

Capítulo 5

Mapeo digital de los suelos de la región de la Hoya del Río Suárez

Viviana Marcela Varón-Ramírez, Ruy Edeymar Vargas Díaz
y Gilberto Alonso Murcia Contreras

Resumen

Conocer la oferta edáfica y su comportamiento espacial es la base del establecimiento de sistemas de producción eficientes y sostenibles. El objetivo de este estudio fue identificar la distribución espacial de nueve propiedades de suelo en zonas paneleras de la Hoya del Río Suárez (Boyacá y Santander) a partir de la exploración de técnicas geoestadísticas (regresión *kriging*) y del aprendizaje automático asistido por computadora (bosques aleatorios, máquinas de soporte vectorial y modelos ensamblados). El estudio contempló 47.204 ha en los municipios de Chitaraque, Moniquirá, Togüí, San José de Pare y Santana, en Boyacá, y Barbosa, Chipatá, Güepa, San Benito, Suaita y Vélez, en Santander, y contó con una base de datos de 933 puntos de muestreo y un conjunto de 41 variables predictoras. Como resultados relevantes, se obtuvieron modelos de predicción con una cantidad de varianza explicada entre 0,255 y 0,519 para contenido de arcilla, densidad aparente, porosidad total, pH y capacidad de intercambio catiónico. Este estudio representa una contribución significativa para el conocimiento de la oferta edáfica y es de potencial uso para ejercicios a escala regional.

Palabras clave: geoestadística, aprendizaje de máquina, zonas paneleras

Introducción

El suelo juega un papel importante en la implementación de los planes de seguridad alimentaria, el sostenimiento de la biodiversidad, la mitigación del cambio climático y la generación de servicios ecosistémicos (Yang et al., 2020). El suelo es un componente de los sistemas de producción agropecuaria, y, por tanto, conocer la oferta edáfica y su comportamiento espacial es la base del establecimiento de sistemas de producción más eficientes y sostenibles (Lal et al., 2021). En Colombia, el mayor núcleo productivo de caña de azúcar para la producción de panela se encuentra en la región de la Hoya del Río Suárez (HRS) (departamentos de Boyacá y Santander), lugar con los mayores rendimientos de producción del país y con buen desarrollo tecnológico en este tipo de cadena productiva, aunque no se cuenta con estudios detallados de suelos (mapas) como herramienta para los tomadores de decisiones y la academia en la región.

Los mapas de suelo en Colombia se han realizado desde la perspectiva del mapeo convencional de suelos realizado por el Instituto Geográfico Agustín Codazzi (IGAC); estos estudios, hasta el momento, han suplido la demanda de información edáfica en la región de la HRS. Sin embargo, esta información tiene diferentes limitantes de uso, pues es necesario tener un conocimiento pedológico para poder interpretar los perfiles representativos de cada unidad cartográfica de suelos y construir mapas de características de estos. Además, la escala de los estudios (1:100.000) evidentemente no representa las condiciones del suelo en áreas más pequeñas.

Atendiendo estas necesidades, el mapeo digital de suelos (MDS) ha surgido como un campo de la ciencia que se basa en la predicción de suelos a través de la construcción de modelos matemáticos y estadísticos que son capaces de explicar parcialmente los patrones de distribución de una propiedad o función de suelo en el espacio geográfico (McBratney et al., 2003). El MDS se ha realizado a través de técnicas geoestadísticas, en busca de

construir un modelo de covarianza en función de la distancia entre puntos identificados, y también se han implementado técnicas de aprendizaje automático asistido por computadora (ML, por sus siglas en inglés: *machine learning*), en las que el principal objetivo es predecir (por regresión o clasificación) el comportamiento de un suelo en un lugar dado a partir de un conjunto de covariables (Minasny & McBratney, 2016).

Diferentes trabajos han mapeado propiedades físicas y químicas de suelos a nivel global (Hengl et al., 2017), continental (Guevara et al., 2018; Zomer et al., 2017) y nacional (Malone et al., 2016; Poggio & Gimona, 2017). En Colombia se han hecho diferentes esfuerzos nacionales y regionales por construir mapas de propiedades de suelos implementando MDS. Desde el enfoque del análisis geoestadístico, Vargas Díaz et al. (2023) exploraron diferentes técnicas determinísticas y probabilísticas para mapear cinco propiedades de suelo en la HRS. A escala nacional, se han implementado modelos ensamblados de ML para predecir la textura del suelo en cinco profundidades estándar en todo el territorio continental (Varón-Ramírez, Araujo-Carrillo, & Guevara Santamaría, 2022); también se han utilizado algoritmos de ML para estimar el contenido de carbono orgánico de suelos en el sistema de páramos a nivel nacional (Gutiérrez Díaz et al., 2020) y, a escala regional, para predecir la textura y la densidad aparente en la región del altiplano cundiboyacense (Araujo-Carrillo et al., 2021). Por otra parte, se han empleado técnicas geoestadísticas para estimar el pH y bases de cambio del suelo en áreas afectadas por inundaciones (Jaramillo-Barrios et al., 2019) o en algunas áreas pequeñas de la sabana de Bogotá (Cortes et al., 2018). Sin embargo, en estos estudios no se identifica un único modelo de MDS que sea aplicable a todas las regiones, por lo que es necesario, para cada caso de estudio, explorar diferentes estrategias de construcción de modelos según el conjunto de datos disponible (Guevara et al., 2018).

En todos estos ejercicios se ha identificado el potencial de uso de las herramientas del MDS que puedan generar productos más detallados de suelos y que además sean de fácil acceso a los usuarios.

Por lo anterior, y dada la necesidad de información detallada de suelos en la HRS y la importancia de identificar el potencial del MDS, el objetivo del presente estudio fue establecer la distribución espacial de algunas propiedades físicas y químicas de los suelos de la HRS, mediante el uso de cuatro modelos (uno geoestadístico y tres de ML). Este esfuerzo por representar a través de un mapa el comportamiento de las propiedades del suelo representa un avance significativo en el conocimiento de la región. Por tanto, se recomienda el empleo de este tipo de productos en ejercicios de zonificación y evaluación de tierras, planificación y ordenamiento productivo, estimación del potencial de los suelos para la prestación de servicios ecosistémicos, entre otros.

