

USO DE LA ESPECTROSCOPIA VISIBLE E INFRARROJO CERCANO PARA ESTIMAR PROPIEDADES DE SUELO EN ARGENTINA

Daniela Ortiz^{1,2}, Juan Martín de Dios Herrero^{1,2}, Nanci Kloster^{1,2*}

¹ Instituto Nacional de Tecnología Agropecuaria, Estación Experimental Agropecuaria "Ing. Ag. Guillermo Covas", Anguil.

² Universidad Nacional de La Pampa. Argentina.

RESUMEN

El objetivo de este estudio fue desarrollar modelos de calibración por espectroscopía visible e infrarrojo cercano (Vis-NIRS) para la predicción del contenido de carbono orgánico (CO), nitrógeno total (N), arcilla+limo, y pH en suelos de Argentina, aplicando diferentes técnicas de pretratamiento matemático de los datos espectrales. Se seleccionaron 154 muestras de suelo con características físico químicas contrastantes, se secaron y tamizaron a 2 mm previo al análisis de CO, N, pH y arcilla+limo mediante métodos de referencia. Luego se obtuvo el espectro Vis-NIR (400 a 2500 nm) de cada muestra en modo reflectancia. El conjunto de muestras fue dividido aleatoriamente en dos grupos: uno para obtener el modelo de calibración (80 %) y el otro para la validación del modelo (20 %). Se utilizaron ocho técnicas de pretratamiento de la información espectral y luego se seleccionó la mejor para cada parámetro, con el criterio de obtener el mínimo error estándar de la validación cruzada (EEVC), máximo coeficiente de determinación de la validación cruzada (R^2_{cv}) y máxima desviación predictiva residual (RPD). Los modelos de calibración se obtuvieron utilizando regresión de cuadrados mínimos parciales modificada (MPLS) y validación cruzada con cinco etapas. Los mejores modelos para la predicción de CO y N se obtuvieron con los espectros sin procesar (RPD=4,69 y 3,65 respectivamente); para pH se obtuvo con la técnica variación normal estándar con segunda derivada (RPD=2,27) y para arcilla+limo con corrección de dispersión múltiple y segunda derivada (RPD=2,83). El desempeño de los modelos aplicados sobre el grupo de testeo y calibración fue similar. Los resultados indican que la elección apropiada de la técnica de pretratamiento de datos espectrales puede optimizar la calibración. Vis-NIRS es una herramienta valiosa para el monitoreo de suelos en Argentina, como complemento de los métodos tradicionales de análisis.

Palabras clave: modelos de calibración local; técnicas de pretratamiento; nitrógeno; arcilla + limo; carbono orgánico; pH del suelo.

USE OF VISIBLE AND NEAR INFRARED REFLECTANCE SPECTROSCOPY TO ESTIMATE SOIL PROPERTIES IN ARGENTINA

ABSTRACT

The aim of this study was to develop visible and near-infrared spectroscopy (Vis-NIRS) calibration models for predicting the content of organic carbon (OC), total nitrogen (N), clay + silt, and pH values in soils from Argentina, employing different mathematical preprocessing techniques for spectral data. A total of 154 soil samples with contrasting physicochemical characteristics was selected, dried, and sieved to 2 mm prior to the analysis of OC, N, pH, and clay + silt using reference methods. Subsequently, the Vis-NIR spectrum (400 to 2500 nm) of each sample was obtained in reflectance mode. The sample set was randomly divided into

* Autor de contacto:
kloster.nanci@inta.gob.ar

Recibido:
16-11-23

Recibido con revisiones:
16-04-23

Aceptado:
17-04-23

two groups: one for calibration model development (80%) and the other for model validation (20%). Eight preprocessing techniques for spectral information were employed, and the best one for each parameter studied was selected based on the criteria of achieving the minimum standard error of cross-validation (EECV), maximum coefficient of determination of cross-validation (R^2_{cv}), and maximum residual predictive deviation (RPD). Calibration models were obtained using modified partial least squares regression (MPLS) and five-fold cross-validation. The best models for predicting OC and N were obtained with the raw spectra (RPD=4.69 and 3.65 respectively); for pH, the best model was obtained using the standard normal variate preprocessing technique with the second derivative (RPD=2.27), and for clay + silt, the best model was obtained with multiple scattering correction and second derivative (RPD=2.83). The performance of the models applied to the calibration and testing groups was similar. The results indicate that the appropriate choice of spectral data preprocessing technique can optimize calibration for the prediction. Vis-NIRS is a valuable tool for soil monitoring in Argentina, complementing traditional analysis methods.

Keywords: local calibration models; preprocessing techniques; nitrogen; clay + silt; organic carbon; soil pH

INTRODUCCIÓN

La espectroscopía visible e infrarrojo cercano (Vis-NIRS, por sus siglas en inglés visible and near infrared spectroscopy) se desarrolló como un método importante para el análisis de suelo en las últimas cuatro décadas (Food and Agriculture Organization [FAO], 2022). El espectro Vis-NIR de suelo provee una 'huella dactilar' de la composición molecular de la matriz del suelo (Viscarra Rosell et al., 2022). La materia orgánica, minerales de arcilla, óxidos de hierro, carbonatos y agua son los constituyentes primarios del suelo que causan la absorción en la región Vis-NIR. Debido a que estos constituyentes primarios se asocian a otras propiedades del suelo, Vis-NIRS es también una técnica aplicada para la medición de propiedades secundarias. Entre ellas la capacidad de intercambio catiónico, el pH, la salinidad (FAO, 2022) y la composición de la materia orgánica en el suelo (por ejemplo, el porcentaje de carbono de grupos carbonilo, aromáticos, y la relación de carbono alquílico a carbono aromático) (Terhoeven-Urselmans et al., 2006; Wetterlind et al., 2022). Actualmente la comunidad científica internacional está en discusión respecto de la capacidad de predicción de propiedades secundarias por Vis-NIRS (McBride, 2021; Viscarra Rossel et al., 2022) y como último recurso se recomienda que las predicciones de este tipo se realicen con calibraciones locales (McBride, 2021).

La tecnología Vis-NIRS tiene numerosas ventajas en comparación con los métodos tradicionales de análisis de suelo, tales como su rapidez, la no destrucción de la muestra y la caracterización simultánea de múltiples propiedades. Además, es una tecnología de bajo costo y amigable con el ambiente, debido a que requiere una mínima preparación de muestra y no utiliza reactivos químicos. Sin embargo, es importante destacar que Vis-NIRS no reemplaza los métodos tradicionales de análisis, sino que los complementa, debido a que su ejecución tiene un costo mucho menor que los métodos tradicionales. Es importante notar que se requiere una inversión inicial para adquirir el espectrómetro Vis-NIR y una base de datos de muestras de suelo analizadas por métodos de referencia de laboratorio (Ge et al., 2020). Aunque presenta importantes ventajas comparada con los métodos tradicionales de análisis, Vis-NIRS es una técnica analítica que frecuentemente requiere de la aplicación de pretratamiento a los datos espectrales, y siempre requiere del desarrollo de modelos matemáticos complejos para establecer una relación entre el espectro y las propiedades del suelo. La suma de un adecuado pretratamiento de los datos y elección de modelo multivariado, generan el modelo más robusto y exacto para la predicción de propiedades del suelo por Vis-NIRS (Carvalho et al., 2022).

