

Montorio Llovería, R.; Nadal Romero, E.; Pérez Cabello, F.; Badía Villas, D. Modelado de las fracciones de carbono orgánico del suelo mediante espectroscopía vis-SWIR

# Modelado de las fracciones de carbono orgánico del suelo mediante espectroscopía vis-SWIR

Montorio Llovería, Raquel <sup>1</sup> Nadal Romero, Estela <sup>2</sup> Pérez Cabello, Fernando <sup>1</sup> Badía Villas, David <sup>1</sup>

<sup>1</sup> Universidad de Zaragoza, España

<sup>2</sup> Instituto Pirenaico de Ecología (IPE-CSIC), España

ORCID: Montorio Llovería 0000-0001-7403-1764 Nadal Romero 0000-0002-4651-7828 Pérez Cabello 0000-0003-4831-4060 Badía Villas 0000-0002-1006-5030

Correspondencia: [montorio@unizar.es](mailto:montorio@unizar.es) [estelanr@ipe.csic.es](mailto:estelanr@ipe.csic.es) [fcabello@unizar.es](mailto:fcabello@unizar.es) [badia@unizar.es](mailto:badia@unizar.es)

## RESUMEN

El estudio de la dinámica del carbono orgánico del suelo (COS) en los ecosistemas requiere de una correcta evaluación de su distribución en diferentes fracciones (lábil y estable), por cuanto estas determinan su potencial como sumidero de carbono y permiten valorar la efectividad de las actuaciones de gestión. Aunque la espectroscopía vis-SWIR ha demostrado su potencial para estimar muchas propiedades del suelo, entre ellas el COS total, su uso para estimar las fracciones de carbono es todavía incipiente y son necesarios más estudios en esta línea. En este trabajo se evalúa la capacidad de modelos de regresión PLSR para estimar el COS total y sus fracciones (libre, ocluida y estable), obtenidas por técnicas densimétricas de laboratorio. Se analizaron 57 muestras de suelo procedentes de áreas de cultivo abandonadas en el Pirineo Central. Los resultados indican que el pretratamiento de los espectros combinando el cálculo de la primera derivada y la normalización SNV es el más efectivo. Los modelos lograron estimar con alta precisión el COS total y la fracción estable ( $R^2_{CV} > 0,8$ ). Por el contrario, las fracciones lábil —libre y ocluida— presentaron un ajuste moderado ( $0,6 < R^2_{CV} < 0,7$ ), que mejoró al combinarse en una única categoría ( $R^2_{CV} = 0,76$ ). Aunque se considera necesario extender esta metodología a nuevas muestras para su validación definitiva, los resultados confirman su potencial para caracterizar la estabilidad del carbono de forma robusta.

**Palabras clave:** reservas de carbono, fraccionamiento densimétrico, espectroscopía de reflectividad difusa, regresión de mínimos cuadrados parciales


Fecha de recepción: 16 febrero 2026 · Fecha de aceptación: 16 febrero 2026

# Modelado de las fracciones de carbono orgánico del suelo mediante espectroscopía vis-SWIR


Montorio Llovería, Raquel <sup>(1)</sup>, Nadal Romero, Estela <sup>(2)</sup>, Pérez Cabello, Fernando <sup>(1)</sup>, Badía Villas, David <sup>(1)</sup>

<sup>(1)</sup> Universidad de Zaragoza, España.

 0000-0001-7403-1764, montorio@unizar.es ;  0000-0003-4831-4060, fcabello@unizar.es

 0000-0002-1006-5030, badia@unizar.es.

<sup>(2)</sup> Instituto Pirenaico de Ecología (IPE-CSIC), España.