Metodología

Muestreo

La zona de estudio comprendía 47.204 ha pertenecientes a la región de la Hoya del Río Suárez (HRS), en los municipios de Chipatá, San Benito, Suaita, Güepso, Vélez y Barbosa, en Santander, y Santana, Chitaraque, San José de Pare, Togüí y Moniquirá, en Boyacá. En la zona de estudio se diseñó un sistema de muestreo en red rígida de 700 m, lo que implicó un muestreo cada 49 ha, para un total de 933 puntos (figura 5.1). El muestreo fue realizado en el periodo 2015-2017.

Para cada punto se tomaron muestras de suelos de entre 0 y 15 cm de profundidad, y se realizaron análisis físicos y químicos en laboratorio. Para las propiedades físicas se usaron los siguientes métodos: textura (fracción de arcilla [Ar]; arena [A]; limo [L], en %), método de Bouyoucos; densidad aparente (ρ_a , en g/cm^3), método de anillo con volumen conocido; porosidad total (porcentaje de $P_T = 1 - [\rho_a/\text{densidad relativa}]$); diámetro medio ponderado (D_{MP} , en mm), método de Yoder; y agua aprovechable (porcentaje de $AA = \text{capacidad de campo} - \text{punto de marchitez permanente}$), a partir de las curvas de retención de agua, con el método de las ollas de presión.

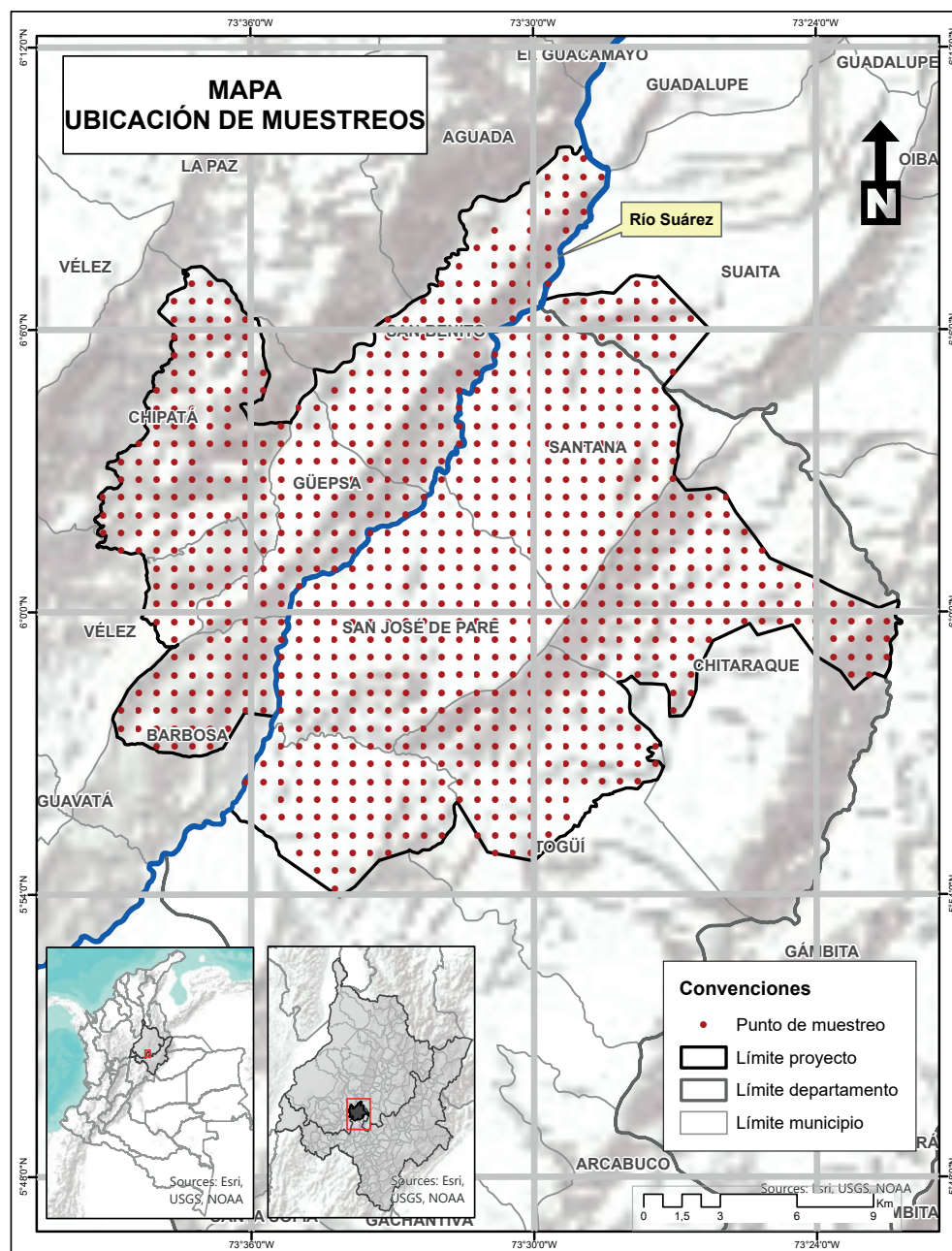


Figura 5.1. Zona de estudio y distribución de los puntos de muestreo.

Fuente: Elaboración propia

Para las variables químicas se usaron los siguientes métodos: pH, método del potenciómetro; capacidad de intercambio catiónico efectivo (CICE, en cmol(+)/kg), método de suma de bases intercambiables; fósforo disponible (P, en mg/kg), método Bray II; y contenido de materia orgánica (en % de MO), método Walkley-Black. En los análisis se utilizó la guía para métodos analíticos del laboratorio de suelos del IGAC (2006).

Análisis de datos

Se hizo un análisis exploratorio de datos de las propiedades físicas y químicas de los suelos para detectar la presencia de datos anómalos y atípicos; inicialmente, se descartaron los datos anómalos, y luego se analizó la relación con las otras variables. Posteriormente, se estimaron los estadísticos descriptivos básicos, como media, máximo, mínimo, asimetría, curtosis y coeficiente de variación (Warrick & Nielsen, 1980). Todos los algoritmos implementados para el análisis de datos, la construcción de modelos y su validación fueron realizados en el *software* libre R, y para cada uno de los procedimientos realizados se indica la librería usada en cada caso.

Mapeo digital de suelos (MDS)

Construcción y selección de covariables

El MDS fue basado en el modelo SCORPAN (McBratney et al., 2003), en el cual se propone predecir el comportamiento de una propiedad de suelo (s) como un resultado de la interacción entre los factores de formación propuestos por el clima (c), los organismos (o), el relieve (R), el material parental (P), la edad (A) y la relación con sus vecinos (N). Por lo tanto, fueron construidas 41 covariables para la zona de estudio; en la tabla 5.1 se muestran los grupos de covariables (binarias o continuas) que explican cada factor de formación de suelo. Las covariables fueron armonizadas a una resolución (tamaño de pixel) de 60 m; esta resolución fue establecida de acuerdo con la relación entre la cantidad de puntos de muestreo y el área de estudio (Hengl, 2006).

