaire multiple (MLR) pour sélectionner un sous-ensemble de bandes spectrales pour le développement de modèles (Ben-Dor et Banin, 1995). Par la suite, des techniques de modélisation telles que la PLSR et la régression en composantes principales (PCR) ont gagné en popularité (voir, par exemple, Chang *et al.*, 2001 et Viscarra Rossel *et al.*, 2006) et ont été largement utilisées en spectroscopie des sols. La PLSR est devenue la méthode standard de facto à laquelle d'autres nouvelles méthodes de modélisation sont comparées.

Les principales différences entre la PLSR/PCR et la régression linéaire multiple sont les suivantes: (1) la PLSR et la PCR utilisent la totalité du spectre pour la modélisation, alors que la MLR ne sélectionne qu'un petit sous-ensemble de bandes d'ondes pour la modélisation; et (2) la PLSR et la PCR utilisent une projection matricielle pour projeter les données spectrales originales et former un ensemble de nouvelles variables connues sous le nom de variables latentes (LV) ou de composantes principales (PC). Ces nouvelles variables sont dé-corrélées les unes des autres, et les quelques premières LV ou PC représentent une grande partie de la variance totale des données spectrales originales. Ces deux propriétés, LV et PC, permettent de résoudre efficacement le problème de la multi-collinéarité dans les données spectrales vis-NIR/MIR, ainsi que le problème du grand  $p$  petit  $n$  pour les petits ensembles d'échantillons. Ici,  $p$  désigne le nombre de variables (bandes spectrales) et  $n$  désigne le nombre d'échantillons dans la modélisation. Les mathématiques détaillées de la PLSR et de la PCR dépassent le cadre de ce document d'introduction. Les lecteurs intéressés peuvent se référer à Wold, Sjöström et Eriksson, 2001 pour le principe et la mise en œuvre des PLSR et PCR.

Plus récemment, une grande famille de méthodes dites «d'apprentissage automatique» (« machine learning ») s'est répandue dans la spéculation vis-NIR et MIR des sols. Des exemples de ces méthodes sont la forêt d'arbre de décision (« Random Forest », RF), la machine à vecteurs de support (« Support Vector Regression », SVR), Cubist, le réseau de neurones artificiels (« Artificial Neural Networks », ANN), etc. (Minasny et McBratney, 2008; Viscarra Rossel *et al.*, 2016). Les modèles formés sur ces méthodes d'apprentissage automatique montrent généralement une précision prédictive plus élevée que les PLSR et PCR, en particulier lorsque le nombre d'échantillons pour le développement du modèle est important. L'une des raisons de leurs performances supérieures est qu'elles peuvent modéliser plus efficacement les relations non linéaires entre les données spectrales et les propriétés du sol. Les méthodes d'apprentissage automatique sont parfois critiquées pour leur nature «boîte noire», qui se traduit par une faible interprétabilité du modèle. Par exemple, il est très difficile de visualiser, d'interpréter ou d'expliquer un modèle SVM ou ANN en termes de bandes d'ondes jouant un rôle plus important dans la prédiction de certaines propriétés du sol. La mise en œuvre de ces méthodes peut également nécessiter davantage de ressources informatiques. C'est pourquoi, dans de nombreuses applications modernes, le développement du modèle est mis en œuvre sur des GPU (Graphics Processing Unit) ou sur un groupe de GPU/CPU.

L'analyse en composantes principales (ACP) est un concept étroitement associé à la PCR et à la PLSR. En fait, la PCR est la régression linéaire multiple des PC résultant de l'ACP. En spectroscopie des sols, l'ACP est généralement réalisée comme première étape pour explorer la variabilité intrinsèque des données spectrales vis-NIR ou MIR et pour identifier les aberrations spectrales. Il est à noter que dans l'espace spectral original, il est assez difficile d'identifier rapidement les valeurs aberrantes d'un groupe de spectres (comme le montre la **figure 3**). La **figure 5A** montre les données spectrales MIR des 540 échantillons. Il est clair qu'avec cette représentation, il est difficile de dire quels échantillons sont aberrants. La **figure 5B** est le diagramme de dispersion des 10 premiers PC décrivant le pourcentage de variance que chaque PC représente (la première PC représente 78 % de la variance totale, la deuxième PC 8,6 %, la troisième PC 4,2 %, et ainsi de suite). Souvent, nous sommes intéressés par le nombre de PC qui, cumulativement, représentent plus de 90 %, 95 % ou 99 % de la variance totale des données spectrales. Dans cet exemple, les trois premières PC représentent >90% et les cinq premières PC représentent >95% de la variance totale. La **figure 5C** montre le nuage de points dans l'espace défini par chaque paire des 3 premières PC (PC1 vs. PC2 vs. PC3), chaque point représentant un échantillon et correspondant à une courbe dans l'espace spectral original de la **figure 5A**. Le cercle bleu représente l'intervalle de confiance de 99 %. Les points qui se situent à l'extérieur sont des aberrations spectrales potentielles qui méritent une étude plus approfondie.