El pretratamiento de los datos espectrales se realiza para eliminar información innecesaria y potenciar la capacidad predictiva de los modelos. Existen distintos pretratamientos entre los que destacan: detección de datos atípicos, selección de variables y pretratamientos matemáticos. Los datos atípicos no están dentro de la distribución normal del conjunto de calibración (Sørensen et al., 2021) y se detectan por supervisión de los espectros, conociendo los datos de manera integral para no eliminar datos extremos que podrían robustecer el modelo. La selección de variables se refiere a la selección de longitudes de onda ya que en los espectros Vis-NIR puede haber colinealidad y redundancia. La aplicación de los pretratamientos matemáticos, previo al modelado se considera un paso crítico para solucionar las variaciones espectrales sistemáticas no deseadas o distorsiones del espectro debidas a la distorsión de la luz. Los pretratamientos matemáticos más utilizados son los multiplicativos o de corrección de la dispersión. Uno de los más difundidos es SNV (Standard Normal Variate, por sus siglas en inglés), se utiliza para corregir variaciones de la línea base producida por la dispersión. Con este pretratamiento se centra cada espectro, restando el valor

medio de la absorbancia del espectro a los valores de absorbancia de cada longitud de onda, dividido todo ello por la desviación estándar. El método SNV suele aplicarse paralelamente al de remoción de tendencias (Detrend, D). Éste se trata de la aplicación de funciones polinómicas, usualmente de segundo grado, para eliminar tendencias curvilíneas que pueden tener los espectros, a medida que se aproximan a las longitudes de onda de las vibraciones fundamentales entre 1100 y 2500 nm (Barnes et al., 1989). Otro método, denominado MSC (Multiplicative Scatter Correction, por sus siglas en inglés), corrige los desplazamientos entre muestras aplicando a los espectros originales los parámetros de la recta, obtenida por mínimos cuadrados, representando los valores de absorbancia de un intervalo de longitudes de onda en el que no absorba el analito frente a los valores de absorbancia de un espectro de referencia (Naes et al., 2002; Sørensen et al. 2021). Finalmente, los métodos que emplean derivadas espectrales, de primer o segundo orden, tienen la capacidad de eliminar tanto los efectos aditivos como los multiplicativos en los espectros. La primera derivada se estima como la diferencia entre dos puntos de medición espectral consecutivos y elimina los desplazamientos de la línea base constantes. Mientras que la segunda derivada se estima mediante el cálculo de la diferencia entre dos puntos sucesivos del espectro de la primera derivada y elimina los desplazamientos que varían linealmente con la longitud de onda (Shenk et al., 2001). Las derivadas por el método de Savitzky-Golay (SG) incorporan un suavizado, por lo que el tamaño de ventana es un parámetro crítico. La notación de estas transformaciones se indica con cuatro dígitos que detallan: el orden de la derivada, la brecha sobre la cual se calcula la derivada, la cantidad de datos que se promedian o primer suavizado y el segundo suavizado (Naes et al., 2002; Olivieri, A. C. 2018). Las técnicas de pretratamiento más utilizadas en espectroscopía de suelo son SNV, D, MSC y SG (FAO, 2022).

En los últimos años incrementó el interés por el uso de grandes librerías de espectros Vis-NIR de suelos para el análisis rápido y otras aplicaciones en el campo de la ciencia del suelo. Sin embargo, el uso de estas librerías con bases de datos de suelos muy diversos frecuentemente resulta en discrepancias en la predicción de resultados a escala local (Chang et al., 2005; Milos et al., 2022; St. Luce et al., 2022). Esta es una de las principales desventajas de Vis-NIRS cuando se intenta hacer predicciones de suelos a partir de tales librerías. En cambio, los modelos desarrollados con muestras sitio-específicas, donde la variabilidad de suelos se reduce, tienden a ser más exactos en la predicción de propiedades de suelo. Ciertas prácticas agrícolas como la agricultura de precisión, que requieren la estimación de propiedades del suelo para determinar la fertilidad y su variación en el campo, requieren especialmente del desarrollo de modelos locales (FAO, 2022). Una de las últimas librerías espectrales a nivel mundial tiene espectros que corresponden a suelos de 92 países, que representan los siete continentes (Viscarra Rossel et al., 2016). Allí los países de Sudamérica están representados de forma muy variable. Brasil, por ejemplo, tiene un servicio online de espectros de suelo (BraSpecS), con 49.753 muestras de suelo con capacidad para la predicción de arcilla y materia orgánica a campo (Dematté et al., 2022). Chile utilizó la tecnología Vis-NIRS para estimar el contenido de isótopos estables de carbono en perfiles de suelo (Sepúlveda et al., 2021). Por su parte, Argentina, utilizó la tecnología Vis-NIRS en las últimas décadas para numerosas aplicaciones, como la estimación del valor nutritivo de granos (Diovisalvi et al., 2021; Juan, et al. 2016a), la calidad forrajera de pasturas (Juan et al., 2016b; Rabotnikof et al., 1995), el perfil lipídico de pechugas de pollo (Ortiz et al., 2020), el diámetro medio de las fibras en la industria textil (Amorena et al., 2021) entre otras. Sin embargo, esta tecnología no se utiliza para el análisis de suelo en Argentina, los avances en la aplicación al análisis de suelo son incipientes (Di Martino y García, 2022) y el país tiene tan sólo 77 espectros que lo representan en la librería mundial (Viscarra Rossel et al., 2016).

Los objetivos de este estudio fueron: 1) comparar diferentes técnicas de pretratamiento matemático de los espectros Vis-NIR, y 2) desarrollar los primeros modelos de calibración por Vis-NIRS para la predicción del contenido de carbono orgánico (CO), nitrógeno (N), arcilla + limo, y valor de pH en suelos de Argentina.

MATERIALES Y MÉTODOS

Muestras de suelo

Las muestras de suelo se obtuvieron a través de un trabajo colaborativo con el 'Sistema nacional para el monitoreo de la degradación de tierras' (PEi040 INTA). Se recolectaron tres submuestras de suelo del horizonte superficial (0-10 cm) y subsuperficial (10-30 cm) a partir de las cuales se obtuvo una muestra compuesta para cada profundidad en 77 sitios obteniendo un total de 154 muestras de suelo. Los suelos predominantes en el área de estudio (Figura 1) son Inceptisoles, Molisoles, Entisoles y Vertisoles (Gaitán et al., 2023).

Preparación de las muestras y análisis físico químico

Las muestras de suelo se secaron al aire y se tamizaron con malla de 2 mm, tal como lo sugiere Viscarra Rossel et al. (2016) para análisis de suelo por Vis-NIRS. Las propiedades del suelo se analizaron utilizando métodos analíticos de referencia. El CO se analizó por el método de Walkley y Black (1934), el N se analizó por el método de Kjeldahl (Bremner, 1996) y el pH se analizó en relación suelo:agua 1:2,5. Para el análisis de CO, N y pH se utilizaron procedimientos normalizados (Instituto Argentino de Normalización y Certificación, IRAM) según normas IRAM-SAGyP 29571-2:2016, 29572:2018 y 29574:2021 respectivamente. La distribución de tamaño de partículas, arcilla, limo y arena se determinó por el método del hidrómetro (Bouyoucos, 1962).



Figura 1: Área de estudio y sitios de muestreo de suelo en Argentina, América del Sur.

