

Cartografía digital de propiedades edáficas superficiales en las zonas agrícolas de El Órbigo y La Cepeda (León, NO España)

Digital Soil Mapping of surface soil properties in El Órbigo and La Cepeda agricultural areas (León, NW Spain)

Judit Rodríguez-Fernández¹, Montserrat Ferrer-Juliá¹ y Sara Alcalde-Aparicio^{2,*}

¹Área de Geodinámica externa, Facultad de Ciencias Biológicas y Ambientales, Campus de Vegazana, Universidad de León 24071, España

²Área de Edafología y química agrícola, Escuela de Ingeniería Agraria y Forestal (EIAF), Avda. Portugal, 41, Universidad de León, 24071, España

³QGE0 Grupo de Investigación, Universidad de León, 24071, España

(*E-mail: salca@unileon.es)

<https://doi.org/10.19084/rca.28410>

RESUMEN

La escasez de información edafológica y cartografía tradicional ha desembocado en la búsqueda de nuevas técnicas que permitan solucionar esta problemática, de forma que se pueda obtener una cobertura completa y homogénea de la superficie. El objetivo principal de este trabajo es la elaboración de una cartografía de propiedades edáficas de interés como son materia orgánica, arena, limo y arcilla del suelo a través de la aplicación de la metodología *Digital Soil Mapping* en el área de Benavides de Órbigo (León). Para ello, se seleccionaron 75 muestreos de suelo ya existentes y distintas covariables ambientales relacionadas con los factores formadores del suelo. Con estas se llevó a cabo el análisis estadístico mediante la combinación de la regresión lineal múltiple y modelos lineales generalizados para obtener el mejor modelo de predicción para cada variable estudiada, además de los residuos generados por el modelo y la estimación del error. De todas las variables, el mejor valor de ajuste $R^2=0,55$ y el error más bajo 2,145 se obtuvo con la materia orgánica, mientras que arena, limo y arcilla obtuvieron unos ajustes más limitados (0,368, 0,459 y 0,426, respectivamente). Por todo ello, se llegó a la conclusión de que, aunque el método es aplicable y útil, presenta todavía grandes limitaciones, por lo que es importante incidir en la búsqueda de soluciones que permitan obtener mejoras sustanciales en los resultados obtenidos.

Palabras clave: cartografía digital edáfica, covariables ambientales, geoestadística, *scorpan*, modelos de predicción.

ABSTRACT

The lack of soil information and soil maps has led into a search for new techniques to solve this problem, so that a complete and homogeneous coverage of the surface can be obtained. The main objective of this study is the mapping of edaphic properties of interest such as organic matter, sand, silt and clay through the application of the Digital Soil Mapping methodology in the area of Benavides de Órbigo (León). For this purpose, 75 existing soil sampling data and different environmental covariates related to soil-forming factors were selected. The statistical analysis was carried out through the combination of multiple linear regression and generalized linear models to obtain the best prediction model for each variable, in addition to the residuals generated by the model and the error estimation. Among all the variables, the best $R^2=0.55$ fitting value and the lowest error 2.145 was obtained for organic matter; while sand, silt and clay reached more limited fitting values (0.368, 0.459 and 0.426, respectively). For this reason, it was concluded that, although the method is widely applicable and useful, it still has great limitations, so it is quite important to point out the search for solutions that allow to obtain considerable improvements in the results.

Keywords: soil mapping, environmental covariates, geostatistics, *scorpan*, prediction models.

INTRODUCCIÓN

El modelo *scorpan* (McBratney *et al.*, 2003; Lagacherie & McBratney, 2006; Minasny & McBratney, 2016) como base para la Cartografía Digital de Suelos (Digital Soil Mapping o DSM), considera que, a partir de la integración de distintos factores formadores del suelo en una función, se pueden predecir las clases de suelo (S_c) o los atributos (S_a). Estos factores comprenden el propio suelo (s), el clima (c), los organismos (o), el relieve o topografía (r), el material parental (p), la edad (a) y la posición espacial (n). Cada uno de estos factores se define a partir de distintas covariables ambientales. De esta forma, en cada ubicación donde se dispone de datos de perfiles de suelo, por un lado, se obtienen los valores de las propiedades edáficas y, por otro, se estima los valores de cada una de las covariables ambientales seleccionadas para el estudio. Todo ello permite la realización de una regresión lineal múltiple, con el objetivo de modelizar las relaciones entre la variable dependiente y las distintas covariables en los puntos de muestreo, y usar esas relaciones para estimar los valores en lugares sin muestrear (Hengl *et al.*, 2007). En España, la aplicación de este tipo de metodologías que se incluyen en el DSM es todavía muy limitada y se ha aplicado principalmente para estimar la distribución del carbono orgánico del suelo, y menos frecuente, en otras propiedades del suelo como la textura. En el presente estudio, se propone estimar (S_a) los porcentajes de materia orgánica del suelo. En cuanto a los factores formadores del suelo se han seleccionado el clima (c), la vegetación (o), la topografía (r) y el material parental (p).