Tabla 5.1. Covariables predictoras según factor de formación

Factor de formación	Covariables	Fuentes
Suelo	Condición de aluminio y fertilidad*	IGAC (2003, 2005)
Clima	Precipitación media multianual (PPT, en mm)** Temperatura media anual (T_MED, en °C)** Brillo solar medio anual (BS, en h/día)** Humedad relativa media anual (HR, en %)**	Corporación Colombiana de Investigación Agropecuaria (Corpoica, 2017)
Organismos	Índice de vegetación de diferencia normalizada (NDVI)**	Google Earth Engine
Relieve	Tipos de relieve* Modelo de elevación digital (DEM)** Derivados del DEM**	IGAC (2003, 2005) Farr et al. (2007)
Material parental	Material parental*	IGAC (2003, 2005)
Edad	Orden de suelo*	IGAC (2003, 2005)

* Covariable binaria; ** covariable continua

Fuente: Elaboración propia

Para seleccionar las mejores covariables con capacidad de predecir cada una de las propiedades de suelo, se implementó una selección recursiva de variables (RFE) de la librería *caret* (Kuhn, 2020). Este algoritmo construye diferentes modelos con subconjuntos de covariables en los cuales se identifica el error medio cuadrático (RMSE). Luego, cada covariable se clasifica según su importancia para el modelo. En cada iteración de selección de características, los predictores más importantes son retenidos y el modelo es reajustado. Finalmente, se selecciona el mejor rendimiento de cada modelo, con un número específico de covariables (Kuhn, 2020).

Modelos

El conjunto de puntos muestreado para cada una de las propiedades de estudio fue dividido en dos subconjuntos: un 75 % de los datos para entrenar los modelos, y un 25 % para hacer una validación externa. Esta selección de los conjuntos se realizó de forma aleatoria, para garantizar que la distribución de los subconjuntos tuviera el mismo comportamiento (Malone et al., 2017).

Para construir los modelos de predicción, se emplearon cuatro algoritmos: un algoritmo basado en técnicas geoestadísticas (regresión *kriging* [RK]) (Odeh et al., 1995); dos algoritmos de ML (bosques aleatorios ensamblados [RF]) y máquinas de soporte vectorial (SVM) (Wadoux et al., 2020), y un modelo ensamblado (*ensemble*), es decir, un modelo lineal entre RF y SVM, mediante el uso de la librería *caret* (Kuhn, 2020).

Evaluación de la calidad de la predicción

Con el conjunto de datos de validación externa (el 25 % del total de los datos), se determinaron los siguientes parámetros de bondad de ajuste: el error medio (ME); la raíz del error medio cuadrático (RMSE); la cantidad de la varianza explicada (AVE) (Kempen et al., 2010); el coeficiente de correlación de concordancia (CCC) propuesto por Lin (1989), que puede tomar valores entre 0 y 1, y el coeficiente de concordancia (IOA) basado en Willmott et al. (2012), que puede tomar valores de -1 a 1, siendo 1 el modelo de mejor desempeño. Por último, se identificó la distribución espacial del error de predicción haciendo uso de técnicas geoestadísticas, específicamente *kriging* ordinario (Giraldo Henao, 2006).

Resultados y discusión

A continuación, se presentan los principales resultados del análisis exploratorio, con la descripción del desempeño de los modelos implementados para el mapeo de las propiedades de suelo y el

comportamiento espacial de aquellas propiedades que tienen un mejor desempeño.

Descripción de los suelos de la HRS

Los estadísticos descriptivos de las propiedades de suelo de la HRS se resumen en la tabla 5.2. Las muestras de suelo presentaron un predominio de las familias texturales francafina y fina (figura 5.2), según la clasificación de la USDA (Soil Survey Staff, 2014); valores de DMP de moderadamente estables a estables (mayores de 1,5 mm); valores de AA de medios a bajos (menores del 25 %), y valores de PT óptimos (entre el 50 y el 55 %). Las calificaciones dadas corresponden a las clasificaciones propuestas por el IGAC (2015).

En cuanto a las variables químicas, se encontraron suelos con MO alta (mayor del 5 %), teniendo en cuenta que la región se encuentra en altitudes entre 1.000 y 2.000 m s. n. m.; además, se observaron pH de extremadamente ácidos a medianamente alcalinos, valores de CICE predominantemente medios-óptimos y, finalmente, valores de P en mayor medida de medios a ideales. Las calificaciones dadas corresponden a las clasificaciones del IGAC (2015).

Tabla 5.2. Estadísticos descriptivos de las propiedades de los suelos (0-15 cm) de la HRS

Propiedad	Muestra	Media	Mediana	Mín.	Máx.	DE	CV	Asim.	Curt.
Ar (%)	933	31,41	31,60	2,00	70,00	11,88	0,38	-0,15	-0,28
DA (g/cm ³)	932	1,09	1,09	0,54	1,74	0,20	0,18	-0,11	-0,25
PT (%)	931	55,70	55,60	31,92	75,36	7,17	0,13	0,03	-0,24
AA (% vol)	932	4,65	4,07	0,13	24,38	2,74	0,59	2,35	9,06
DMP (mm)	933	2,62	2,56	0,56	5,21	0,81	0,31	0,29	-0,39
MO (%)	931	6,42	5,85	0,41	18,10	2,78	0,43	1,12	1,58
pH	933	5,59	5,20	4,02	8,41	1,09	0,19	0,98	-0,19
CICE (cmol(+)/ kg)	924	13,92	10,89	0,91	53,01	9,71	0,70	1,80	3,34
P (mg/km)	898	12,01	8,74	0,13	49,72	9,79	0,81	1,62	2,40

Nota: mínimo (Mín.), máximo (Máx.), desviación estándar (DE), coeficiente de variación (CV), asimetría (Asim.), curtosis (Curt.), Ar (arcilla), DA (densidad aparente), PT (porosidad total), AA (agua aprovechable), DMP (diámetro medio ponderado), MO (materia orgánica), CICE (capacidad de intercambio catiónico efectivo), P (fósforo disponible)