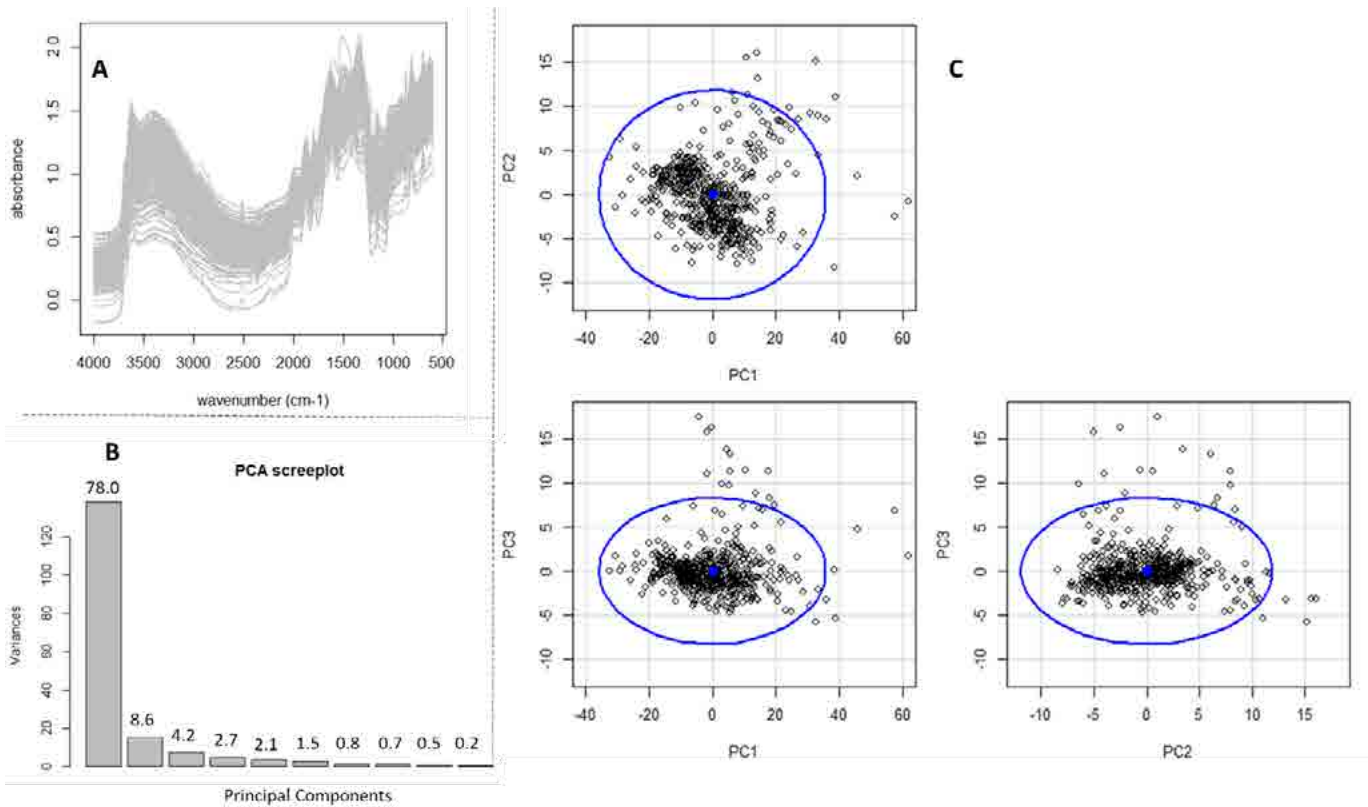


Figure 5. Illustration de l'ACP sur les données spectrales MIR du sol: (A) les données MIR du sol dans l'espace spectral; (B) le diagramme de dispersion des 10 premières composantes principales et le pourcentage de variance que chaque PC représente; et (C) le nuage de points dans l'espace défini par chaque paire des 3 premières composantes principales. Source : Auteurs du présent document

### 3.6 Évaluation du modèle

Une évaluation objective des performances des modèles vis-NIR et MIR pour prédire les propriétés de nouveaux échantillons de sol inconnus est nécessaire. À cette fin, un jeu de données test indépendant dont les propriétés du sol sont mesurées avec les méthodes de référence chimique est utilisé. Les performances sont généralement évaluées en comparant les prédictions vis-NIR ou MIR aux valeurs de référence, en calculant des statistiques telles que le biais (erreur moyenne, ME), la racine carrée de l'erreur quadratique moyenne (RMSE), le coefficient de corrélation  $r$  de Pearson et le coefficient de détermination ( $R^2$ ).

Erreur moyenne:

$$ME = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)$$

où  $y$  et  $\hat{y}$  désignent respectivement les valeurs mesurées et prédites par le modèle vis-NIR ou MIR de la propriété du sol, et  $N$  est le nombre total de valeurs mesurées.

Racine carrée de l'erreur quadratique moyenne:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

Le coefficient de corrélation  $r$  de Pearson:

$$r = \frac{\sum_{i=1}^N (y_i - \bar{y}_i)(\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}}_i)^2}}$$

Où  $\bar{y}$  est la moyenne des valeurs mesurées et  $\bar{\hat{y}}$  est la moyenne des valeurs prédites.

La ME représente le biais, et la RMSE l'ampleur de l'erreur. La RMSE est une mesure de la proximité de la prédiction par rapport à la valeur de référence pour un échantillon individuel et donc une indication de la précision de la prédiction attendue lorsque le modèle est appliqué à un échantillon de sol inconnu. Le biais, quant à lui, est une mesure de la tendance du modèle à surestimer ou sous-estimer certaines propriétés du sol. Tous deux ont une valeur optimale de 0, mais la ME peut être négative. Le coefficient de corrélation  $r$  est une mesure de la corrélation linéaire entre les valeurs prédites et les valeurs de référence, et varie de -1 (corrélation linéaire négative parfaite) à 1 (corrélation linéaire positive parfaite). L'un des inconvénients de l'utilisation de la RMSE comme

mesure de la performance du modèle est qu'elle dépend de la propriété et de l'unité. Il est donc difficile de comparer les modèles vis-NIR ou MIR entre différentes propriétés de sol et différentes études. Par exemple, un modèle vis-NIR de pH du sol a une RMSE de 0,45 unité de pH; un modèle vis-NIR de CO du sol peut avoir une RMSE de 0,35 %; mais un modèle vis-NIR d'argile a une RMSE de 5 %. Sans informations supplémentaires, il est difficile de déterminer quel modèle vis-NIR est le plus performant.

Enfin, le coefficient de détermination est estimé par:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

qui est équivalent à  $1 - \text{RMSE}^2 / \sigma^2$ , où  $\sigma^2$  est la variance des valeurs mesurées. Dans la littérature, le  $R^2$  est également appelé efficacité de modélisation ou coefficient de Nash-Sutcliffe. La valeur optimale de  $R^2$  est 1 et elle peut être négative. Lorsque  $R^2 < 0$ , cela indique que la moyenne des observations est un meilleur prédicteur que le modèle. Une valeur positive peut être interprétée comme une quantité de variance expliquée par le modèle. Par exemple, un modèle vis-NIR de pH du sol avec  $R^2 = 0,3$  signifie que 30 % de la variance des valeurs de pH mesurées est expliquée par le modèle vis-NIR. Notez que le  $R^2$  est souvent calculé comme le carré du coefficient de corrélation  $r$  de Pearson (c'est-à-dire  $r^2$ ). Le  $R^2$  et le  $r^2$  ne sont équivalents que dans le cas d'un modèle linéaire avec interception. Le  $R^2$  évalue l'écart par rapport à la ligne 1:1, tandis que le  $r^2$  évalue l'écart par rapport à la ligne de régression linéaire ajustée entre le mesuré et le prédit. Nous soulignons que pour une évaluation réaliste de la performance des modèles spectroscopiques, le coefficient de détermination devrait être calculé comme  $1 - \text{RMSE}^2 / \sigma^2$ .