Área de estudio

El área de estudio se corresponde con una extensión aproximada de 530 km² en el área de Benavides de Órbigo, en el noroeste de la Península Ibérica (Figura 1), caracterizada por un clima continental templado. Se distinguen relieves paleozoicos situados al oeste y norte constituidos por pizarras, areniscas y cuarcitas. Se encuentran principalmente en estas formaciones de la zona de relieves en pendiente asociaciones de Regosoles lépticos/Umbrisoles háplicos y Leptosoles úmbricos con inclusiones de Leptosoles líticos (Alonso Herrero *et al.*, 2010). El resto de la zona está formada por

materiales silíceos asociados a conglomerados, arenas, limos y arcillas del Mioceno y del Cuaternario. Las asociaciones de Umbrisoles húmicos/Regosoles dísticos junto con Cambisoles crómicos se encuentran en las zonas de laderas y vertientes. En la zona centro y noreste, en las terrazas medias y altas, dominan los Acrisoles úmbricos y abrupticos con inclusiones de Acrisoles profúndicos y gléicos. Destacan asociaciones de Cambisoles eútricos y dísticos en zonas de terrazas bajas, junto con Fluvisoles y Regosoles dísticos, cerca de los cursos fluviales (Alonso Herrero *et al.*, 2010).

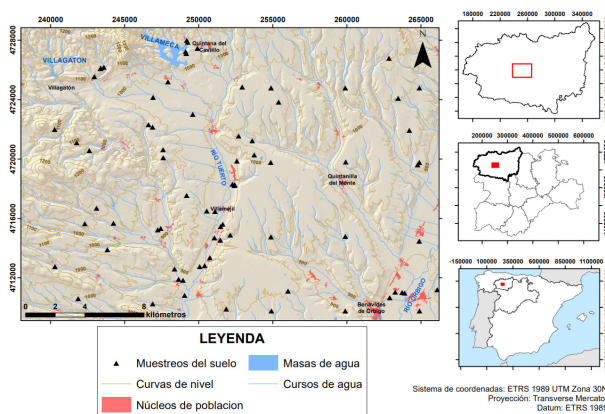


Figura 1 - Localización del área de estudio.

MATERIAL Y MÉTODOS

Las fuentes de datos y los materiales utilizados para estimar las covariables de clima, vegetación, topografía y material parental siguiendo el método propuesto, se explican a continuación.

- 45 muestreos de los primeros 25-30 cm de suelo pertenecientes a la Base de Datos de los Suelos de Castilla y León (ITACyL, 2012) con datos de % materia orgánica, arcilla, limo y arena.

- 30 perfiles realizados para el Plan Piloto del Mapa de suelos de la Provincia de León (Alonso Herrero *et al.*, 2010), de los que se ha seleccionado solo los datos del horizonte A: el valor único para dichos horizontes se estimó como media ponderada en función de la profundidad de cada subhorizonte A.

- Modelo de Elevación Digital (MED) del Centro Nacional de Información Geográfica (CNIG, 2014), con una resolución espacial de 25m x 25m para la estimación de todas las variables topográficas: altitud (m), pendiente (grados), orientación (radianes) e índice topográfico de humedad (TWI).

- Los datos climáticos se estimaron a partir de los mapas ráster (500x500m) derivados del modelo de simulación SIMPA (MITECO, 2019), obteniendo la precipitación media anual (mm), la temperatura media anual (°C) y la evapotranspiración potencial media anual (mm) del período hidrológico 1940/1-2017/8; excepto la radiación solar anual derivada a partir del MED.

- 35 imágenes del satélite Sentinel2 de junio-septiembre 2017 y octubre-mayo de 2018 de la Agencia Espacial Europea (ESA, 2021) con las que se estimó el valor del Índice de Vegetación Normalizado (NDVI) mínimo, máximo y promedio para el período seco (junio, julio, agosto y septiembre) y el período húmedo (resto meses), después de analizar la evolución pluviométrica en el periodo simulado por el modelo SIMPA.

- Cartografía de usos de suelo del Sistema de Información de Ocupación del Suelo de España (SIOSE) del año 2014 (escala 1:25k) que se reclasificó en cuatro clases: suelo desnudo, agrícola-prado-pastizal, matorral y bosque. Al ser esta covariable de tipo categórico, se transformó cada clase a una variable numérica ordinal.