Fuente: Elaboración propia

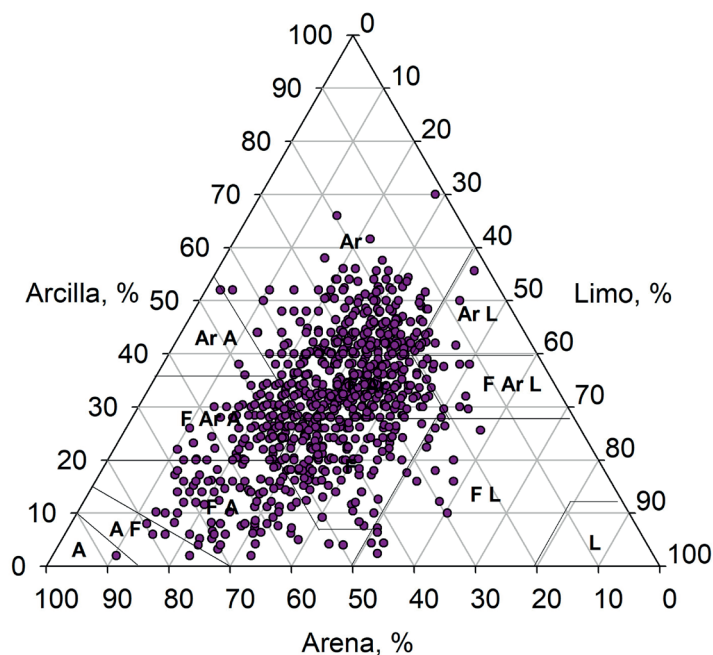


Figura 5.2. Triángulo textural de las fracciones de arcilla (Ar), arena (A) y limo (L) de los suelos de la HRS, según la clasificación de la USDA.

Fuente: Elaboración propia

Selección de covariables

En cuanto a la importancia de las variables predictoras, en la tabla 5.3 se puede observar que covariables climáticas como brillo solar, humedad relativa y precipitación fueron seleccionadas frecuentemente; además, covariables de relieve como el modelo de elevación digital (DEM) y sus derivados, que indican las distancias horizontal y vertical a la red de drenaje del área de estudio, fueron las representativas de dicho grupo de covariables. Es interesante resaltar que covariables binarias, como la presencia o ausencia de aluminio o el nivel de fertilidad, se seleccionaron con frecuencia para predecir los valores de propiedades químicas del suelo, como el pH.

Tabla 5.3. Importancia de las covariables (top 5) para las propiedades mapeadas

Propiedad	Covs.	RMSE	Top 5
Ar (%)	23	10,30	BS, DEM, HR, PPT, T_MED
DA (g/cm ³)	18	0,15	BS, PPT, NivelBaseRedCanales, DEM, HR
PT (%)	40	5,62	BS, PPT, NivelBaseRedCanales, ProfundidadValles, DEM
AA (% vol)	40	2,65	BS, DistanciaVerticalRedCanales, PPT, HR, ProfundidadValles
DMP (mm)	10	0,79	T_MAX, DEM, BS, PPT, T_MED
MO (%)	33	2,62	BS, PPT, DEM, HR, NivelBaseRedCanales
pH	39	0,74	BS, PPT, HR, Fertilidad_Baja, NivelBaseRedCanales
CICE (cmol(+)/kg)	30	7,61	BS, Aluminio_Alto, PPT, Fertilidad_Baja, HR
P (ppm)	39	9,83	ProfundidadValles, DEM, T_MED, NivelBaseRedCanales, PosiciónRelativaPendiente

Nota: número de covariables seleccionadas para la construcción del modelo de predicción (Covs.), error medio cuadrático final de las predicciones realizadas con las covariables seleccionadas (RMSE), arcilla (Ar), densidad aparente (DA), porosidad total (PT), agua aprovechable (AA), diámetro medio ponderado (DMP), materia orgánica (MO), capacidad de intercambio catiónico efectivo (CICE), fósforo disponible (P). Las covariables NivelBaseRedCanales, ProfundidadValles, DistanciaVerticalRedCanales y PosiciónRelativaPendiente corresponden a derivados del DEM

Fuente: Elaboración propia

Calidad de las predicciones

En la tabla 5.4 se pueden observar los parámetros de bondad de ajuste obtenidos mediante una validación externa (el 25 % del conjunto de datos) para cada modelo entrenado. Con color naranja se indican los modelos seleccionados para representar la distribución espacial de cada propiedad de suelo y los modelos con mejor desempeño para cada parámetro.

Tabla 5.4. Parámetros de bondad de ajuste para cada modelo entrenado en cada propiedad de estudio

Propiedad	Modelo	ME	RMSE	AVE	CCC	IOA
Ar (%)	RK	0,057	10,494	0,226	0,408	0,561
	RF	-0,512	10,300	0,255	0,407	0,567
	SVM	-0,354	10,669	0,200	0,379	0,561
	<i>Ensemble</i>	-0,512	10,300	0,255	0,394	0,568
DA (g/cm ³)	RK	0,015	0,158	0,267	0,524	0,579
	RF	0,006	0,152	0,326	0,551	0,598
	SVM	0,011	0,161	0,241	0,512	0,577
	<i>Ensemble</i>	0,014	0,157	0,281	0,519	0,584
PT (%)	RK	-0,429	5,620	0,348	0,542	0,603
	RF	-0,485	5,408	0,397	0,577	0,614
	SVM	-0,234	5,546	0,365	0,564	0,605
	<i>Ensemble</i>	-0,542	5,584	0,359	0,532	0,605
AA (% vol)	RK	-0,533	2,581	0,022	0,223	0,511
	RF	-0,042	2,624	-0,010	0,150	0,501
	SVM	-0,700	2,730	-0,094	0,098	0,499
	<i>Ensemble</i>	-0,233	2,587	0,018	0,112	0,508
DMP (mm)	RK	-0,104	0,762	0,021	0,142	0,509
	RF	-0,052	0,745	0,063	0,175	0,519
	SVM	-0,085	0,755	0,039	0,154	0,518
	<i>Ensemble</i>	-0,066	0,747	0,058	0,116	0,519
MO (%)	RK	-0,292	2,512	0,083	0,236	0,532
	RF	0,240	2,424	0,146	0,257	0,525
	SVM	-0,247	2,462	0,119	0,257	0,532
	<i>Ensemble</i>	0,098	2,433	0,145	0,224	0,526