Nous notons également que certaines publications évaluent le modèle sur la base du ratio de performance par rapport à l'écart (RPD), qui est un indice de performance du modèle similaire au  $R^2$ . Par ailleurs, le ratio de performance par rapport à la distance interquartile (RPIQ) a été proposé pour tenir compte des données non normales, en remplaçant l'écart type dans l'équation RPD par l'écart interquartile des données. Lors de l'évaluation d'un modèle, il n'y a pas d'étalon-type à partir duquel on dit qu'un modèle est suffisant, l'utilité et la qualité du modèle dépendent de son application.

La **figure 6** utilise des données simulées pour montrer comment les valeurs prédites vis-NIR/MIR se comparent aux valeurs mesurées en laboratoire, pour certaines valeurs de ME, RMSE,  $r$  et  $R^2$ . Notez la précision de l'ajustement des points autour de la ligne de régression et la façon dont les lignes de régression s'écartent de la ligne 1:1 (évaluée par le  $R^2$ ).

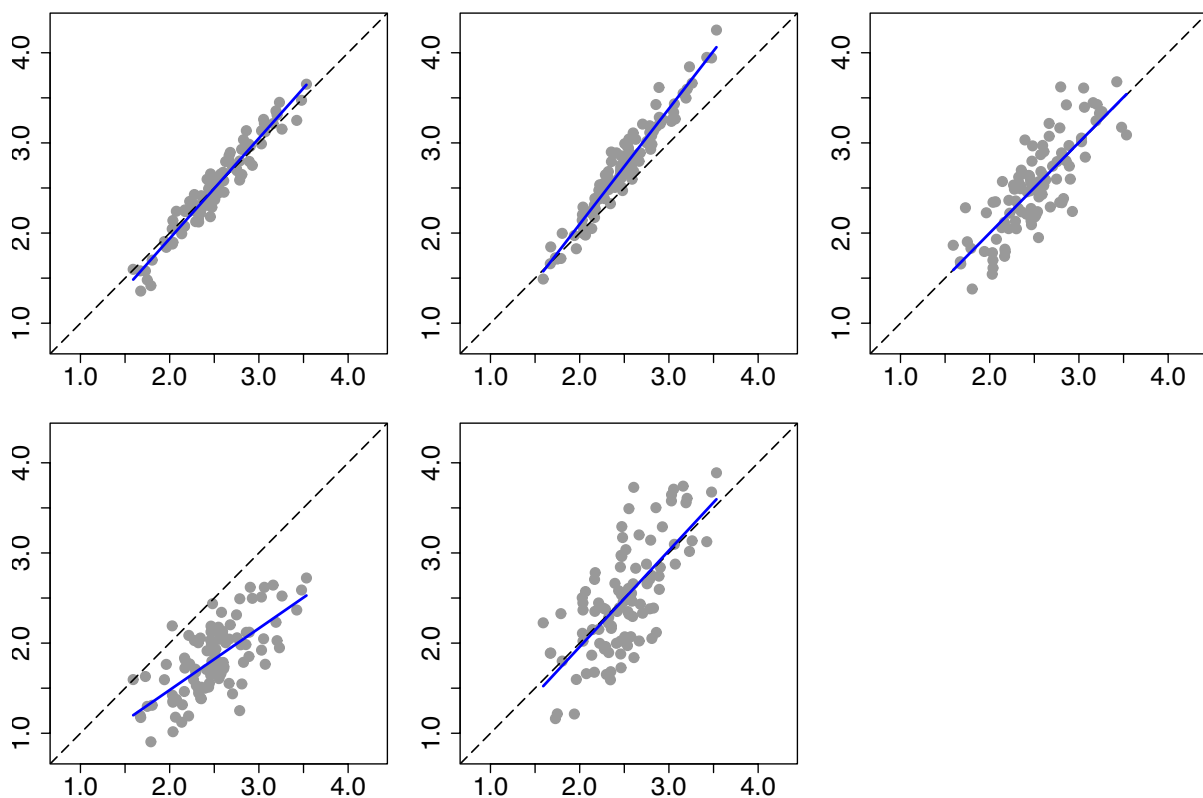


Figure 6. Données simulées pour montrer la comparaison des propriétés du sol prédites par le vis-NIR ou le MIR à celles mesurées en laboratoire dans des diagrammes de dispersion. Différents niveaux de performance du modèle sont donnés et leurs critères d'évaluation sont fournis. Source : Auteurs du présent document

Voici un exemple d'utilisation de jeux de données vis-NIR et MIR du laboratoire d'étude des sols de Kellogg de l'USDA-NRCS pour estimer trois propriétés du sol: OC, argile et pH. Le jeu de données vis-NIR contient 201 échantillons au total, et il a été divisé en 70 % pour le développement (n=141) et 30 % pour le test (n=60). Le jeu de données MIR contient 540 échantillons et a été divisé en 80 % pour le développement (n= 432) et 20 % pour le test (n=108). La modélisation PLSR a été utilisée. Les résultats des tests sont présentés sous forme de diagrammes de dispersion à la **figure 7** (pour vis-NIR) et à la **figure 8** (pour MIR); les valeurs numériques de ME, RMSE, r et R<sup>2</sup> sont indiquées au tableau 1. En outre, le tableau 1 présente également les performances des modèles de machine à vecteur de support (SVR) pour ces deux jeux de données. Le SVR est un modèle d'apprentissage automatique courant dans les études de spectroscopie du sol et nous l'avons utilisé ici à titre d'exemple. Notez que la **figure 7**, la **figure 8** et le tableau 1 sont des façons courantes de présenter les résultats de la modélisation vis- NIR/MIR.