Figure 1: Studied area and soil sampling sites in Argentina, South America.

Obtención de espectros y pretratamiento

Los espectros se obtuvieron utilizando un espectrómetro Vis-NIR modelo DS2500 (FOSS, Hillerød, Dinamarca), en modo absorbancia como $[\log(1/R)]$, siendo R la reflectancia, utilizando una cubeta portamuestra de 7 cm de diámetro con una cantidad aproximada de 20 g de suelo para cada lectura. El espectro se registró en el rango de 400 a 2500 nm, con una resolución de 2 nm. De acuerdo con el protocolo para análisis de suelo sugerido por Viscarra Rossel et al. (2016), todas las muestras de suelo se midieron por triplicado dando origen a 3 espectros por muestra, donde cada espectro es el promedio de 49 escaneos que realiza el equipo de manera automática. El espectrómetro Vis-NIR se conectó a una computadora con el software ISISCAN NOVA™ (FOSS) para la adquisición de los espectros. Luego se utilizaron otros dos softwares para configurar la forma de trabajo del equipo y realizar la exportación de los espectros (MOSAIC), y para el pretratamiento espectral y la obtención del modelo de calibración (WinISI III). Los tres softwares mencionados son proporcionados por el fabricante del equipo y cada uno cumple funciones específicas y complementarias en el proceso de análisis y gestión de datos.

Los datos espectrales del presente estudio se trataron con diferentes técnicas para reducir el ruido aleatorio: SNV+D, MSC y derivadas por el método SG: 1,4,4,1; 2,4,4,1; 1,10,10,1 y 2,10,10,1. También se utilizaron los espectros puros $[\log(1/R)]$ para el desarrollo de las calibraciones, debido a que algunos autores hallaron que el pretratamiento no siempre mejora el desempeño del modelo (Ge et al., 2020). Para cada parámetro se seleccionó el mejor método de pretratamiento, con el criterio del mínimo error estándar de la validación cruzada (EECV) [Eq. 1], el máximo coeficiente de determinación de la validación cruzada (R^2_{cv}) [Eq. 2] y la máxima relación entre la desviación estándar del conjunto de calibración y el EECV (RPD, por sus siglas en inglés, Relative Prediction Deviation) [Eq. 3] (Seema et al., 2020).

[Eq. 1]

$$EECV = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{predicho} - y_{medido})^2}$$

[Eq. 2]

$$R^2_{cv} = \frac{\sum_{i=1}^N (y_{predicho} - y_{medido})^2}{\sum_{i=1}^N (y_{observado} - y_{medio})^2}$$

[Eq. 3]

$$RPD = \frac{DS}{EECV}$$

Donde $y_{predicho}$ representa el valor predicho, y_{medido} el valor medido, N el número total de valores medidos, y_{medio} es la media de los valores medidos y DS es el desvío estándar de los valores medidos.

Modelo de calibración y validación

En la Figura 2 se representa un esquema que resume el proceso de obtención del modelo, el cual se traduce como una función matemática con dos variables: la dependiente (Y) y la independiente (X). En aplicaciones NIRS, uno o más espectros de las muestras se organiza en forma de una matriz X_{ij} , correspondiente a valor de la absorbancia $[\log(1/R)]$ de la muestra "i" en alguna longitud de onda "j" (en nm), por lo que el dato espectral X se expresa:

[Eq. 4]

$$X = [x_{11} \ x_{12} \ x_{13} \dots \ x_{1j}]$$

La nueva matriz transpuesta comprimida en variables latentes y preprocesada se combina con los datos de referencia. Ello desemboca en un análisis de regresión, donde la calibración para un parámetro se define como:

[Eq. 5]

$$y = X\beta + \tau$$

Siendo β el vector de regresión y τ el error del modelo. Mientras que para una calibración multivariada sería:

[Eq. 6]

$$y = \tau + X_1 \beta_1 + X_2 \beta_2 + \dots + X_n \beta_n$$

Siendo n el número de variables independientes (parámetros calibrados).

Para la predicción de una muestra nueva X_2 el software comprime y aplica tratamientos matemáticos sobre el espectro de la muestra incógnita y luego lo multiplica por cada vector de regresión con el fin de predecir el parámetro calibrado.

[Eq. 7]

$$Y_{predicha} = X_2\beta$$

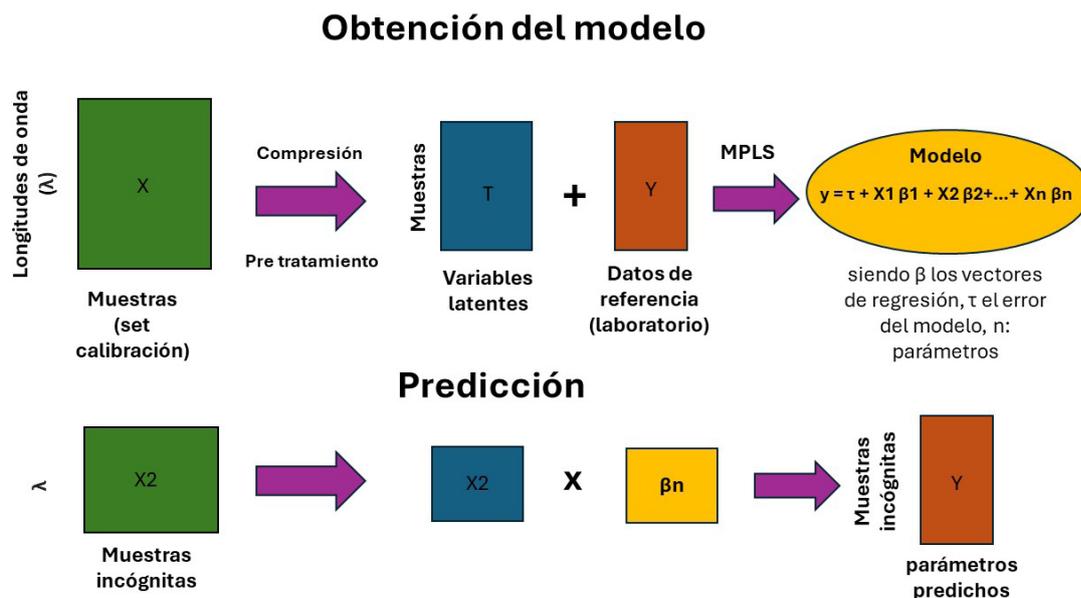


Figura 2: Esquema del proceso de obtención del modelo para la predicción de parámetros utilizando NIRS.

Figure 2: Scheme of the process of model development for parameter prediction using NIRS.

El conjunto de muestras se dividió aleatoriamente en dos grupos, utilizando la forma sugerida de división más común 80:20 (Ciurczak et al., 2021): un grupo de muestras destinadas a elaborar el modelo de calibración (80 % de las muestras) y otro grupo de muestras destinado al testeo del modelo (20 % de las muestras). El modelo de calibración se obtuvo utilizando la regresión por mínimos cuadrados parciales modificada (MPLS) y la validación cruzada con cinco etapas de validación. Se seleccionó el modelo MPLS debido a que es el análisis multivariado aplicado en la mayoría de los estudios internacionales, para la predicción de parámetros en suelo (Milinovic et al., 2023) y provee la mejor exactitud en la predicción del CO (Seema et al., 2020) y otras propiedades del suelo (Askari et al., 2015). Se definieron valores atípicos como muestras con espectros por fuera de la población espectral (outlier-H, donde el límite se estableció en un valor igual a diez) o muestras para las cuales la diferencia entre el valor de referencia y el valor predicho fue mucho mayor que el error estándar de la validación cruzada (outlier-t, donde el límite se estableció en 2,5) (WINISI software). Se realizaron dos etapas de eliminación de valores atípicos.