 0000-0002-4651-7828, estelanr@ipe.csic.es

**Resumen:** El estudio de la dinámica del carbono orgánico del suelo (COS) en los ecosistemas requiere de una correcta evaluación de su distribución en diferentes fracciones (lábil y estable), por cuanto estas determinan su potencial como sumidero de carbono y permiten valorar la efectividad de las actuaciones de gestión. Aunque la espectroscopía vis-SWIR ha demostrado su potencial para estimar muchas propiedades del suelo, entre ellas el COS total, su uso para estimar las fracciones de carbono es todavía incipiente y son necesarios más estudios en esta línea. En este trabajo se evalúa la capacidad de modelos de regresión PLSR para estimar el COS total y sus fracciones (libre, ocluida y estable), obtenidas por técnicas densimétricas de laboratorio. Se analizaron 57 muestras de suelo procedentes de áreas de cultivo abandonadas en el Pirineo Central. Los resultados indican que el pretratamiento de los espectros combinando el cálculo de la primera derivada y la normalización SNV es el más efectivo. Los modelos lograron estimar con alta precisión el COS total y la fracción estable ( $R^2CV > 0,8$ ). Por el contrario, las fracciones lábil —libre y ocluida— presentaron un ajuste moderado ( $0,6 < R^2CV < 0,7$ ), que mejoró al combinarse en una única categoría ( $R^2CV = 0,76$ ). Aunque se considera necesario extender esta metodología a nuevas muestras para su validación definitiva, los resultados confirman su potencial para caracterizar la estabilidad del carbono de forma robusta.

**Palabras clave:** reservas de carbono, fraccionamiento densimétrico, espectroscopía de reflectividad difusa, regresión de mínimos cuadrados parciales

## ***Modelling soil organic carbon fractions using vis-SWIR spectroscopy***

**Abstract:** *The study of soil organic carbon (SOC) dynamics in ecosystems requires a proper assessment of its distribution across different fractions (labile and stable), as these determine its potential as a carbon sink and allow for evaluating the effectiveness of management actions. Although vis-SWIR spectroscopy has demonstrated its potential to estimate many soil properties, including total SOC, its use for estimating carbon fractions is still incipient, and more research is needed in this line. This study evaluates the capacity of PLSR models to estimate total SOC and its fractions (free, occluded, and heavy), obtained through laboratory densimetric methods. A total of 57 soil samples from abandoned croplands in the Central Pyrenees were analyzed. Results indicate that pretreatment combining the first derivative and SNV normalization is the most effective. The models were able to estimate total SOC and the stable fraction with high precision ( $R^2CV > 0.8$ ). Conversely, free and occluded light fractions showed a moderate fit ( $0.6 < R^2CV < 0.7$ ), which improved when combined into a labile pool ( $R^2CV = 0.76$ ). Although it is necessary to extend this methodology to other samples and regions for its definitive validation, the results confirm its potential for a robust characterization of carbon stability.*

**Keywords:** *carbon pools, density fractionation, diffuse reflectance spectroscopy, partial least squares regression*

## **1. INTRODUCCIÓN**

El carbono orgánico (COS) es un componente fundamental del suelo, esencial para la funcionalidad del

ecosistema y la provisión de servicios (Lal, 2016). En la actualidad, el mantenimiento y la mejora del contenido de COS mediante la adopción de prácticas de gestión recomendadas es crucial para aspectos tan relevantes

como la seguridad alimentaria, la mitigación del cambio climático y otros retos reconocidos en los Objetivos de Desarrollo Sostenible (Keesstra *et al.*, 2016).

El COS varía en función de factores ambientales (p.ej. clima, tipo de suelo o vegetación), perturbaciones (p.ej. incendios forestales) y factores inducidos por el hombre (p.ej. cambios de uso y cobertura y su gestión), afectando de forma distinta a sus componentes (O'Rourke *et al.*, 2015). Así, mientras que la fracción lábil, con tasas de renovación cortas, responde rápidamente, la fracción estabilizada determina el almacenamiento a largo plazo. El balance entre ambas determina el papel del suelo como potencial sumidero de carbono, una información muy valiosa para el desarrollo de políticas de gestión. En consecuencia, una monitorización del COS debe incluir una evaluación cuantitativa de su composición en fracciones o reservorios.