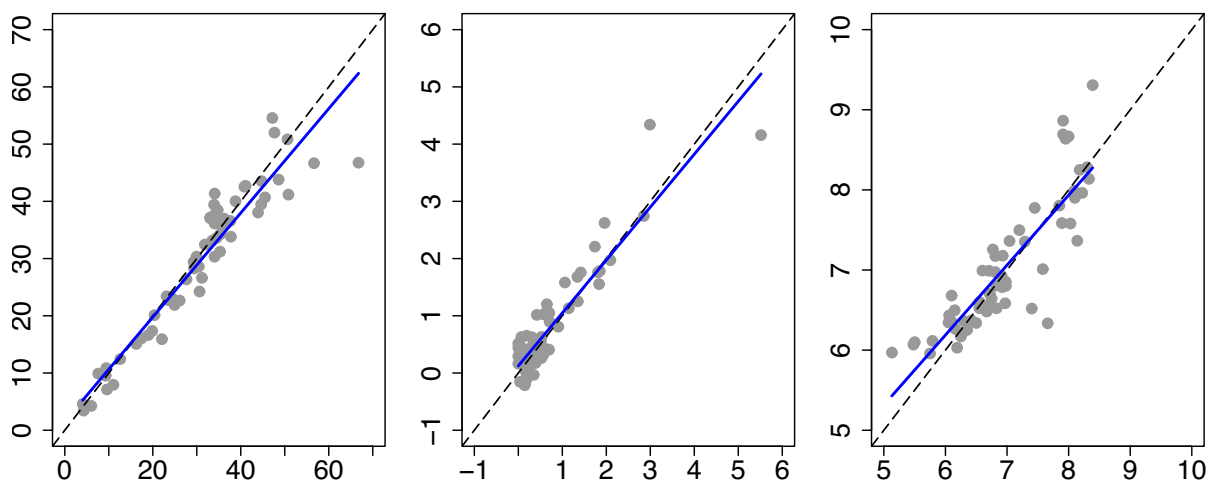


Figure 7. Diagrammes de dispersion des propriétés du sol prédites par le système vis-NIR et mesurées en laboratoire pour le jeu de données test (n = 60) à l'aide de la régression partielle par les moindres carrés. La ligne grise en pointillés est une ligne d'égalité entre le mesuré au laboratoire et le prédit, tandis que la ligne bleue est la ligne ajustée aux mesures. Source : Auteurs du présent document

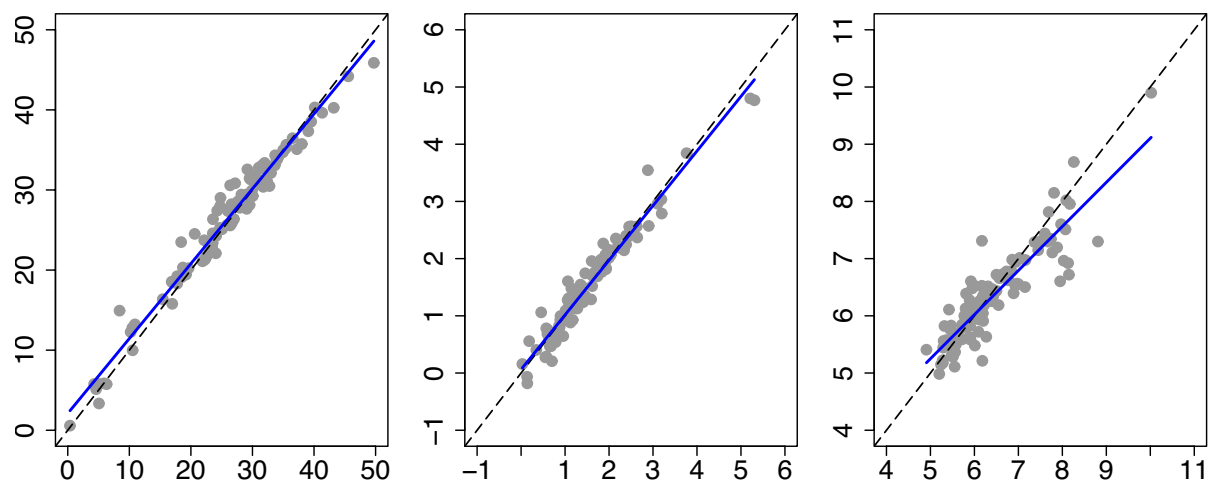


Figure 8. Diagrammes de dispersion des propriétés du sol prédites par le MIR et mesurées en laboratoire pour le jeu de données test (n = 108) en utilisant la régression partielle par les moindres carrés. Source : Auteurs du présent document

Propriété du sol vis-NIR		vis-NIR (n = 60)				MIR (n = 108)			
		ME	RMSE	r	R <sup>2</sup>	ME	RMSE	r	R <sup>2</sup>
PLSR	Argile	-1,16	4,35	0,95	0,9	0,41	1,7	0,98	0,97
	OC	0,07	0,38	0,92	0,83	-0,01	0,2	0,97	0,95
	pH	0,07	0,42	0,87	0,72	-0,09	0,43	0,89	0,79
SVR	Argile	-1,68	5,53	0,92	0,83	0,39	2,35	0,97	0,94
	OC	-0,1	0,44	0,92	0,77	0,05	0,2	0,98	0,95
	pH	-0,08	0,49	0,8	0,63	-0,02	0,44	0,88	0,77

Cours en vidéo d'analyse de données spectrales du sol (niveau I) :

**Introduction:** <https://www.youtube.com/watch?v=KoK4WVVAedI&list=PLEE7AAD33ADCF6038&index=23>

**Prétraitement spectral:** <https://www.youtube.com/watch?v=fNa-G7ziCwI&list=PLEE7AAD33ADCF6038&index=29>

**Analyse exploratoire de spectres:** <https://www.youtube.com/watch?v=74RiPUoG6-8&list=PLEE7AAD33ADCF6038&index=29>

**Sélection des échantillons pour les analyses labos:** [https://www.youtube.com/watch?v=xqKpYE\\_Y6b8&list=PLEE7AAD33ADCF6038&index=27](https://www.youtube.com/watch?v=xqKpYE_Y6b8&list=PLEE7AAD33ADCF6038&index=27)

**Estimation des propriétés des sols à partir des spectres:** <https://www.youtube.com/watch?v=4BmFzEyQ4DA&list=PLEE7AAD33ADCF6038&index=26>

## 4 | Les bibliothèques spectrales vis-NIR et MIR du sol à l'échelle régionale, continentale et mondiale: motivations, avantages et réserves