Una vez obtenido el modelo de calibración, se evaluó su desempeño para la predicción de propiedades del suelo sobre muestras desconocidas, es decir realizar una evaluación imparcial del rendimiento del modelo obtenido, para lo cual se utilizó el grupo de muestras de testeo. El desempeño del modelo de calibración se evaluó a partir de los resultados de EECV, R^2_{cv} y RPD. Este último indicador otorga un índice normalizado para comparar diferentes modelos Vis-NIRS para distintas variables de suelo y en diferentes estudios. El valor de RPD se categorizó siguiendo la escala propuesta por Ge et al. (2020), donde un $RPD < 2$ identifica un modelo de ajuste pobre que solo puede utilizarse con propósitos de sondeo, $2 < RPD < 3$ identifica un modelo bueno, con capacidad de predicción cuantitativa y un $RPD > 3$ identifica un modelo muy bueno.

RESULTADOS Y DISCUSIÓN

Estadística descriptiva y características espectrales de los suelos

El contenido de CO y N en las muestras de suelo varió en un rango amplio (0,4-13,3 % y 0,04-0,88 %), lo cual refleja la variabilidad de los suelos estudiados. Las muestras también mostraron un rango amplio en el valor de pH (5,2-10), contenido de arena (26-86 %) y limo (12-62 %), y un rango más acotado en el contenido de arcilla (2-28 %) (Tabla 1). En este estudio se utilizó como parámetro de estudio el contenido de arcilla + limo, debido a su correlación positiva con el contenido de CO en el suelo (Quiroga et al., 2006). Por lo que la suma de arcilla y limo también varió en un rango amplio (14-74 %). De acuerdo con el esquema de clasificación USDA (Gee y Bauder, 1986) las muestras de suelo clasificaron en los siguientes grupos texturales:

franco limoso, franco, franco arenoso y arenoso franco (Figura 3). Esta variación en los parámetros de suelo analizados se atribuye a la variación en el material parental, orden de suelo y clases texturales, así como también al uso y distintas prácticas de manejo del suelo en el área de estudio (Milos et al., 2022).

Tabla 1: Descripción estadística de parámetros del suelo: carbono orgánico (CO), nitrógeno (N), pH, arcilla, limo, arena y arcilla + limo.

Table 1: Descriptive statistics of soil parameters: organic carbon (OC), nitrogen (N), pH, clay, silt, sand, and clay+silt.

Parámetro	n	Media	DE	Mín.	Máx.
CO (%)	154	2,7	2,7	0,4	13,3
N (%)	154	0,20	0,13	0,04	0,88
pH	154	6,8	0,9	5,2	10,0
Arcilla (%)	150	10	8	2	28
Limo (%)	150	36	12	12	56
Arena (%)	150	54	18	26	86
Arcilla+limo (%)	150	46	18	14	74

n: número de muestras; Media: valor promedio de cada parámetro; DE: desvío estándar de cada medida; mín.: valor mínimo; máx.: valor máximo.

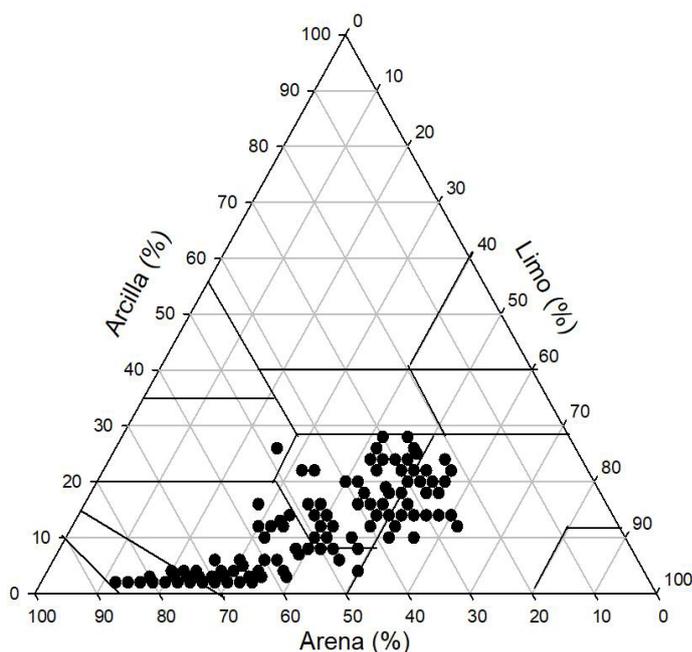


Figura 3: Distribución de tamaño de partículas, arcilla, limo y arena (%) de las muestras de suelo de acuerdo con la clasificación USDA (Gee y Bauder, 1986).

Figure 3: Particle size (clay, silt, and sand (%)) distribution of the soil samples according to USDA classification scheme (Gee y Bauder, 1986).

Los espectros Vis-NIR de las muestras de suelo registrados en el modo absorbancia como $[\log(1/R)]$, tienen una forma típica, similar a la obtenida en otros estudios de suelos (Figura 4) (Ge et al., 2020; Ludwing et al., 2018; Seema et al., 2020; Viscarra Rossel et al., 2016;). La materia orgánica, los óxidos de hierro y la

humedad de la muestra son los constituyentes principales que determinan la forma de los espectros. En la región visible del espectro (400 a 780 nm) la absorción se atribuye a grupos cromóforos del suelo tales como los óxidos de hierro y el color oscuro originado por la materia orgánica. Mientras que en la región del infrarrojo cercano (780 a 2500 nm) la absorción se atribuye a moléculas de agua, grupos hidroxilo y minerales de arcilla (FAO, 2022).

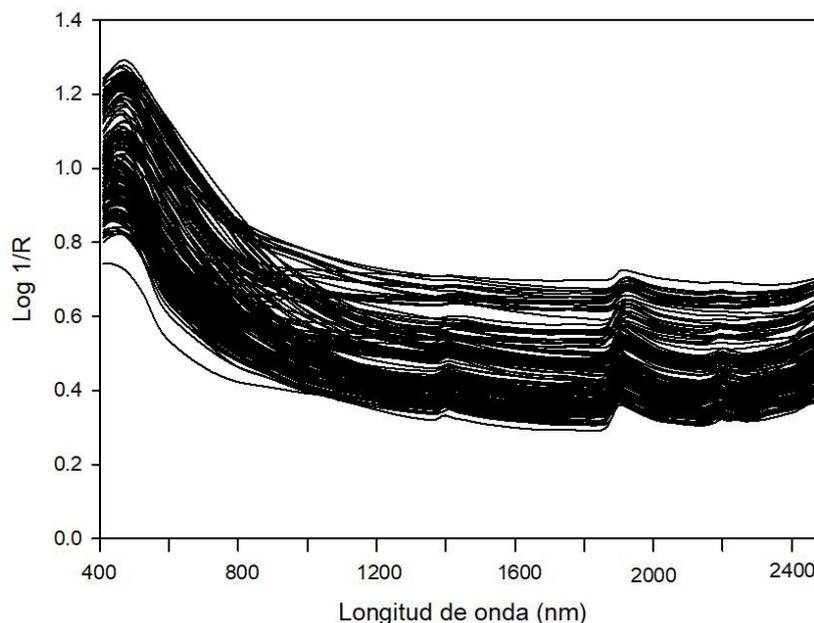


Figura 4: Espectros Vis-NIR de las muestras de suelo en modo absorbancia [$\log(1/R)$]. Se representan 154 espectros que son el promedio de las muestras medidas por triplicado.