Frente a los métodos químicos convencionales —complejos, costosos y lentos—, la espectroscopía de reflectancia difusa (DRS, por sus siglas en inglés) ha demostrado ampliamente su potencial para estimar el COS con gran precisión de una manera rentable y rápida (Demattê *et al.*, 2019). Sin embargo, la estimación de las fracciones de COS se encuentra todavía en una fase inicial y son pocos los estudios que la han abordado con datos en el rango vis-SWIR.

El objetivo de este trabajo es evaluar el potencial de la espectroscopía vis-SWIR para la estimación de los contenidos de COS total y su distribución en fracciones (lábil y estable) mediante modelos de regresión aplicados a los espectros de suelos no fraccionados, analizando además el impacto de diferentes técnicas de pretratamiento espectral en la precisión de los modelos.

## 2. MATERIAL Y MÉTODOS

### 2.1. Área de estudio y muestreo

El estudio analiza un conjunto representativo de 57 muestras de suelo superficial (0-5 cm) procedentes de dos zonas del Pirineo Central español: la cuenca de Aragón y el municipio de Bestué, situadas entre los 800 y 1200 m.s.n.m. Las zonas presentan un clima submediterráneo y una cobertura vegetal caracterizada por un mosaico de usos del suelo tras el abandono agrícola: repoblaciones forestales (*Pinus nigra* y *Pinus sylvestris*), matorral de recolonización natural (*Genista scorpius*, *Buxus sempervirens*), pastizales y zonas de cárcavas. Los suelos, desarrollados sobre margas y calizas, se clasifican como regosoles lépticos calcáricos y cambisoles endolépticos calcáricos según la taxonomía WRB, con texturas que varían de franco-limosa a franco-arcillosa.

### 2.2. Fraccionamiento físico y análisis químico

Las muestras de suelo fueron separadas en tres fracciones aplicando el método densimétrico adaptado por Cerli *et al.* (2012):

- Fracción libre (FLF): material orgánico joven y de ciclo rápido.
- Fracción ocluida (OLF): material orgánico atrapado en agregados.

- Fracción estable (HF): material asociado a la fracción mineral, resistente y de ciclo lento.

Para el análisis estadístico, las fracciones FLF y OLF se analizaron también agrupadas en una única categoría —fracción lábil—. Para el cálculo del COS, se determinó el carbono total mediante un analizador elemental. En aquellas muestras con presencia de  $\text{CaCO}_3$ , estas se analizaron con el método calcímetro de Bernard.

### 2.3. Adquisición y pretratamiento espectral

Las medidas radiométricas se realizaron sobre las muestras de suelo (seco y tamizado) utilizando un espectrómetro ASD FieldSpec 4 (350-2500 nm). Las lecturas se tomaron en laboratorio bajo condiciones de iluminación controladas, utilizando una sonda de contacto y un panel de referencia Spectralon para la transformación de los valores a reflectividad relativa.

Para optimizar la calidad de la señal, se aplicó como tratamiento básico un suavizado mediante el filtro de Savitzky-Golay (SG) con una ventana de 11 nm. Tomando como referencia estas firmas filtradas, se evaluaron diferentes técnicas de pretratamiento que incluyeron métodos de linealización, corrección de la dispersión y derivadas espectrales (Rinnan *et al.*, 2009) para determinar el tratamiento óptimo en la modelización del COS. Concretamente, se aplicaron las técnicas *Standard Normal Variate* (SNV), *Multiplicative Scatter Correction* (MSC) y primera derivada espectral (FDS), junto con sus combinaciones.

### 2.4. Modelización estadística

La relación entre los espectros del suelo y los contenidos de COS (total y fraccionado) se modeló mediante Regresión por Mínimos Cuadrados Parciales (PLSR, por sus siglas en inglés), técnica de referencia en el análisis quimiométrico de suelos (Gozukara *et al.*, 2025).

El flujo de trabajo incluyó la transformación logarítmica de los datos de COS para corregir su asimetría positiva. Debido al tamaño de la muestra, los modelos se desarrollaron mediante una validación cruzada de 5 particiones, determinando el número óptimo de variables latentes (VL) mediante la aproximación 1-sigma. El rendimiento se evaluó a través del coeficiente de determinación ( $R^2$ ), el error cuadrático medio de predicción (RMSE) y la ratio de desviación del rendimiento (RPD).