Continúa

Continuación tabla 5.4

Propiedad	Modelo	ME	RMSE	AVE	CCC	IOA
pH	RK	-0,110	0,750	0,468	0,655	0,663
	RF	0,023	0,713	0,519	0,691	0,669
	SVM	-0,053	0,761	0,453	0,678	0,651
	Ensemble	0,005	0,717	0,510	0,693	0,672
CICE (cmol(+)/kg)	RK	-1,006	6,416	0,407	0,583	0,639
	RF	0,845	6,607	0,371	0,605	0,618
	SVM	-0,621	6,507	0,390	0,582	0,634
	Ensemble	0,695	6,541	0,381	0,596	0,613
P (ppm)	RK	-2,523	9,318	-0,019	0,110	0,541
	RF	0,537	9,439	-0,045	0,038	0,483
	SVM	-2,722	9,440	-0,046	0,099	0,535
	Ensemble	0,043	9,136	0,011	0,039	0,503

Nota: arcilla (Ar), densidad aparente (DA), porosidad total (PT), agua aprovechable (AA), diámetro medio ponderado (DMP), materia orgánica (MO), capacidad de intercambio catiónico efectivo (CICE), fósforo disponible (P), regresión *kriging* (RK), bosques aleatorios ensamblados (RF), máquinas de soporte vectorial (SVM), modelo ensamblado (*ensemble*), error medio (ME), raíz del error medio cuadrático (RMSE), varianzas explicadas (AVE), coeficientes de correlación de concordancia (CCC), coeficientes de concordancia (IOA)

Fuente: Elaboración propia

El desempeño de cada uno de los modelos evaluados es similar para cada propiedad, pues se observa que presentan RMSE y ME parecidos, condición típica en la evaluación de modelos predictivos (Organización de las Naciones Unidas para la Alimentación y la Agricultura [FAO], 2018). En general, las mejores predicciones se obtuvieron con los modelos de RK y RF, y el desempeño bajo lo presentó el modelo de SVM, situación similar a la encontrada en ejercicios de mapeo en el altiplano cundiboyacense (Araujo-Carrillo et al., 2021). Para propiedades como Ar, DA, PT, pH y CICE, se obtuvieron las mejores predicciones, con AVE entre 0,255 y 0,519, CCC entre 0,408 y 0,691 e IOA entre 0,568 y 0,672. Los modelos para la predicción de propiedades como AA, DMP, MO y P presentaron un bajo desempeño, por lo que se utilizó el método de interpolación

de la distancia inversa ponderada (IDW) (Bivand et al., 2013) para representar la distribución espacial de las variables mencionadas.

Interpretación de la distribución espacial de las propiedades de suelo

En la figura 5.3 se muestra la distribución espacial de las propiedades de Ar, DA, PT, pH y CICE, empleando los modelos descritos anteriormente. Estos mapas se han representado en una paleta de color en la que los valores bajos se muestran en color morado, los valores medios en color *beige* y los valores altos en color verde. También se presenta la distribución del error de predicción, errores que se encontraron con valores positivos (en rojo) y negativos (en verde), siendo aquellas zonas rojas los sitios donde la estimación del modelo fue mayor (sobreestimación), y aquellas con color verde, los sitios donde la estimación del modelo fue menor (subestimación); las zonas amarillas se presentan como aquellos sitios donde los modelos obtuvieron los menores errores de predicción. Por otra parte, la figura 5.4 muestra la representación espacial de las demás propiedades de suelo (AA, DMP, MO y P) que no se ajustaron a un modelo predictivo.

En general, se puede observar que todas las propiedades del suelo, con excepción del P, el DMP y la MO, presentaron un comportamiento homogéneo en función de la vertiente del río Suárez. Por ejemplo, se encuentran suelos predominantemente de francofinos a finos, con mayor contenido de arcilla (Ar mayor al 35 %), en la vertiente occidental del río Suárez, y suelos de francos a francogruesos (Ar menor al 35 %) en la margen oriental del río. Debido a que, precisamente, el río Suárez corresponde al límite departamental en el área de estudio, de ahora en adelante este documento se referirá a los suelos por departamento, para facilitar la descripción.