La partie la plus coûteuse de la spectroscopie vis-NIR et MIR pour l'analyse des sols est l'acquisition d'un jeu de données pour le développement du modèle, qui doit non seulement être scanné par les instruments vis-NIR et MIR, mais aussi être analysé avec une méthode de laboratoire de référence. Dans le cadre de la recherche, une approche commune consiste à diviser l'ensemble des données en un jeu pour le développement du modèle et un jeu pour le test du modèle. Cette approche n'est cependant ni pratique ni économique pour chaque projet individuel. Une approche plus viable serait de développer des bibliothèques spectrales de sols qui peuvent être partagées et utilisées par de nombreux utilisateurs et projets. Cela motive essentiellement le développement de bibliothèques spectrales vis-NIR et MIR du sol à l'échelle régionale, nationale, continentale et mondiale. Il y a plusieurs façons de construire de telles bibliothèques spectrales. Tout d'abord, les échantillons de sol existants peuvent être extraits des archives pédologiques et scannés avec un instrument vis-NIR et/ou MIR. Puisque les échantillons de sol sont déjà séchés, broyés et que certaines données chimiques sont déjà disponibles, cela représente un moyen moins coûteux de construire une bibliothèque spectrale. Un léger inconvénient des anciens échantillons de sol est que toutes les propriétés du sol ne sont pas disponibles ou analysées avec les méthodes de référence souhaitées. Dans ce cas, une nouvelle analyse de certains échantillons anciens peut être effectuée. Il est également possible de concevoir et de mettre en œuvre des campagnes d'échantillonnage pour collecter de nouveaux échantillons d'une manière plus systématique et contrôlée afin de constituer une bibliothèque spectrale des sols. Les bibliothèques spectrales vis-NIR et MIR du sol contiennent souvent un grand nombre d'échantillons de sol (par exemple des dizaines de milliers), couvrent de grandes régions géographiques et fusionnent probablement des échantillons provenant de différentes campagnes et projets d'échantillonnage. Ces facteurs entraînent des problèmes supplémentaires concernant le développement et l'utilisation des bibliothèques spectrales. Tout d'abord, la même propriété du sol peut être mesurée par différentes méthodes de laboratoire de référence. Par exemple, la «combustion sèche» et le «Walkley-Black» sont utilisés comme méthodes de référence pour le carbone organique du sol. La teneur en carbone organique du sol mesurée par «Walkley-Black» peut être utilisée pour estimer la matière organique du sol en utilisant un facteur de conversion générique de 1,724. Mais, Walkley-Black est basé sur une

oxydation utilisant une solution de  $K_2Cr_2O_7$  et peut ne pas oxyder tout le CO. Par conséquent, le véritable facteur de conversion peut être spécifique au sol. D'autre part, la combustion sèche mesure toutes les formes de carbone (y compris le C inorganique). Le pH du sol peut être mesuré dans l'eau, dans une solution de KCl ou de  $CaCl_2$  à différents rapports sol/solution; et diverses techniques d'extraction sont utilisées pour les nutriments du sol comme le phosphore et le potassium. Si des échantillons de sol mesurés par différentes méthodes de référence ou provenant de différents laboratoires avec des protocoles différents sont introduits dans la même bibliothèque spectrale, la base de données sera incohérente. Si elle n'est pas traitée correctement, cette incohérence sera amplifiée lors de la modélisation spectroscopique et donnera lieu à de mauvaises performances du modèle.

Un deuxième problème des bibliothèques spectrales à grande échelle réside dans la très grande variabilité des échantillons de sol (matériaux parentaux, suites minéralogiques, propriétés physiques et chimiques). Cette variabilité entraîne également une grande variabilité des données spectrales vis-NIR et MIR. D'autre part, en ce qui concerne les applications, le besoin d'estimer les propriétés du sol se produit souvent à l'échelle locale (par exemple, dans l'agriculture de précision pour estimer la fertilité du sol et sa variation dans le champ). Pour utiliser plus efficacement les bibliothèques spectrales, il est nécessaire de choisir et d'optimiser le sous-ensemble d'échantillons pour la calibration du modèle. Pour résoudre ce problème, des méthodes d'apprentissage automatique plus avancées (par exemple, la recherche de similarité spectrale et l'apprentissage basé sur la mémoire) et des schémas de modélisation (par exemple, l'utilisation d'un poids supplémentaire) sont développés pour améliorer la performance de la prédiction des propriétés du sol en utilisant des spectres qui sont spectralement similaires, au lieu d'utiliser toutes les données spectrales d'une bibliothèque. Il est également courant de tester si le nouvel échantillon que l'on souhaite prédire se situe effectivement dans l'espace couvert par les spectres de la bibliothèque. S'il est déterminé que le nouvel échantillon ne tombe pas dans l'espace de la bibliothèque, alors cet échantillon ne doit pas être prédit par la bibliothèque mais passer par l'analyse traditionnelle en laboratoire. Cela peut être évalué en utilisant la technique de l'ACP et en dessinant une enveloppe convexe autour de l'espace couvert par les spectres de la bibliothèque. Une telle méthode a été récemment décrite dans Wadoux *et al.*, 2021.

Un troisième problème est lié à la variabilité introduite dans les bibliothèques spectrales à grande échelle par l'utilisation de différents instruments vis-NIR et MIR. Les laboratoires participants qui contribuent à la bibliothèque sont susceptibles d'utiliser différents fabricants d'instruments ou différents modèles du même fabricant d'instruments. Les protocoles selon lesquels les instruments sont utilisés et les échantillons de sol sont présentés aux capteurs, ainsi que l'environnement du laboratoire (par exemple, la température, l'humidité et la ventilation) dans lequel les instruments sont situés, peuvent introduire des différences significatives dans les données spectrales (Ge, Thomasson et Sui, 2011). L'élaboration de modes opératoires standards pour le fonctionnement des instruments et l'acquisition spectrale, à partager entre les laboratoires, aidera à harmoniser les données spectrales provenant de diverses sources. Le transfert de calibrage (Feudale *et al.*, 2002), qui permet aux modèles calibrés sur un instrument d'être utilisés pour les données spectrales d'un autre instrument, devrait également être envisagé pour atténuer ce problème. Malgré les mises en garde mentionnées ci-dessus, ces bibliothèques spectrales représentent une avancée majeure dans l'obtention de données pédologiques quantifiables et peu coûteuses pour des applications telles que l'agriculture de précision et la séquestration du carbone dans le sol. Par exemple, les agriculteurs ont besoin de développer des cartes de fertilité et de pH du sol pour guider l'application à taux variable d'engrais et de chaux par le biais d'une collecte intensive d'échantillons de sol (par exemple, un échantillon par acre ou 2,5 échantillons par hectare). Pour un grand champ, le nombre total d'échantillons de sol serait de quelques centaines. Avec les bibliothèques spectrales vis-NIR et MIR, il suffit de scanner les échantillons avec un spectromètre vis-NIR ou MIR et les propriétés du sol comme la matière organique et le pH peuvent être estimées à partir des modèles existants. Cet effort permet d'économiser considérablement le temps et les coûts associés à la préparation et à l'analyse des échantillons. Dans le tableau 2, nous avons résumé les bibliothèques spectrales vis-NIR et MIR publiées à grande échelle pour les sols aux échelles régionale, continentale et mondiale.