## 3. RESULTADOS Y DISCUSIÓN

### 3.1. Caracterización del COS y sus fracciones

El contenido total de COS varió entre 10,3 y 72,9  $\text{mg g}^{-1}$ , con un valor promedio de 34,6  $\text{mg g}^{-1}$  (Tabla 1). Tras el fraccionamiento, la recuperación de masa fue casi total (99,9%), mientras que se recuperó el 93,2% del COS original. El contenido de COS presentó una asimetría positiva y una distribución de frecuencia log-normal, en concordancia con las distribuciones reportadas en estudios previos (Baldock *et al.*, 2013; Henaka Arachchi *et al.*, 2016).

**Tabla 1.** Estadísticas descriptivas del COS (mg g<sup>-1</sup>)

COS	Min	Max	Media	DE	Asimetría
Total	10,26	72,94	34,55	13,71	0,77
FLF	1,20	27,65	6,50	5,55	1,98
OLF	0,29	21,41	6,65	4,65	0,95
HF	5,59	37,68	18,61	7,81	1,43

El análisis de correlación de Pearson (Tabla 2) reveló dinámicas distintas entre las fracciones. El COS total mostró una fuerte correlación con la fracción pesada (HF), lo cual es coherente dado que la HF representa, en promedio, el 96% de la masa de suelo recuperada. Las fracciones ligeras (FLF y OLF) presentaron una correlación muy alta entre sí, pero no mostraron una relación significativa con los niveles de la fracción pesada (HF). Se observaron patrones similares en los estudios de Knox *et al.* (2015) y Vasques *et al.* (2009), aunque con valores de correlación significativamente más altos entre las fracciones lábiles y resistentes.

**Tabla 2.** Matriz de correlación de Pearson

COS	Total	FLF	OLF	HF
Total	1,00			
FLF	0,45**	1,00		
OLF	0,36**	0,79**	1,00	
HF	0,80**	-0,07	-0,09	1,00

\*\* correlaciones significativas al nivel  $p < 0,01$

### 3.2. Evaluación de los pretratamientos

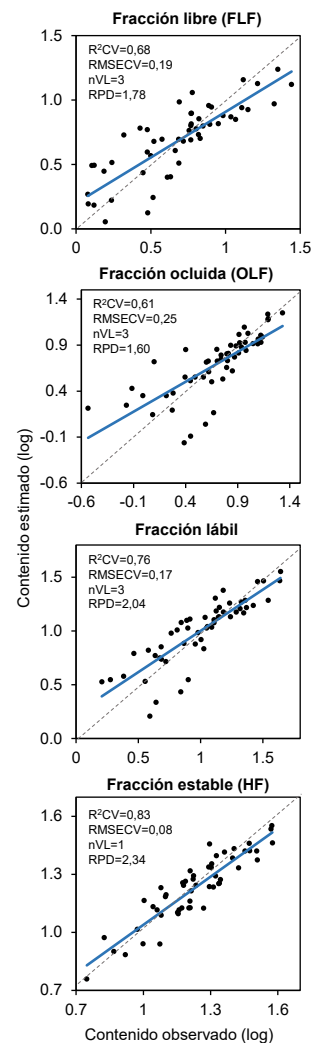
La aplicación de técnicas de pretratamiento mejoró en todos los casos la precisión de las estimaciones de COS total (Tabla 3). Más que los tratamientos individuales, fue el uso combinado de derivadas con técnicas de corrección de dispersión (SNV o MSC) el que generó el mayor impacto. En promedio, estas combinaciones mejoraron el RPD en 1,18 puntos respecto al filtrado básico. Específicamente, la combinación de primera derivada y SNV (FDS-SNV) se identificó como el pretratamiento óptimo, ofreciendo una mejora de 1,37 puntos respecto al modelo base. Estos resultados concuerdan con estudios previos que han demostrado que la aplicación de estas dos técnicas mejora significativamente la capacidad predictiva de los modelos (Dotto *et al.*, 2018; Vasques *et al.*, 2008).