Figure 4: Vis-NIR spectra of the soil samples in absorbance mode [$\log(1/R)$]. One hundred fifty-four spectra are represented, which are the average of the samples measured in triplicate.

Selección de la técnica de pretratamiento para cada parámetro

La aplicación de diferentes técnicas de pretratamiento tuvo diferentes efectos sobre el desempeño del modelo MPLS para cada una de las propiedades de suelo estudiadas (Tabla 2) (Milos et al., 2022). Para la estimación de CO y N se observó que la aplicación de las técnicas de pretratamiento SNV+D, MSC, primera y segunda derivada, no influyeron en las características espectrales en comparación con los espectros obtenidos en el modo absorbancia [$\log(1/R)$]. Por lo tanto, el modelo para la estimación de CO y N se elaboró a partir de los espectros sin procesar. De esta forma los resultados obtenidos por el modelo, para todas las muestras, son para CO: $R^2_{cv}=0,92$, EECV= 0,57 y RPD= 3,65 y para N: $R^2_{cv}=0,92$, EECV= 0,03 y RPD= 3,57. Seema et al. (2020) obtuvieron resultados similares para CO. Aunque otros estudios propusieron que un menor desempeño se obtuvo con los espectros sin procesar, y la aplicación de la primera derivada como técnica de pretratamiento arrojó el mejor resultado (Milos et al., 2022). Las diferencias halladas por distintos autores se atribuyen a la composición orgánica y mineral variable de los suelos estudiados (Milos et al., 2022; Moura-Bueno et al., 2020).

Para la predicción de pH, el mejor modelo se obtuvo aplicando SNV+D, con segunda derivada (2,10,10,1). Los resultados obtenidos por el modelo aplicando esta técnica de pretratamiento son $R^2_{cv}=0,82$, EECV= 0,32 y RPD= 2,36. La estimación del pH a partir del modelo es aproximadamente cuantitativa (RPD > 2), lo cual se asocia al amplio rango de pH determinado en las muestras en estudio (pH 5,2 a 10,0) que contribuye a la presencia de contenidos variables de ácidos orgánicos y carbonatos que afectan la absorbancia en el espectro Vis-NIR. En la literatura, los resultados para la estimación de pH a partir de Vis-NIRS son ambiguos. Algunos autores obtuvieron estimaciones exitosas utilizando el espectro completo (RPD = 2,24) o longitudes de onda seleccionadas (RPD = 2,49) (Wang & Wang, 2022). Los autores asociaron el éxito en la predicción de pH a la relación de este parámetro con el contenido de óxidos de hierro y materia orgánica en

suelos con altos contenidos de materia orgánica. Otros autores obtuvieron estimaciones de pH ineficaces (RPD = 1,60) (Ludwing et al., 2018), RPD=1,58 (Levi et al., 2020), RPD=1,81 (Nduwamungu et al., 2009). La menor exactitud en la estimación de pH comparada con la obtenida para CO y N se asocia a que el pH no responde a la absorción espectral en la región Vis-NIR de forma directa, sino que se encuentra dentro de las propiedades secundarias que pueden estimarse con esta tecnología.

Para la predicción de arcilla + limo, el mejor desempeño en la predicción se obtuvo aplicando MSC con segunda derivada (2,10,10,1). Los resultados obtenidos por el modelo aplicando esta técnica de pretratamiento son: R^2_{cv} = 0,85, EECV= 7,01 y RPD= 2,55. Los resultados obtenidos están en concordancia con estudios previos para la estimación de arcilla. Milos et al. (2022) indicaron un modelo obtenido utilizando primera derivada o MSC, con resultados de RPD igual a 2,2 y 1,8 respectivamente. Aunque no se hallaron en la literatura otros estudios que realicen la estimación sobre el parámetro arcilla + limo.

Tabla 2: Desempeño de diferentes técnicas de pretratamiento espectral de los parámetros carbono orgánico (CO), nitrógeno (N), pH y arcilla+limo. Los resultados en negrita indican la mejor técnica de pretratamiento para cada parámetro.

Table 2: Performance of different spectral preprocessing techniques of soil organic carbon (CO), nitrogen (N), pH and clay+silt. Bold results indicate the best preprocessing technique for each parameter.

Parámetro	Pretratamiento espectral		EECV	R^2_{cv}	RPD	
CO	Sin procesar		0,57	0,92	3,65	
	SVN+D	1,4,4,1	0,63	0,93	3,72	
		2,4,4,1	0,70	0,85	2,60	
		1,10,10,1	0,61	0,93	3,80	
		2,10,10,1	0,74	0,87	2,73	
	MSC	1,4,4,1	0,55	0,92	3,53	
		2,4,4,1	0,69	0,86	2,67	
		1,10,10,1	0,55	0,92	3,54	
		2,10,10,1	0,69	0,88	2,93	
	N	Sin procesar		0,03	0,92	3,57
		SVN+D	1,4,4,1	0,03	0,86	2,63
			2,4,4,1	0,04	0,79	2,18
1,10,10,1			0,03	0,88	2,81	
2,10,10,1			0,04	0,80	2,20	
MSC		1,4,4,1	0,04	0,83	2,44	
		2,4,4,1	0,04	0,80	2,23	
		1,10,10,1	0,04	0,83	2,40	
		2,10,10,1	0,04	0,79	2,78	
pH		Sin procesar		0,38	0,74	1,98
		SVN+D	1,4,4,1	0,33	0,80	2,24
			2,4,4,1	0,34	0,79	2,19
	1,10,10,1		0,33	0,80	2,24	
	2,10,10,1		0,32	0,82	2,36	
	MSC	1,4,4,1	0,36	0,77	2,09	
		2,4,4,1	0,32	0,82	2,34	
		1,10,10,1	0,36	0,77	2,09	
		2,10,10,1	0,33	0,80	2,26	
	Arcilla+limo	Sin procesar		7,44	0,83	2,40
		SVN+D	1,4,4,1	7,84	0,81	2,28
			2,4,4,1	8,26	0,79	2,16
1,10,10,1			7,80	0,81	2,29	
2,10,10,1			7,78	0,81	2,30	
MSC		1,4,4,1	7,55	0,82	2,37	
		2,4,4,1	7,75	0,81	2,31	
		1,10,10,1	7,56	0,82	2,37	
		2,10,10,1	7,01	0,85	2,55	

EECV: error estándar de la validación cruzada, R^2_{cv} : coeficiente de determinación de la validación cruzada, RPD: relación entre la desviación estándar y el EECV.