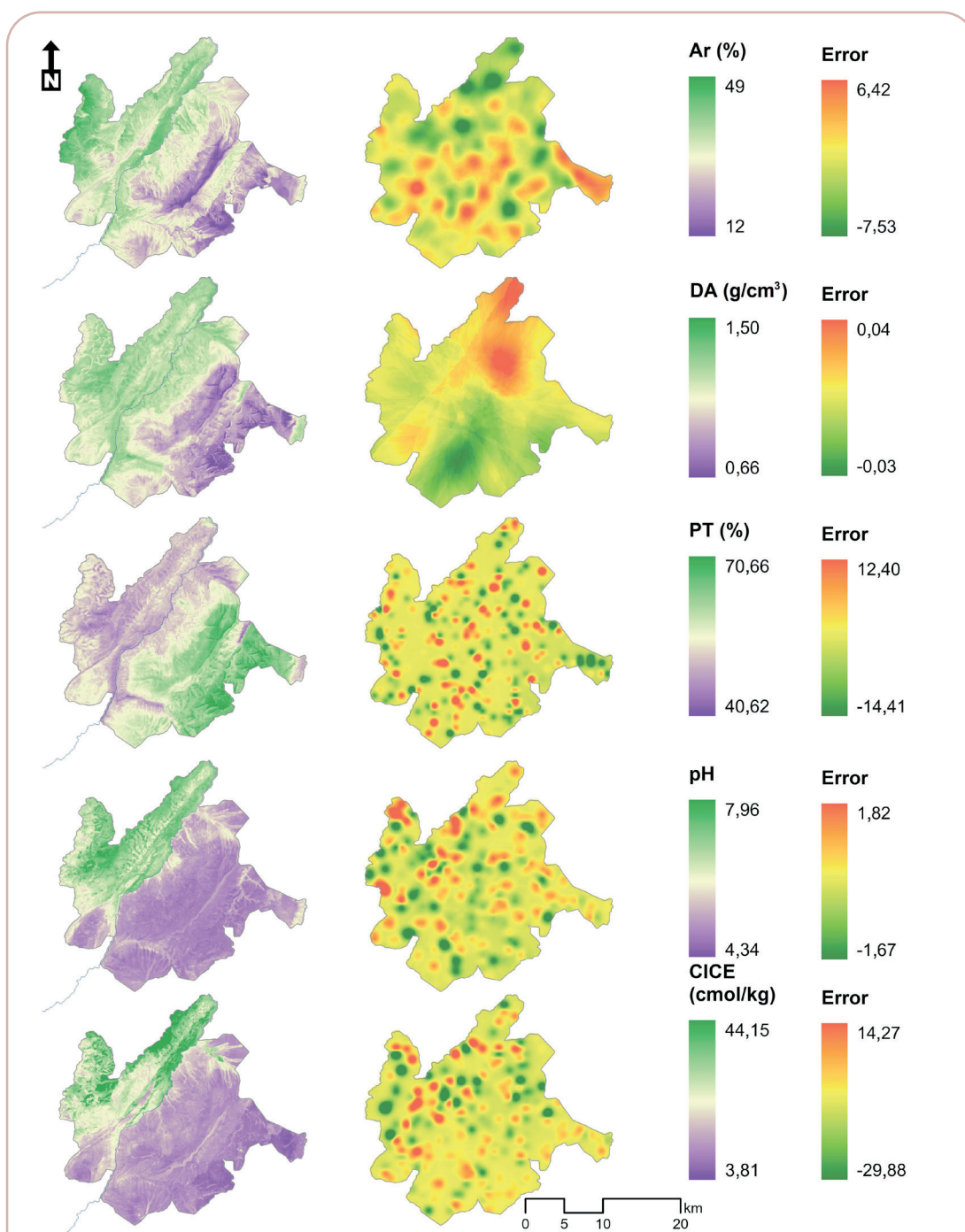
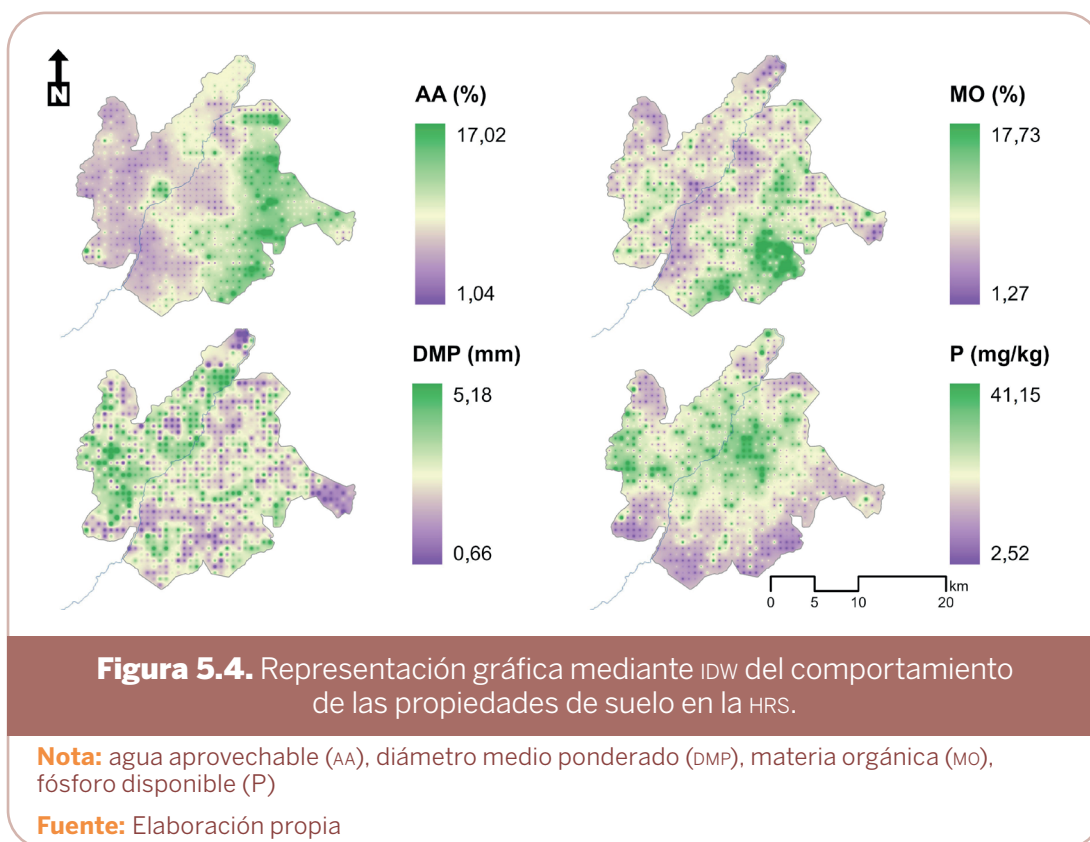


Figura 5.3. Distribución espacial y distribución del error de predicción de las propiedades de suelo en la HRS.

Nota: arcilla (Ar), densidad aparente (DA), porosidad total (PT), capacidad de intercambio catiónico efectivo (CICE)

Fuente: Elaboración propia



Las condiciones ideales de ρ_a para el desarrollo de la caña de azúcar se reportan entre 1,1 y 1,2 g/cm³ (Anschau et al., 2009); en los municipios del departamento de Santander (Barbosa, Chipatá, Güepesa, San Benito, Suaita y Vélez) se obtuvieron valores de ρ_a entre 1,0 y 1,5 g/cm³, que están entre los rangos ideales, pero es importante tener en cuenta que aquellos sitios con ρ_a mayores de 1,4 g/cm³ y texturas finas pueden presentar problemas de adensamiento (Medina et al., 2011) y, por tanto, una restricción para el desarrollo radicular de las plantas de caña de azúcar (Baquero et al., 2012). Asimismo, al observar la ρ_t con valores menores al 55 %, se confirma un proceso de disminución del espacio poroso; sin embargo, los valores de ρ_t no alcanzan un valor crítico para el desarrollo radicular de las plantas de caña de azúcar (menores al 40 %) (Montenegro et al., 2010). En cuanto al DMP y el AA, se observa que estas propiedades tienen más zonas con valores

bajos en comparación con el departamento de Boyacá, lo que indica que estas zonas son de menor estabilidad y tienen una baja capacidad de almacenamiento de agua, por lo que son más susceptibles a procesos erosivos (Montenegro & Malagón, 1990).