**Tabla 3.** Estadísticos de la validación cruzada (CV) de los modelos PLSR para la estimación del COS total según el pretratamiento espectral aplicado. Se resaltan en negrita los valores del método con mejor capacidad predictiva.

Técnica	nVL	R <sup>2</sup> CV	RMSECV	RPD
SG	3	0,738	0,092	1,945
SNV	2	0,849	0,069	2,593
MSC	2	0,823	0,075	2,385
FDS	2	0,827	0,074	2,418
<b>FDS-SNV</b>	<b>2</b>	<b>0,909</b>	<b>0,054</b>	<b>3,313</b>
FDS-MSC	2	0,889	0,061	2,933

### 3.3. Rendimiento de los modelos de regresión para la estimación de las fracciones de COS

Los modelos PLSR demostraron una capacidad notable para inferir el contenido de las fracciones de COS (Fig. 1). La fracción estable (HF), representativa del carbono mineralizado más resistente, fue estimada con la mayor precisión, explicando el 83% de la varianza con un RPD de 2,34, lo que califica al modelo como excelente. Por su parte, los modelos para las fracciones lábiles (libre —FLF— y ocluida —OLF—) obtuvieron una precisión moderada, explicando entre el 61-68% de la varianza. Al combinar ambas en la categoría de fracción lábil, el rendimiento mejoró significativamente, alcanzando un R<sup>2</sup>CV de 0,76.



**Figura 1.** Modelos de regresión para la estimación de las fracciones de COS.

El rendimiento diferencial entre fracciones guarda una estrecha relación con las dinámicas observadas en el análisis de correlación (Tabla 2). El ajuste superior de la fracción estable es consistente con el hecho de que esta representa el mayor porcentaje de la masa del suelo y presenta la correlación más alta con el COS total. Por el contrario, las fracciones lábiles tienen un peso relativo menor en la muestra, lo que reduce su influencia directa en la señal espectral del suelo integral.

Considerando las diferencias en los métodos de fraccionamiento y el procesamiento de datos, los modelos obtenidos presentan una precisión comparable a la reportada en estudios previos (Jaconi *et al.*, 2019; Knox *et al.*, 2015; Vasques *et al.*, 2009; Viscarra Rossel and Hicks, 2015), utilizando además un número significativamente menor de factores (2 variables latentes), lo que demuestra la ausencia de sobreajuste de los modelos.

#### 4. CONCLUSIONES

El presente estudio demuestra que la espectroscopía vis-SWIR es una herramienta eficaz para monitorizar no solo el carbono total, sino también su distribución en reservorios con distintas estabildades. La optimización mediante el pretratamiento que combina el cálculo de la derivada espectral y la normalización SNV permitió obtener estimaciones de precisión moderada ( $R^2CV = 0,61-0,68$ ) para las fracciones lábiles individuales (FLF y OLF), que mejoraron al agruparse en una categoría combinada ( $R^2CV = 0,76$ ). Los modelos más robustos se alcanzaron en la estimación de la fracción estable (HF) y el COS total ( $R^2CV = 0,83$  y  $0,91$ , respectivamente), las cuales se encuentran altamente correlacionadas. Si bien el tamaño limitado de la muestra sugiere la necesidad de validar estos hallazgos con bases de datos más extensas, los resultados obtenidos en este contexto de montaña indican que la técnica captura eficazmente la información de los distintos reservorios. Esta metodología se postula, por tanto, como una alternativa rápida y no destructiva a los métodos tradicionales, facilitando el seguimiento de la estabilidad del carbono frente a cambios de uso del suelo. Además, los modelos derivados y la información sobre sus bandas espectrales clave sientan las bases para su transferencia a sensores aeroportados y satelitales, facilitando el mapeo de la dinámica del carbono a escala regional o nacional.

#### 5. AGRADECIMIENTOS

Este trabajo ha sido financiado por el proyecto Sens4Soil (PID2024-160889OA-I00, Ministerio de Ciencia, Innovación y Universidades).