Desarrollo del modelo para la estimación de los parámetros estudiados

Una vez seleccionada la técnica de pretratamiento para cada parámetro se procedió a elaborar el modelo, a partir del grupo de muestras de calibración (80 % del total de muestras). El modelo para la predicción de CO y N utilizando el espectro sin procesar y MPLS arrojó los siguientes resultados: $R^2_{cv} = 0,96$, EECV= 0,46, RPD= 4,69 y $R^2_{cv} = 0,92$, EECV= 0,03, RPD= 3,55 para CO y N respectivamente (Tabla 3). Se conoce que el CO tiene una respuesta espectral directa, lo cual se relaciona con los buenos índices obtenidos para la estimación a partir de Vis- NIR. Mientras que los buenos resultados obtenidos para la estimación de N se asocian a su elevada relación con el CO del suelo (Ludwing et al., 2018). Posiblemente el modelo obtenido pueda mejorarse a partir de la incorporación de muestras de suelo con contenidos de CO mayores a 6 % (Figura 4.a). Carvalho et al. (2022) obtuvieron un desempeño poco favorable en su modelo para la predicción de CO, quizás debido al escaso número de muestras con contenidos de CO mayores a 6 %, muestras que son difíciles de cuantificar para CO, no sólo por Vis-NIRS sino también por los métodos tradicionales de análisis.

El modelo para la predicción de pH utilizando SNV+D, con segunda derivada (2,10,10,1) y MPLS arrojó los siguientes resultados: $R^2_{cv} = 0,81$, EECV= 0,34 y RPD= 2,27 (Tabla 3). Estos resultados están en línea con los informados por Ge et al. (2020), quienes obtuvieron valores de RPD= 2,4. El desempeño del modelo para la estimación del pH es menor que el obtenido para CO y N, debido a que el pH es una propiedad de tipo secundaria en su predicción a partir de Vis-NIRS, la cual se deriva de la presencia de óxidos de hierro y materia orgánica (Wang & Wang, 2022). El modelo para la predicción de arcilla + limo utilizando MSC con segunda derivada (2,10,10,1) y MPLS arrojó los siguientes resultados: $R^2_{cv} = 0,88$, EECV= 6,40 y RPD= 2,83 (Tabla 3).

Tabla 3: Desempeño del modelo seleccionado para la estimación de carbono orgánico (CO), nitrógeno (N), pH y arcilla+limo.

Table 3: Performance of the selected model for estimation of organic carbon (OC), nitrogen (N), pH and clay+silt.

Parámetro	n	Media	DE	Pretratamiento espectral		EECV	R^2_{cv}	RPD
CO (%)	116	2,26	2,17	Sin procesar		0,46	0,96	4,69
N (%)	118	0,17	0,12	Sin procesar		0,03	0,92	3,55
pH	113	6,7	0,8	SVN+D	2,10,10,1	0,34	0,81	2,27
Arcilla+limo (%)	116	46,1	18,1	MSC	2,10,10,1	6,40	0,88	2,83

Media: valor promedio de cada parámetro; DE: desvío estándar de cada medida; EECV: error estándar de la validación cruzada, R^2_{cv} : coeficiente de determinación de la validación cruzada, RPD: desviación predictiva residual, DE/EECV.

Se utilizó un conjunto de muestras independientes (20 % de las muestras, conjunto de muestras de testeo) para evaluar el desempeño del modelo obtenido, tal cual lo sugerido por Ge et al. (2020). Los resultados demostraron que el desempeño del modelo aplicado a las muestras independientes es aceptable y similar al obtenido para las muestras de calibración (Figura 5)

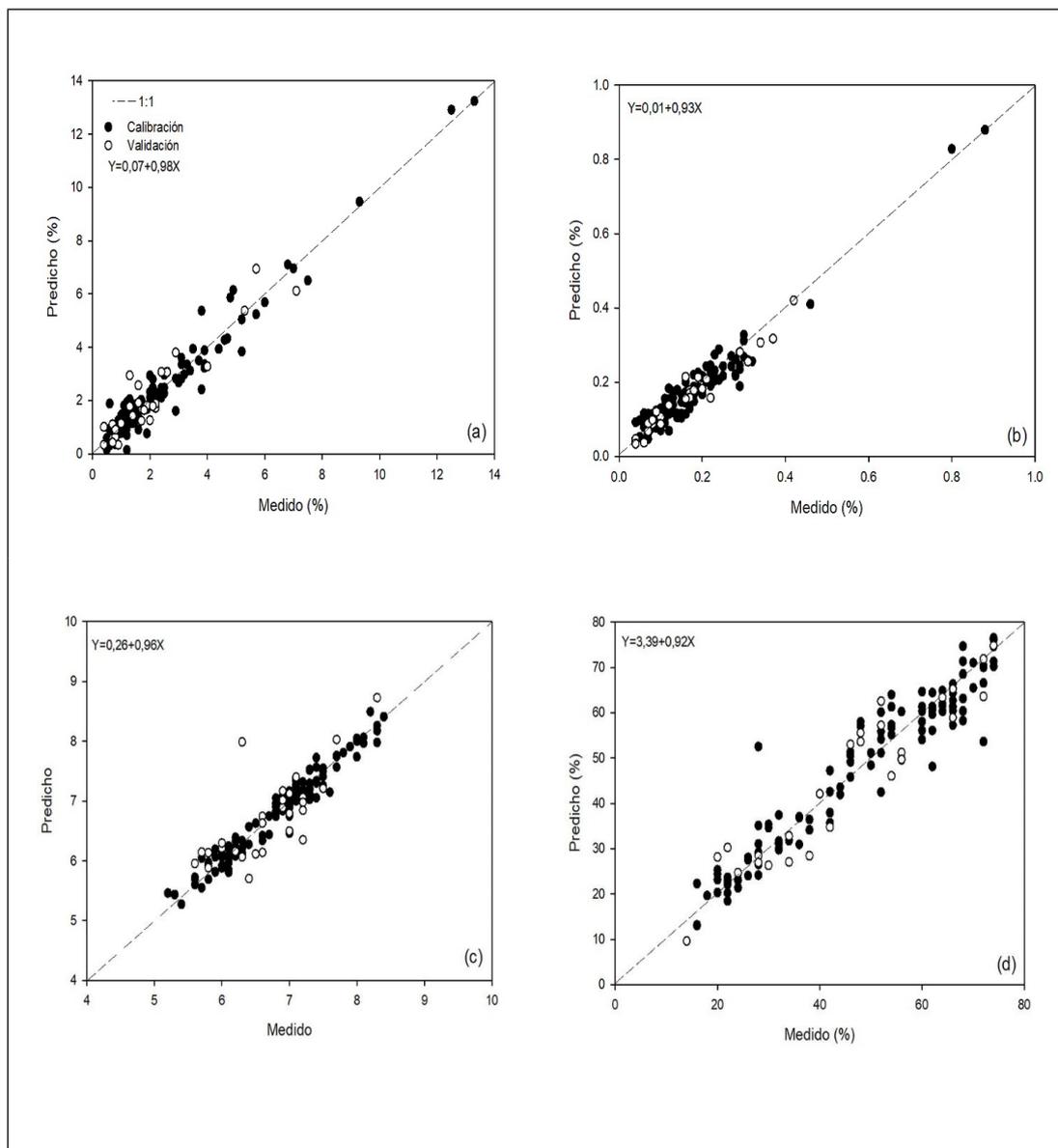


Figura 5: Relación entre valores medidos y predichos de los datos espectrales usando el modelo de calibración (círculos llenos) y de validación (círculos vacíos) para: a) carbono orgánico (CO), b) nitrógeno (N), c) pH, y d) arcilla + limo. La línea punteada representa la recta $y=x$.