En cuanto a las propiedades químicas en los municipios del departamento de Santander, el pH presentó valores entre 6 y 8, lo que indica condiciones de medianamente ácidas a moderadamente alcalinas; los valores altos de pH son contradictorios en esta zona, pues los tipos de roca reportados en estos lugares no contienen altos contenidos de calcio (rocas carbonatadas) (IGAC, 2003); sin embargo, la mayor parte de los suelos en este departamento presentaron condiciones óptimas de pH para el desarrollo de las plantas de caña (Tarazona Parra, 2011), aunque pH mayores de 7,2 pueden afectar el crecimiento y la concentración de sacarosa en el cultivo (García et al., 2007). Por otra parte, los valores de CICE para estos municipios son predominantemente mayores de 25 cmol/kg, lo que indica valores altos y adecuados (IGAC, 2015). En cuanto a MO, se encontraron valores de bajos (< 3 %) a muy altos (> 5 %), coincidiendo con los valores reportados por Manrique et al. (2000). El P no mostró un patrón acorde a la topografía, debido a que este elemento en el suelo es influenciado posiblemente por las enmiendas realizadas al suelo cuando se establecen los cultivos de caña de azúcar para la producción de panela (García et al., 2007).

Para los municipios del departamento de Boyacá (Chitaraque, Moniquirá, Togüí, San José de Pare y Santana), se observan valores de D_A menores de 1 g/cm³, e incluso, en algunos sitios de Chitaraque y Togüí, se registraron los menores valores de D_A , lo cual está muy relacionado con la presencia de ceniza volcánica en algunas zonas (IGAC, 2005). Con relación a lo anterior, en estas zonas se encontró una P_T alta, con valores superiores al 55 %, pues la ceniza volcánica facilita la estructuración y organización de las partículas del suelo (IGAC, 2015); sin embargo, esta condición de alta porosidad no alcanza los valores críticos (mayores al 75 %) para el adecuado desarrollo radicular de las plantas de caña

de azúcar (Montenegro et al., 2010). También para estas zonas se encontraron suelos con una mayor estabilidad de agregados (DMP), lo que indica una condición óptima de la estructura del suelo para el crecimiento de la caña de azúcar para la producción de panela, pues dicha estructura favorece el almacenamiento de agua y aire en el suelo (Osorio, 2007). Sin embargo, aun cuando en esta zona se presentaron las mejores condiciones físicas de suelo, el AA se clasifica dentro de un rango medio, lo que evidencia la necesidad de incorporar MO para mejorar la capacidad de retención de agua en el suelo y favorecer, así, el desarrollo vegetativo de las plantas.

En cuanto a las propiedades químicas para los municipios del departamento de Boyacá, se encontraron suelos de moderadamente ácidos a fuertemente ácidos ($\text{pH} < 6$), siendo estos últimos restrictivos para el adecuado desarrollo radicular del cultivo de caña de azúcar para la producción de panela (Tarazona Parra, 2011); la CICE presentó valores altos-muy altos (mayores de 24 cmol/kg); se encontraron valores de MO de bajos a muy altos, y, finalmente, contenidos de P de bajos a medios para los municipios de Moniquirá, Chitaraque y Togüí, y contenidos de medios a ideales ($> 20 \text{ mg/kg}$) en Santana y San José de Pare, para el cultivo de caña de azúcar para la producción de panela (Tarazona Parra, 2011).

En cuanto al error de las predicciones (figura 5.3), se observa que para toda el área de estudio no se presenta un patrón de sobreestimación o subestimación distintivo por departamentos o vertiente del río Suárez, con excepción de la DA, cuyas predicciones fueron sobreestimadas en los municipios de Santana, Suaita y San Benito, y subestimadas en los municipios de Togüí, San José de Pare y Moniquirá. Es importante notar que, para las propiedades de PT, pH y CICE, se aprecia una mayor área representada con color amarillo, lo que indica que los errores de predicción son bajos y brinda, así, confianza en estos productos del MDS en estas áreas (los productos del presente estudio se encuentran disponibles para libre acceso y consulta) (Varón-Ramírez, Vargas Díaz, & Murcia Contreras, 2022).

Conclusiones

A través de las herramientas del MDS, fue posible representar en el espacio las características de los suelos de la región de la HRS. Se identificó que propiedades como el contenido de arcilla, la densidad aparente, la porosidad total, el pH y la capacidad de intercambio catiónico efectivo tienen un patrón de distribución estrechamente ligado a la topografía y al material parental del área de estudio. Los modelos de regresión *kriging* y los bosques aleatorios lograron explicar la varianza de los datos; sin embargo, los modelos podrían aumentar su capacidad predictiva con la inclusión de una mayor cantidad de covariables, la exploración de otro tipo de modelos de aprendizaje de máquina o el aumento de las bases de datos de suelos. Propiedades como la estabilidad, el agua aprovechable, el fósforo disponible y la materia orgánica no demostraron correlación o dependencia con las covariables ambientales exploradas.

Agradecimientos

Los autores agradecen al Ministerio de Agricultura y Desarrollo Rural, a AGROSAVIA y al equipo de trabajo de ejecución del proyecto “243 recomendaciones técnicas preliminares de manejo de suelos en ladera para el sistema de producción de caña panelera en la HRS”.

Referencias

- Anschau, R., Flores, N., Carballo, S., & Hilbert, J. (2009). *Evaluación del potencial de producción de biocombustibles en Argentina, con criterios de sustentabilidad social, ecológica y económica, y gestión ordenada del territorio. El caso de la caña de azúcar y el bioetanol* [informe técnico]. Instituto Nacional de Tecnología Agropecuaria (INTA).
- Araujo-Carrillo, G. A., Varón-Ramírez, V. M., Jaramillo-Barrios, C. I., Estupiñan-Casallas, J. M., Silva-Arero, E. A., Gómez-Latorre, D. A., & Martínez-Maldonado, F. E. (2021). IRAKA: The first Colombian soil