Tableau 2. Résumé des bibliothèques spectrales vis-NIR et MIR du sol publiées aux niveaux national, continental et mondial. Source : Auteurs du présent document

Auteurs et année	Échelle	Nombre d'échantillons	Données de laboratoire rapportées	Notes
<b>Vis-NIR</b>				
Brown <i>et al.</i> , 2006	Global	3,768 des États-Unis et 416 du monde entier	Argile, CO, Cl, Fe, CEC, classes de minéralogie des argiles	Echantillons d'archives entre 1988 et 1999 caractérisés par US National Soil Survey Center - Soil Survey Laboratory; ASD FieldSpec
Brodský <i>et al.</i> , 2011	National (tchèque)	5,223	pH, CEC, TC, Cl, Mehlich-3 P, K, Ca, Mg	Echantillons d'archives du département de la science du sol et de la protection du sol de l'Université tchèque des sciences de la vie de Prague; ASD FieldSpec 3
Viscarra Rosset Webster, 2012	National (Australie)	21,493	24 propriétés du sol	Principalement CSRIO National Soil Archive, daté jusqu'à 70 ans; ASD LabSpec
Stevens <i>et al.</i> , 2013	Continental (Europe)	~20,000	Argile, limon, sable, pH, CEC, CO, Cl, Nt, P, K	23 pays de l'UE, échantillons collectés dans le cadre de l'enquête LUCAS de 2008 à 2012; échantillons de surface (0-30 cm); systèmes NIR FOSS
Araújo <i>et al.</i> , 2014	National (Brésil)	7,172	MO, argile, limon, sable	Université de Sao Paulo; échantillons de profil; ASD Field Spec Pro FR
Shi <i>et al.</i> , 2014	National (Chine)	1,581	MO	Quatorze provinces en Chine; échantillons de surface (0-20cm); ASD FieldSpec Pro FR
Clairotte <i>et al.</i> , 2016	National (français)	~3,800	CO	Echantillons provenant d'une grille de parcelles sur le territoire de France métropolitaine, Deux profondeurs (0-30, 30-50 cm); ASD LabSpec 2500

Auteurs et année	Échelle	Nombre d'échantillons	Données de laboratoire rapportées	Notes
<b>Vis-NIR</b>				
Wijewardane <i>et al.</i> , 2016	National (U.S.)	~20,000	CO, TC	Projet d'évaluation rapide du carbone par l'USDA-NRCS; échantillons de profil; ASD LabSpec
Viscarra Rossel <i>et al.</i> , 2016	Global	23,631	CO, CI, pH, CEC, Fe, argile, limon, sable	Échantillons provenant de 92 pays du monde entier; séries ASD (FieldSpec, AgriSpec, TerraSpec, LabSpec)
<b>MIR</b>				
Terhoeven-Urselmans <i>et al.</i> , 2010	Global	971	pH, CO, CEC, Mg, argile, sable, Ca	Échantillons de 18 pays, Centre international de référence et d'information pédologique; échantillons de profil; Bruker Tensor 27 FTIR
Viscarra Rossel <i>et al.</i> , 2008	National (Australie)	1,878 (213 avec données de laboratoire)	pH, EC, CEC, CO, K, Na, Mg, Ca, argile, limon, sable	Échantillons historiques, quatre grandes régions productrices de coton; échantillons de profil; Bruker Tensor 37 FTIR
Grinand <i>et al.</i> , 2012	National (France)	~2,000	CO, CI	Échantillons recueillis de 2002 à 2009; échantillons de surface (0-30 cm); Thermo Nicolet 6700
Clairotte <i>et al.</i> , 2016	National (France)	~3,800	CO	Échantillons provenant d'une grille de parcelles sur le territoire de France métropolitaine; Deux profondeurs (0-30, 30-50 cm); Thermo Nicolet 6700
Wijewardane <i>et al.</i> , 2018	National (U.S.)	~20,000	CO, CI, TC, Nt, TS, argile, limon, sable, CEC, K, P, pH	USDA-NRCS National Soil Survey Center, échantillons collectés de 2001 à 2018; échantillons de profil; Bruker Vertex 70 FTIR



## 5 | Instruments communs pour l'acquisition spectrale par Vis-NIR et MIR du sol

Si, en principe, tous les spectromètres dotés d'un accessoire de réflectance diffuse peuvent être utilisés, la communauté des chercheurs en science du sol partage certaines marques d'instruments favorables à l'analyse des sols. Pour le vis-NIR, les spectroradiomètres de la série ASD (FieldSpec et Lab-Spec) semblent être les plus utilisés dans l'analyse des sols. L'ASD était autrefois connu sous le nom d'Analytical Spectral Devices (basé à Boulder, Colorado, USA) et fait maintenant partie de Malvern Panalytical. Les autres principaux instruments vis-NIR sont FOSS, Spectral Evolution et Thermo Nicolet (Benedetti et van Egmond, 2021).

Pour le MIR, les systèmes FTIR fabriqués par Bruker Optics (Alpha, HTS-XT) et Thermo Nicolet (6700) sont souvent utilisés pour l'analyse des sols. Agilent a commercialisé des instruments portatifs (4100 Exoscan et 4300 Handheld) qui peuvent mesurer les sols en champs (Benedetti et van Egmond, 2021).

Une liste détaillée des instruments vis-NIR et MIR courants pour l'analyse des sols figure dans une évaluation mondiale de la spectroscopie des sols récemment publiée par le Partenariat mondial pour les sols de la FAO (Benedetti et van Egmond, 2021).

Comme nous l'avons décrit précédemment, il existe de nouveaux spectromètres NIR portables à faible coût qui fonctionnent sur une certaine gamme de longueurs d'onde. Des appareils tels que le scanner Neospectra (Si-Ware), un FT-NIR d'une longueur d'onde de 1250 à 2500 nm, se sont avérés être une bonne alternative aux spectromètres de qualité « recherche ».