Figure 5: Model performance for calibration (filled circles) and testing dataset (empty circles) to predict organic carbon (CO) (a), nitrogen (N) (b), pH (c), and clay+silt (d). The dashed line represents the $y=x$ line.

CONCLUSIONES

A partir de este trabajo se obtuvieron los primeros modelos de calibración Vis-NIRS para suelos en Argentina, que permiten estimar los parámetros CO, N, pH y arcilla + limo. Los modelos de predicción obtenidos para CO y N clasificaron como muy buenos ($RPD > 3,0$). Mientras que los modelos de predicción obtenidos para pH y arcilla + limo clasificaron como buenos, con estimación cuantitativa ($2,0 < RPD < 3,0$). Los resultados confirman la necesidad de realizar estudios preliminares para definir la mejor técnica de pretratamiento espectral de acuerdo con el parámetro estudiado. Vis-NIRS es una herramienta útil para la estimación de CO, N, pH y arcilla + limo, de forma rápida y económica, que podría utilizarse como complemento de los métodos tradicionales de análisis. Su aplicación será de gran utilidad para el monitoreo de propiedades de suelo en Argentina, tal como ocurre a nivel mundial. Aunque la transferencia de modelos Vis-NIRS, es decir que un modelo Vis-NIRS desarrollado por un laboratorio sea usado por otro laboratorio, es un área de in-

investigación a desarrollar. Hay pocos estudios que comparen el análisis Vis-NIRS de diferentes laboratorios con diferentes equipos. Este es el punto crucial para que la tecnología pase de ser de aplicación estricta a la investigación, a ser de aplicación práctica en los análisis de rutina.

AGRADECIMIENTOS

Los autores agradecen al proyecto PEi040-INTA por proveer las muestras de suelo, a M. Prost, M. Pérez, V. Colotti y R. Maldonado por su colaboración en el análisis físico y químico de las muestras, y a N. Juan por las observaciones sobre el manuscrito. Este trabajo fue financiado por INTA.

REFERENCIAS

- Amorena, J., Álvarez, D. & Fernández-Ahumada, E. (2021). Development of Calibration Models to Predict Mean Fibre Diameter in Llama (Lama glama) Fleeces with Near Infrared Spectroscopy. *Animals*, 11(7), 1998.
- Askari, M.S., O'Rourke, S.M. & Holden, N.M. (2015). Evaluation of soil quality for agricultural production using visible–near-infrared spectroscopy. *Geoderma*, 243–244, 80–91. <https://doi.org/10.1016/j.geoderma.2014.12.012>
- Barnes, R. J., Dhanoa, M. S. & Lister, S. J. (1989). Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, 43(5), 772–777.
- Bouyoucos, G.J. (1962). Hydrometer method for making particle size analysis soils. *Agronomy Journal*, 54, 464-465.
- Bremner, J.M. (1996). Total Nitrogen, in: Sparks, D. L., (ed), *Methods of Soil Analysis. Part 3: Chemical Methods*. Soil Science Society America, Madison, Wisconsin, pp. 1149-1176.
- Carvalho, J.C., Moura-Bueno, J.M., Ramon, R., Almeida, T.F., Naibo, G., Martins, A.P., Santos, L.S., Gianello, C. & Tiecher, T. (2022). Combining different pre-processing and multivariate methods for prediction of soil organic matter by near infrared spectroscopy (NIRS) in Southern Brazil. *Geoderma Regional*, 29, e00530. <https://doi.org/10.1016/j.geodrs.2022.e00530>
- Chang, C.W., Laird, D.A. & Hurburgh, C.R. (2005). Influence of soil moisture on near-infrared reflectance spectroscopic measurement of soil properties. *Soil Science*, 170, 244–255.
- Ciurczak, E. W., Igne, B., Workman Jr, J. & Burns, D. A. (Eds.). (2021). *Handbook of near-infrared analysis*. CRC press.
- Demattê, J.A.M., Paiva, A.F., Poppiel, R.R., Rosin, N.A., Ruiz, L.F.C., Mello, F.A., Minasny, B., Grunwald, S., Ge, Y., Ben Dor, E., et al. (2022). The Brazilian Soil Spectral Service (BraSpecS): A User-Friendly System for Global Soil Spectra Communication. *Remote Sensing*, 14, 740. <https://doi.org/10.3390/rs14030740>
- Di Martino, A. M. y García L. (2022). Análisis de materia orgánica en suelos por espectroscopia de infrarrojo cercano. *Revista de tecnología Agropecuaria, INTA Ediciones*, 10(41), 27-31.
- Diovisalvi, N.V., Izquierdo, N., Echeverría, H., García, F. & Reussi Calvo, N. (2021). Methods to determine nitrogen in sunflower grains. *Ciencia del Suelo*, 39(2), 217-232. <https://ojs.suelos.org.ar/index.php/cds/article/view/620>
- Food and Agriculture Organization (FAO). (2022). A primer on soil analysis using visible and near-infrared (Vis-NIR) and mid-infrared (MIR) spectroscopy, Rome. <https://doi.org/10.4060/cb9005en>
- Gaitán, J.J., Wingeyer, A.B., Peri, P., Moavro, E., Peralta, G., Fritz, F., Berhongaray, G., Adema, E., Albarracín, S., Álvarez, C., Álvarez Cortés, D.J., Bacigaluppo, S., Balducci, E., Ballón, M., Banegas, N., Barbaro, S., Barral, P., Behr, S.J., Beider, A.M., ...y Sasal, M.C. (2023). Mapa de almacenamiento de C en los suelos de la República Argentina. Asociación Argentina de Productores en Siembra Directa (Aapresid), Consorcio Regional de Experimentación Agrícola (CREA), Instituto Nacional de Tecnología Agropecuaria (INTA), Secretaría de Agricultura, Ganadería y Pesca de la Nación. Ciudad Autónoma de Buenos Aires. Argentina.
- Gee, G. & Bauder, J. (1986). Particle size analysis. In: *Methods of soil analysis, part 1, physical and mineralogical methods*, in: Klute, A., (Ed.). Soil Science Society of America. Madison, Wisconsin USA, pp. 383-411.
- Ge, Y., Morgan, C.L. & Wijewardane, N.K. (2020). Visible and near infrared reflectance spectroscopy analysis of soils. *Soil Science Society of America Journal*, 84, 1495-1502. <https://doi.org/10.1002/saj2.20158>
- Juan, N.A., Ortiz, D.A., Pordomingo, A.B. y Funaro, D.O. (2016*). Tecnología NIRS para estimar el valor nutritivo de planta entera de maíz y sorgo para silaje y sus fracciones (tallo, hoja, panoja/espiga). *Revista de la Asociación Argentina de Producción Animal*, 36(1), 230.
- Juan, N.A., Ortiz, D.A. y Ruiz, M.A. (2016b). Análisis de calidad de gramíneas forrajeras perennes de ciclo otoño-inverno-primaveral con tecnología NIRS. *Revista de la Asociación Argentina de Producción Animal*, 36(1), 229.
- Levi, N., Karnielia, A. & Paz-Kaganb, T. (2020). Using reflectance spectroscopy for detecting land-use effects on soil quality in drylands. *Soil and Tillage Research*, 199, 104571. <https://doi.org/10.1016/j.still.2020.104571>
- Ludwig, B., Murugan, R., Parama, V.R.R. & Vohland, M. (2018). Use of different chemometric approaches for an estimation of soil properties at field scale with near infrared spectroscopy. *Journal of Plant Nutrition and Soil Science*, 181, 704-713. <https://doi.org/10.1002/jpln.201800130>