#### 6. REFERENCIAS

Baldock, J.A., Hawke, B., Sanderman, J., & Macdonald, L.M. (2013). Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. *Soil Res.* 51, 577. <https://doi.org/10.1071/SR13077>

Cerli, C., Celi, L., Kalbitz, K., Guggenberger, G., & Kaiser, K. (2012). Separation of light and heavy organic matter fractions in soil — Testing for proper density cut-off and dispersion level. *Geoderma* 170, 403–416. <https://doi.org/10.1016/j.geoderma.2011.10.009>

Demattê, J. A. M., Dotto, A. C., Bedin, L. G., Sayão, V. M., & Souza, A. B. (2019). Soil analytical quality control by traditional and spectroscopy techniques: Constructing the future of a hybrid laboratory for low environmental impact. *Geoderma*, 337, 111–121. <https://doi.org/10.1016/j.geoderma.2018.09.010>

Dotto, A.C., Dalmolin, R.S.D., ten Caten, A., & Grunwald, S. (2018). A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. *Geoderma* 314, 262–274. <https://doi.org/10.1016/j.geoderma.2017.11.006>

Gozukara, G., Hartemink, A. E., Huang, J. & Demattê, J. A. M. (2025). Prediction accuracy of pXRF, MIR, and Vis-NIR spectra for soil properties—A review. *Soil Science Society of America Journal*, 89(2), e70028. <https://doi.org/10.1002/saj2.70028>

Henaka Arachchi, M.P.N.K., Field, D.J., & McBratney, A.B. (2016). Quantification of soil carbon from bulk soil samples to predict the aggregate-carbon fractions within using near- and mid-infrared spectroscopic techniques. *Geoderma* 267, 207–214. <https://doi.org/10.1016/j.geoderma.2015.12.030>

Jaconi, A., Poeplau, C., Ramirez-Lopez, L., Van Wesemael, B., & Don, A. (2019). Log-ratio transformation is the key to determining soil organic carbon fractions with near-infrared spectroscopy. *European Journal of Soil Science*, 70(1), 127–139. <https://doi.org/10.1111/ejss.12761>

Keesstra, S. D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., Cerdà, A., ... & Fresco, L. O. (2016). The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. *SOIL*, 2(2), 111–128. <https://doi.org/10.5194/soil-2-111-2016>

Knox, N.M., Grunwald, S., McDowell, M.L., Bruland, G.L., Myers, D.B., & Harris, W.G. (2015). Modelling soil carbon fractions with visible near-infrared and mid-infrared spectroscopy. *Geoderma* 239-240, 229–239. <https://doi.org/10.1016/j.geoderma.2014.10.019>

Lal, R. (2016). Soil health and carbon management. *Food and Energy Security*, 5(4), 212–222. <https://doi.org/10.1002/fes3.96>

O'Rourke, S. M., Angers, D. A., Holden, N. M., & McBratney, A. B. (2015). Soil organic carbon across scales. *Global Change Biology*, 21(10), 3561–3574. <https://doi.org/10.1111/gcb.12959>

Rinnan, Å., Berg, F. van den, & Engelsen, S.B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal. Chem.* 28, 1201–1222. <https://doi.org/10.1016/j.trac.2009.07.007>

Vasques, G.M., Grunwald, S., & Sickman, J.O. (2008). Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. *Geoderma* 146, 14–25. <https://doi.org/10.1016/j.geoderma.2008.04.007>

Vasques, G.M., Grunwald, S., & Sickman, J.O. (2009). Modeling of Soil Organic Carbon Fractions Using VNIR Spectroscopy. *Soil Sci. Soc. Am. J.* 73, 176–184. <https://doi.org/10.2136/sssaj2008.0015>

Viscarra Rossel, R. A., & Hicks, W. S. (2015). Soil organic carbon and its fractions estimated by visible-near infrared transfer functions. *European Journal of Soil Science*, 66(3), 438–450. <https://doi.org/10.1111/ejss.12237>