- Cartografía geológica continua del proyecto GEODE escala 1:50k del IGME (2008) para la extracción de la litología, que también se reclasificó para simplificar el número de clases litológicas en función del grado de consolidación, tamaño de grano y naturaleza litológica y se transformó a variable numérica ordinal.

Mediante operaciones de SIG se extrajo el valor de cada una de estas covariables para los 75 puntos con datos de % materia orgánica y se realizó un análisis estadístico para derivar el modelo correspondiente a DSM. Previamente, se dividió de manera aleatoria el número de puntos en dos grupos: 55 se utilizaron para calibrar el modelo y 20 para validarlo. En primer lugar, se realizó un análisis de las correlaciones entre las distintas covariables de

NDVI, para eliminar aquellas que mostrasen una dependencia lineal elevada (coeficiente de Pearson superior a 0,65, $p < 0,01$). Con todas covariables resultantes (tanto de NDVI como del resto de factores) se estimaron diversas regresiones lineales múltiples (MLR) siguiendo el método de “eliminación hacia atrás”. De todos los modelos resultantes, se seleccionaron dos que presentasen un valor de R^2 adecuado teniendo en cuenta su valor de significancia y el número de covariables menor posible. Tras esta selección, se estimaron modelos lineales generalizados (GLM) utilizando las covariables que representaban cada uno de los modelos seleccionados, obteniéndose un valor del Criterio de Información de Akaike (AIC), el cual permite evaluar la idoneidad de un modelo estadístico concreto en comparación con otros modelos. Posteriormente, se seleccionó aquel modelo que menor valor de AIC presentaba o, en caso de que presentasen índices similares, se seleccionó el que menos covariables contenía. Se validaron los resultados comparando en los 20 puntos seleccionados para validación el valor predicho (estimado) con el valor real (medido), mediante la estimación del error cuadrático medio. Finalmente, la interpolación espacial de los residuos se realizó siguiendo el método del inverso de la distancia (IDW) y se aplicó la ecuación del modelo.

RESULTADOS Y DISCUSIÓN

Tras la realización de la MLR, se seleccionaron los modelos *d* y *g*, con valores de R^2 de 0,552 y 0,543, respectivamente y, tras el análisis GLM se seleccionó el modelo *g*, al tener ambos modelos el mismo AIC y ser este el que menos covariables considera (Tabla 1).

Tabla 1 - Resultados de la MLR y del GLM

Modelo	R^2	AIC	Sig.
d	0,552	260,586	0,000
g	0,543	260,586	0,000

A su vez, analizando el RMSE, el valor es similar para los muestreos de calibración como para los de validación, así como para el conjunto total de los datos (Tabla 2).

Tabla 2 - RMSE obtenido para los muestreos de calibración, los de validación y totales

	Error (RMSE)		
	55 muestreos	20 muestreos	75 muestreos
	2,130	2,186	2,145

Estos resultados son mejores que los obtenidos por otros autores, que presentaban un R^2 menor y RMSE mayor. Comparando los valores medidos con los estimados para ambos conjuntos de datos, se observa que no existe una gran diferencia entre ambos (Figura 2).

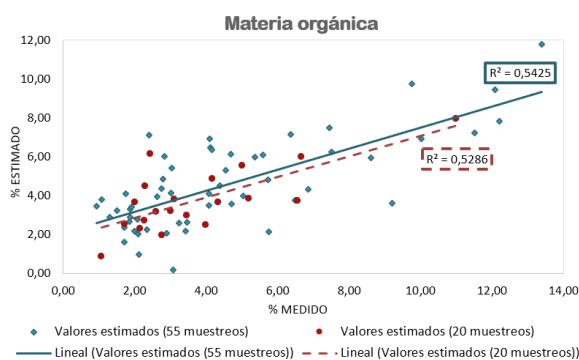


Figura 2 - Valores medios estimados y observados para la variable de materia orgánica.

En cuanto al análisis geoestadístico, los residuos de la materia orgánica muestran una ligera sobreestimación en el centro y noreste del área de estudio, mientras que se subestiman, principalmente, al sur y oeste (Figura 3).

Como resultado final, los mayores porcentajes de materia orgánica se encuentran al norte de la zona siendo más elevados al noroeste (Figura 4).