- McBride, M.B. (2021). Estimating soil chemical properties by diffuse reflectance spectroscopy: promise versus reality. *Eur. J. Soil Sci.*, 73, e13192. <https://doi.org/10.1111/ejss.13192>
- Milinic, J., Vale, C. & Azenha, M. (2023). Recent advances in multivariate analysis coupled with chemical analysis for soil surveys: a review. *Journal of Soils and Sediments*, 23, 1085–1098. <https://doi.org/10.1007/s11368-022-03377-8>
- Milos, B., Bensa, B. & Japundzic-Palenkic, B. (2022). Evaluation of Vis-NIR preprocessing combined with PLS regression for estimation soil organic carbon, cation exchange capacity and clay from eastern Croatia. *Geoderma Regional*, 30, e00558. <https://doi.org/10.1016/j.geodrs.2022.e00558>
- Moura-Bueno, J.M., Dalmolin, R.S.D., Zborowski Horst-Heinen, T., ten Caten, A., Vasques, G.M., Carneletto Dotto, A. & Grunwald, S. (2020). When does stratification of a subtropical soil spectral library improve predictions of soil organic carbon content?. *Science of the Total Environment*, 737, 139895. <https://doi.org/10.1016/j.scitotenv.2020.139895>
- Næs, T., Isaksson, T., Fearn, T., & Davies, T. (2002). A user-friendly guide to multivariate calibration and classification. NIR Publications. Chichester, UK.
- Nduwamungu, C., Ziadi, N., Tremblay, G.F. & Parent, L.E. (2009). Near-Infrared Reflectance Spectroscopy Prediction of Soil Properties: Effects of Sample Cups and Preparation. *Soil Science Society of America Journal*, 73(6), 1896-1903. <https://doi.org/10.2136/sssaj2008.0213>
- Olivieri, A. C. (2018). Introduction to multivariate calibration: A practical approach. Springer. Switzerland.
- Ortiz, D.A., Camiletti, F.K., Pordomingo, A.B., Cunzolo, S.A., Pighin, D.G., Pordomingo, A.J. y Juan, N.A. (2020). Tecnología NIRS para la determinación de la composición química de pechuga de pollo. *Revista Argentina de Producción Animal*, 40(1).
- Ozaki, Y., McClure, W. F. & Christy, A. A. (Eds.). (2006). "Near-infrared spectroscopy in food science and technology". John Wiley & Sons.
- Prananto, J. A., Minasny, B., & Weaver, T. (2020). Near infrared (NIR) spectroscopy as a rapid and cost-effective method for nutrient analysis of plant leaf tissues. *Advances in agronomy*, 164, 1-49. <https://doi.org/10.1016/bs.agron.2020.06.001>
- Quiroga, A., Funaro, D., Noellemeyer, E. & Peinemann, N. (2006). Barley yield response to soil organic matter and texture in the Pampas of Argentina. *Soil and Tillage Research*, 90, 63-8. <https://doi.org/10.1016/j.still.2005.08.019>.
- Rabotnikof, C.L., Planas, G.M., Colomer, J.S. & Stritzler, N.P. (1995). Near infrared reflectance spectroscopy (NIRS) for predicting forage quality of perennial warm-season grasses in La Pampa, Argentina. *Annales de Zootechnie*, 44(1), 97-100.
- Seema, A.K., Ghosh A.K., Das B.S. & Reddy N. (2020). Application of Vis-NIR spectroscopy for estimation of soil organic carbon using different preprocessing techniques and multivariate methods in the middle Indo-Gangetic plains of India. *Geoderma Regional*, 23, e00349. <https://doi.org/10.1016/j.geodrs.2020.e00349>
- Sepúlveda, M. A., Hidalgo, M., Araya, J., Casanova, M., Muñoz, C., Doetterl, S., Wasner, D., Colpaert, B., Bodé, S., Boeckx, P. & Zagal, E. 2021. Near-infrared spectroscopy: Alternative method for assessment of stable carbon isotopes in various soil profiles in Chile. *Geoderma Regional*, 25, e00397. <https://doi.org/10.1016/j.geodrs.2021.e00397>.
- Shenk, J.S., Workman, J.J., Westerhaus, M.O., Burns, D.A. & Ciurczak, E.W. (2001). Application of NIR Spectroscopy to Agricultural Products. En *Handbook of near-infrared analysis*. M. Dekker.
- Sørensen, K.M., van den Berg, F. & Engelsen, S.B. (2021). NIR Data Exploration and Regression by Chemometrics—A Primer. In: Ozaki, Y., Huck, C., Tsuchikawa, S., Engelsen, S.B. (eds). *Near-Infrared Spectroscopy*. Springer, Singapore. https://doi.org/10.1007/978-981-15-8648-4_7
- St. Luce, M., Ziadi, N. & Viscarra Rossel, R.A. (2022). GLOBAL-LOCAL: A new approach for local predictions of soil organic carbon content using large soil spectral libraries. *Geoderma*, 425, 116048. <https://doi.org/10.1016/j.geoderma.2022.116048>
- Terhoeven-Urselmans, T., Michel K., Helfrich M., Flessa H. & Ludwig, B. (2006). Near-infrared spectroscopy can predict the composition of organic matter in soil and litter. *Journal of Plant Nutrition and Soil Science*, 169, 168-174. <https://doi.org/10.1002/jpln.200521712>
- Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Dematté, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aichi, H., Barthès, B.G., Bartholomeus, H.M., Bayer, A.D., Bernoux, M., Böttcher, K., Brodský L., Du, C.W., Chappell, A. & Ji, W. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155, 98-230. <http://dx.doi.org/10.1016/j.earscirev.2016.01.012>
- Viscarra Rossel, R., Behrens, T., Ben-Dor, E., Chabrilat, S., Melo Dematté, J. A., Ge, Y., Gomez, C., Guerrero, C., Peng, Y., Ramirez-Lopez, L., Shi, Z., Stenberg, B., Webster, R., Winowiecki, L. & Shen, Z. (2022). Diffuse reflectance spectroscopy for estimating soil properties: A technology for the 21st century. *European Journal of Soil Science*, 73 (4), e13271. <https://doi.org/10.1111/ejss.13271>
- Wang, L. & Wang, R. (2022). Determination of soil pH from Vis-NIR spectroscopy by extreme learning machine and variable selection: A case study in lime concretion black soil. *Spectrochemical Acta Part A: Molecular and Biomolecular Spectroscopy*, 283, 121707, <https://doi.org/10.1016/j.saa.2022.121707>.
- Walkley, A. & Black, I.A. (1934). An examination of the Degtjareff method for determining soil organic matter, a proposed modification of the chromic acid titration method. *Soil Science*, 37, 29-38.
- Wetterlind, J., Viscarra Rossel, R. A. & Steffens, M. (2022). Diffuse reflectance spectroscopy characterizes the functional chemistry of soil organic carbon in agricultural soils. *European Journal of Soil Science*, 73, e13263. <https://doi.org/10.1111/ejss.13263>