## 6 | Remarques finales

La spectroscopie de réflectance Vis-NIR et MIR a été proposée et étudiée depuis longtemps comme une méthode rapide, peu coûteuse et fiable pour l'analyse quantitative des sols en laboratoire. De nombreuses études de cas dans le monde entier ont montré que la spectroscopie vis-NIR et MIR peut estimer avec précision une série de propriétés physiques et chimiques du sol. Par rapport aux méthodes analytiques traditionnelles, le vis-NIR et le MIR ne nécessitent que le séchage à l'air et le broyage des échantillons, et sont donc favorables en termes de compétences requises pour la préparation des échantillons. Le fonctionnement des instruments vis-NIR et MIR suit un ensemble de procédures établies (et normalisées) et des efforts harmonisés pour garantir des mesures spectrales de haute qualité. Les données spectrales contiennent des milliers de points de données, à partir desquels des modèles statistiques doivent être développés pour estimer les propriétés du sol. Il s'agit de la différence la plus importante entre les méthodes vis-NIR/MIR et les autres méthodes d'analyse de laboratoire. C'est pourquoi le personnel de laboratoire doit être équipé d'un ensemble différent de compétences en termes d'interprétation des résultats, de vérification des erreurs et de contrôle/assurance qualité. Avec la création de grandes bibliothèques spectrales vis-NIR et MIR à l'échelle locale, régionale et continentale, un réseau mondial de laboratoires d'analyse des sols répartis dans l'espace et basés uniquement sur les méthodes vis-NIR et MIR devient réalisable. Le partage des bibliothèques spectrales et des modèles au sein du réseau de laboratoires d'analyse des sols améliorera considérablement la vitesse et réduira le coût de l'analyse des sols, tout en rendant les résultats plus comparables et reproductibles entre les laboratoires en réseau.

# Bibliographie

- Araújo, S.R., Wetterlind, J., Demattê, J.A.M. et Stenberg, B.** 2014. Amélioration de la performance de prédiction d'une grande bibliothèque de sols tropicaux spectroscopiques vis- NIR du Brésil par le regroupement en sous-ensembles plus petits ou l'utilisation de techniques de calibration de l'exploration de données. *European Journal of Soil Science*, 65(5): 718-729. <https://doi.org/10.1111/ejss.12165>
- Baumgardner, M.F., Silva, L.F., Biehl, L.L. et Stoner, E.R.** 1986. Propriétés de réflectance des sols. *Advances in Agronomy*, pp. 1-44. Elsevier. [https://doi.org/10.1016/S0065-2113\(08\)60672-0](https://doi.org/10.1016/S0065-2113(08)60672-0)
- Ben-Dor, E. et Banin, A.** 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. *Soil Science Society of America Journal*, 59(2): 364-372. <https://doi.org/10.2136/sssaj1995.03615995005900020014x>
- Benedetti, F. et van Egmond, F.** 2021. *Global Soil Spectroscopy Assessment*. FAO. <https://doi.org/10.4060/cb6265en>
- Brodský, L., Klement, A., Penížek, V., Kodešová, R. et Borůvka, L.** 2011. Construction d'une bibliothèque spectrale des sols tchèques pour la cartographie numérique quantitative des sols. *Soil and Water Research*, 6(4): 165- 172.
- Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, M.D. et Reinsch, T.G.** 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*, 132(3-4): 273-290. <https://doi.org/10.1016/j.geoderma.2005.04.025>
- Chang, C.W., Laird, D.A., Mausbach, M.J. et Hurburgh, C.R.** 2001. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65(2): 480-490.
- Clairotte, M., Grinand, C., Kouakoua, E., Thébault, A., Saby, N.P.A., Bernoux, M. et Barthès, B.G.** 2016. Calibrage national de la concentration en carbone organique des sols par spectroscopie de réflectance infrarouge diffuse. *Geoderma*, 276: 41-52. <https://doi.org/10.1016/j.geoderma.2016.04.021>
- Feudale, R.N., Woody, N.A., Tan, H., Myles, A.J., Brown, S.D. et Ferré, J.** 2002. Transfert de modèles d'étalonnage multivariés: une revue. *Chemometrics and Intelligent Laboratory Systems*, 64(2): 181-192. [https://doi.org/10.1016/S0169-7439\(02\)00085-0](https://doi.org/10.1016/S0169-7439(02)00085-0)
- Ge, Y.F., Thomasson, J.A. et Sui, R.X.** 2011. La télédétection des propriétés du sol dans l'agriculture de précision: A review. *Frontiers of Earth Science*, 5(3): 229-238. <https://doi.org/10.1007/s11707-011-0175-0>
- Grinand, C., Barthes, B.G., Brunet, D., Kouakoua, E., Arrouays, D., Jolivet, C., Caria, G. et al.** 2012. Prédiction des teneurs en carbone organique et inorganique des sols à l'échelle nationale (France) par spectroscopie de réflectance dans l'infrarouge moyen (MIRS). *European Journal of Soil Science*, 63(2): 141-151. <https://doi.org/10.1111/j.1365-2389.2012.01429.x>
- Janik, L.J., Merry, R.H. et Skjemstad, J.O.** 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture*, 38(7): 681. <https://doi.org/10.1071/EA97144>
- Minasny, B. et McBratney, A.B.** 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 94(1): 72-79. <https://doi.org/10.1016/j.chemolab.2008.06.003>
- Nawar, S., Corstanje, R., Halcro, G., Mulla, D. et Mouazen, A.M.** 2017. Délimitation des zones de gestion des sols pour la fertilisation à taux variable. *Advances in Agronomy*, p. 175-245. Elsevier. <https://doi.org/10.1016/bs.agron.2017.01.003>
- Nduwamungu, C., Ziadi, N., Tremblay, G.F. et Parent, L.-É.** 2009. Prédiction des propriétés des sols par spectroscopie de réflectance dans le proche infrarouge: Effects of Sample Cups and Preparation. *Soil Science Society of America Journal*, 73(6): 1896-1903. <https://doi.org/10.2136/sssaj2008.0213>
- Nocita, M., tevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Ben Dor, E. et al.** 2015. Chapitre 4 - La spectroscopie des sols: Une alternative à la chimie humide pour la surveillance des sols. Dans L.S. Donald, ed. *Advances in Agronomy*, pp. 139-159. Academic Press.
- Reeves, J.B.** 2010. Spectroscopie de réflectance diffuse dans l'infrarouge proche ou moyen pour l'analyse des sols, en mettant l'accent sur le carbone et l'analyse en laboratoire ou sur place: Où en sommes-nous et que reste-t-il à faire? *Geoderma*, 158(1-2): 3-14. <https://doi.org/10.1016/j.geoderma.2009.04.005>
- Savitzky, A. et Golay, M.J.E.** 1964. Smoothing + Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8): 1627-. <https://doi.org/10.1021/Ac60214a047>
- Shariffar, A., Singh, K., Jones, E., Ginting, F.I. et Minasny, B.** 2019. Évaluation d'un spectromètre NIR portable à faible coût pour la prédiction du carbone organique et total du sol en utilisant différents modèles d'étalonnage. *Soil Use and Management*, 35(4): 607-616. <https://doi.org/10.1111/sum.12537>
- Shi, Z., Wang, Q., Peng, J., Ji, W., Liu, H., Li, X. et Viscarra Rossel, R.A.** 2014. Développement d'une bibliothèque nationale de spectres de sol VNIR pour la classification des sols et la prédiction des concentrations de matière organique. *Science China Earth Sciences*, 57(7): 1671-1680. <https://doi.org/10.1007/s11430-013-4808-x>
- Smith, P., Soussana, J., Angers, D., Schipper, L., Chenu, C., Rasse, D.P., Batjes, N.H. et al.** 2020. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Global Change Biology*, 26(1): 219-241. <https://doi.org/10.1111/gcb.14815>
- Soriano-Disla, J.M., Janik, L.J., Rossel, R.A.V., Macdonald, L.M. et McLaughlin, M.J.** 2014. The Performance of Visible, Near-, and Spectroscopie de réflectance dans l'infrarouge moyen pour la prédiction des propriétés physiques, chimiques et biologiques des sols. *Applied Spectroscopy Reviews*, 49(2): 139-186. <https://doi.org/10.1080/05704928.2013.811081>
- Stevens, A., Nocita, M., Tóth, G., Montanarella, L. et van Wesemael, B.** 2013. Prédiction du carbone organique du sol à l'échelle européenne par la spectroscopie de réflectance dans le visible et le proche infrarouge. *PLoS ONE*, 8(6): e66409. <https://doi.org/10.1371/journal.pone.0066409>
- Tang, Y., Jones, E. et Minasny, B.** 2020. Évaluation des capteurs infrarouges proches portables à faible coût pour l'analyse rapide des sols du sud-est de l'Australie. *Geoderma Regional*, 20: e00240. <https://doi.org/10.1016/j.geodrs.2019.e00240>
- Terhoeven-Urselmans, T., Vagen, T.G., Spaargaren, O. et Shepherd, K.D.** 2010. Prediction of Soil Fertility Properties from a Globally Distributed Soil Mid-Infrared Spectral Library. *Soil Science Society of America Journal*, 74(5): 1792-1799. <https://doi.org/10.2136/sssaj2009.0218>

**Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Demattê, J.A.M., Shepherd, K.D., Shi, Z. et al.** 2016. Une bibliothèque spectrale globale pour caractériser les sols de la planète. *Earth-Science Reviews*, 155: 198–230. <http://dx.doi.org/10.1016/j.earscirev.2016.01.012>

**Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. et Skjemstad, J.O.** 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1-2): 59-75. <https://doi.org/10.1016/j.geoderma.2005.03.007>

**Viscarra Rossel, R.A.V., Jeon, Y.S., Odeh, I.O.A. et McBratney, A.B.** 2008. Using a legacy soil sample to develop a mid-IR spectral library. *Soil Research*, 46(1): 1. <https://doi.org/10.1071/SR07099>

**Viscarra Rossel, R.A.V. et Webster, R.** 2012. Prédiction des propriétés du sol à partir de la base de données spectroscopiques visible-néo-infrarouge du sol australien. *European Journal of Soil Science*, 63(6): 848-860. <https://doi.org/DOI.10.1111/j.1365-2389.2012.01495.x>

**Wadoux, A.M.J-C., Malone, B., Minasny, B., Fajardo, M. et McBratney, A.B.** (2021). *Soil Spectral Inference with R - Analysing Digital Soil Spectra using the R Programming Environment*. Progress in Soil Science. Springer, Cham, 310 pages.

**Wijewardane, N.K., Ge, Y., Sanderman, J. et Ferguson, R.** 2021. Un broyage fin est nécessaire pour maintenir la haute précision de la spectroscopie de réflectance diffuse dans le rouge moyen pour l'estimation des propriétés du sol. *Soil Science Society of America Journal*, 85(2): 263-272. <https://doi.org/10.1002/saj2.20194>

**Wijewardane, N.K., Ge, Y., Wills, S. et Libohova, Z.** 2018. Prédiction des propriétés physiques et chimiques des sols américains avec une bibliothèque spectrale de réflectance dans l'infrarouge moyen. *Soil Science Society of America Journal*, 82(3): 722-731.

**Wijewardane, N.K., Ge, Y., Wills, S. et Loecke, T.** 2016. Prédiction du carbone du sol dans les États-Unis continentaux: Analyse par spectroscopie de réflectance dans le visible et le proche infrarouge du projet d'évaluation rapide du carbone. *Soil Science Society of America Journal*, 80(4): 973-982. <https://doi.org/10.2136/sssaj2016.02.0052>

**Wold, S., Sjöström, M. et Eriksson, L.** 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2): 109-130. [http://dx.doi.org/10.1016/S0169-7439\(01\)00155-1](http://dx.doi.org/10.1016/S0169-7439(01)00155-1)





Le partenariat mondial sur les sols (GSP) a été établi en 2012 comme un mécanisme reconnu au niveau mondial pour positionner les sols au cœur de l'agenda mondial grâce à une action collective. Les principaux objectifs du GSP sont de promouvoir la gestion durable des sols et d'améliorer la gouvernance des sols afin de garantir des sols sains et productifs et de soutenir la provision des services essentiels des écosystèmes pour la sécurité alimentaire et une meilleure nutrition, l'adaptation au changement climatique et l'atténuation de ses effets, et le développement durable.

Land and Water Division  
GSP-secretariat@fao.org  
[www.fao.org/global-soil-partnership](http://www.fao.org/global-soil-partnership)

Food and Agriculture Organization of the United Nations  
Rome, Italy

#### **GLOSOLAN-SPEC**

Initiative du réseau mondial des laboratoires  
des sols sur la spectroscopie du sol

GLOSOLAN est un Réseau mondial des laboratoires des sols qui vise à harmoniser les méthodes et les données d'analyse des sols afin que les informations sur les sols soient comparables et interprétables dans tous les laboratoires, pays et régions.

Grâce au soutien financier de