CONCLUSIONES

1) Los resultados obtenidos son mejorables, pues se han hallado valores de R^2 bastante bajos. Se debe revisar el origen de los muestreos con el fin de ver la afección a los resultados. 2) No hay grandes diferencias entre el grupo de calibración y validación, por lo que se corrobora la utilidad de la segmentación del conjunto de datos. 3) El muestreo de la

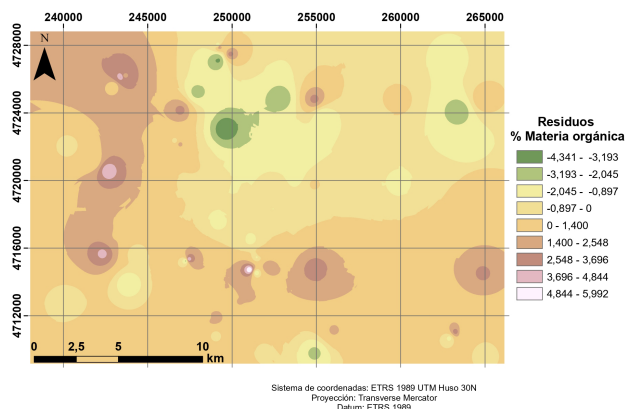


Figura 3 - Interpolación de los residuos de materia orgánica.

cartografía convencional edáfica sigue siendo necesario, pues es un elemento clave para reducir los errores. 4) El método del DSM presenta aún grandes limitaciones a pesar de su gran aplicabilidad.

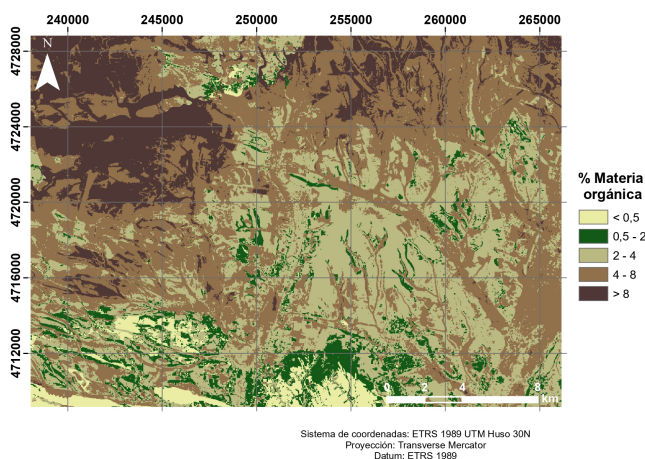


Figura 4 - Distribución espacial de la materia orgánica (%) en la zona con los resultados del modelo.

AGRADECIMIENTOS

Investigación financiada por FEDER/Ministerios de Ciencia e Innovación-Agencia Estatal de Investigación/ Proyecto ESP2017-89045-R.

REFERENCIAS BIBLIOGRÁFICAS

- Alonso Herrero, E.; Vidal-Bardán, M.; Villa-Bermejo, J.E.; Fuertes-Guitérrez, I.; Alcalde-Aparicio, S.; Robles-Llamazares, J.A.; Rodríguez-Blanco, J. & Juan-Gonzalez, L.A. (2010) - *Plan Piloto del Mapa de Suelos de la Provincia de León*. E1:50k. Memoria Hoja160 (Benavides). Universidad de León.
- CNIG (2014) - *Centro de Descargas*. Instituto Geográfico Nacional [cit. 2021.01.12] <<http://centrodedescargas.cnig.es/CentroDescargas/>>
- ESA (2021) - *Sentinel-2 MSI Introduction*. European Space Agency. [cit. 2021.01.12] <<https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi>>
- Hengl, T.; Heuvelink, G.B. & Rossiter, D.G. (2007) - About regression-kriging: From equations to case studies. *Computers & Geosciences*, vol. 33, n. 10, p. 1301-15. <https://doi.org/10.1016/j.cageo.2007.05.001>
- IGME (2008) - *Mapa Geológico Digital continuo E. 1:50.000, Zona Asturoccidental-Leonesa (Zona 1100)*. Instituto Geológico y Minero de España <<http://info.igme.es/cartografiadigital/geologica/geodezona.as>>
- ITACyL (2012) - *Base de datos*. Instituto Tecnológico Agrario de Castilla y León. [cit. 2021.01.12] <http://suelos.itacyl.es/base_datos>
- Lagacherie, P. & McBratney, A.B. (2006) - Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. *Developments in Soil Science*, vol. 31, p. 3-22. [https://doi.org/10.1016/S0166-2481\(06\)31001-X](https://doi.org/10.1016/S0166-2481(06)31001-X)
- McBratney, A.B.; Santos, M.M. & Minasny, B. (2003) - On digital soil mapping. *Geoderma*, vol. 117, n. 1-2, p. 3-52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Minasny, B. & McBratney, A.B. (2016) - Digital soil mapping: A brief history and some lessons. *Geoderma*, vol. 264, part B, p. 301-311. <https://doi.org/10.1016/j.geoderma.2015.07.017>
- MITECO (2019) - *Modelo SIMPA 2019*. Ministerio para la Transición Ecológica y el Reto Demográfico. [cit. 2021.01.08] <<https://www.miteco.gob.es/es/agua/temas/evaluacion-de-los-recursos-hidricos/evaluacion-recursos-hidricos-regimen-natural/>>