- information system with digital soil mapping products. *Catena*, 196, artículo 104940. <https://doi.org/10.1016/j.catena.2020.104940>
- Baquero, J. E., Ralisch, R., Medina, C. de C., Filho, J. T., & Guimarães, M. de F. (2012). Soil physical properties and sugarcane root growth in a red oxiso. *Revista Brasileira de Ciência Do Solo*, 36(1), 63-70. <https://doi.org/10.1590/S0100-06832012000100007>
- Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R* (2.ª ed.). Springer. <https://doi.org/10.1007/978-1-4614-7618-4>
- Corporación Colombiana de Investigación Agropecuaria (Corpoica). (2017). Modelos de Adaptación Agroclimática – MAPA. *Boletín Técnico Corpoica*, (19). <https://repository.agrosavia.co/handle/20.500.12324/34920>
- Cortes, D. L., Camacho-Tamayo, J. H., & Giraldo, R. (2018). Spatial prediction of soil infiltration using functional geostatistics. *Acta Universitatis Carolinae Geographica*, 53(2), 149-155. <https://doi.org/10.14712/23361980.2018.15>
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. E. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45(2), artículo RG2004. <https://doi.org/10.1029/2005RG000183>
- García, H. R., Albarracín, L. C., Toscano Latorre, A., Santana, N. J., & Insuasty, O. (2007). *Guía tecnológica para el manejo integral del sistema productivo de la caña panelera*. Corporación Colombiana de Investigación Agropecuaria (Corpoica). <https://repository.agrosavia.co/handle/20.500.12324/13454>
- Giraldo Henao, R. (2006). *Introducción a la geoestadística. Teoría y aplicación* (1.ª ed.). Universidad Nacional de Colombia.
- Guevara, M., Olmedo, G. F., Stell, E., Yigini, Y., Aguilar Duarte, Y., Arellano Hernández, C., Arévalo, G. E., Arroyo-Cruz, C. E., Bolívar, A., Bunning, S., Bustamante Cañas, N., Cruz-Gaistardo, C. O., Davila, F., Dell Acqua, M., Encina, A., Figueredo Tacona, H., Fontes, F., Hernández Herrera, J. A., Ibelle Navarro, A. R., ... Vargas, R. (2018). No silver bullet for digital soil mapping: Country-specific soil organic carbon estimates across Latin America. *Soil*, 4(3), 173-193. <https://doi.org/10.5194/soil-4-173-2018>
- Gutiérrez Díaz, J. S., Ordoñez Delgado, N., Bolívar Gamboa, A., Bunning, S., Guevara, M., Medina, E., Olivera, C., Olmedo, G., Rodríguez, L. M.,

- Sevilla, V., & Vargas, R. (2020). Estimación del carbono orgánico en los suelos de ecosistema de páramo en Colombia. *Ecosistemas*, 29(1). <https://doi.org/10.7818/ECOS.1855>
- Hengl, T. (2006). Finding the right pixel size. *Computers & Geosciences*, 32(9), 1.283-1.298. <https://doi.org/10.1016/j.cageo.2005.11.008>
- Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, 12(2), artículo e0169748. <https://doi.org/10.1371/journal.pone.0169748>
- Instituto Geográfico Agustín Codazzi (IGAC). (2003). *Estudio general de suelos y zonificación de tierras: departamento de Santander*.
- Instituto Geográfico Agustín Codazzi (IGAC). (2005). *Estudio general de suelos y zonificación de tierras: departamento de Boyacá*.
- Instituto Geográfico Agustín Codazzi (IGAC). (2006). *Métodos analíticos del laboratorio de suelos* (6.ª ed.).
- Instituto Geográfico Agustín Codazzi (IGAC). (2015). *Suelos y tierras de Colombia*.
- Jaramillo-Barrios, C. I., Ortiz-Rico, A. F., Araujo Carrillo, G. A., & Varón-Ramírez, V. M. (2019). Establishment of homogeneous zones in a soil of alluvial origin. *Ciencia e Agrotecnología*, 43. <https://doi.org/10.1590/1413-7054201943026318>
- Kempen, B., Heuvelink, G. B. M., Brus, D. J., & Stoorvogel, J. J. (2010). Pedometric mapping of soil organic matter using a soil map with quantified uncertainty. *European Journal of Soil Science*, 61(3), 333-347. <https://doi.org/10.1111/j.1365-2389.2010.01232.x>
- Kuhn, M. (2020). *Caret: Classification and regression training (R package version 6.0-85)*. <https://cran.r-project.org/package=caret>
- Lal, R., Bouma, J., Brevik, E., Dawson, L., Field, D. J., Glaser, B., Hatano, R., Hartemink, A. E., Kosaki, T., Lascelles, B., Monger, C., Muggler, C., Ndzana, G. M., Norra, S., Pan, X., Paradelo, R., Reyes-Sánchez, L. B., Sandén, T., Singh, B. R., ... Zhang, J. (2021). Soils and sustainable development goals of the United Nations: An International Union of Soil Sciences perspective. *Geoderma Regional*, 25, artículo e00398. <https://doi.org/10.1016/j.geodrs.2021.e00398>

- Lin, L. I.-K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45(1), 255-268. <https://doi.org/10.2307/2532051>
- Malone, B. P., Jha, S. K., Minasny, B., & McBratney, A. B. (2016). Comparing regression-based digital soil mapping and multiple-point geostatistics for the spatial extrapolation of soil data. *Geoderma*, 262, 243-253. <https://doi.org/10.1016/j.geoderma.2015.08.037>
- Malone, B. P., Minasny, B., & McBratney, A. B. (2017). *Using R for digital soil mapping*. Springer. <https://doi.org/10.1007/978-3-319-44327-0>
- Manrique, R., Insuasty, O., Mora, C., Rodríguez, G., Blanco, R., Mejía, L., Pinto, J., & Sandoval, G. (2000). *Manual de caña de azúcar para la producción de panela*. Corporación Colombiana de Investigación Agropecuaria (Corpoica). <https://repository.agrosavia.co/handle/20.500.12324/18504>
- McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1-2), 3-52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Medina, C., Camacho-Tamayo, J. H., & Cortés, C. A. (2011). Soil penetration resistance analysis by multivariate and geostatistical methods. *Soil and Water Engineering*, 32(1), 91-101. <https://www.scielo.br/j/eagri/a/H6yXJxZXZJttJvBPdQY78Qn/?lang=en>
- Minasny, B., & McBratney, A. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264(parte B), 301-311. <https://doi.org/10.1016/j.geoderma.2015.07.017>
- Montenegro, H., & Malagón, D. (1990). *Propiedades físicas de los suelos*. Instituto Geográfico Agustín Codazzi (IGAC).
- Montenegro, M., Ruiz, H., & Tafur, H. (2010). Tiempo de uso del suelo cultivado con caña panelera y su efecto sobre algunas propiedades físicas en Sandoná, Nariño. *Revista de Agronomía*, 27(1), 85-92. <https://dialnet.unirioja.es/descarga/articulo/5104153.pdf>
- Odeh, I. O. A., McBratney, A. B., & Chittleborough, D. J. (1995). Further results on prediction of soil properties from terrain attributes: Heterotopic cokriging and regression-kriging. *Geoderma*, 67(3-4), 215-226. [https://doi.org/10.1016/0016-7061\(95\)00007-B](https://doi.org/10.1016/0016-7061(95)00007-B)
- Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO). (2018). *Soil organic carbon mapping* (2.ª ed.). <http://www.fao.org/documents/card/es/c/l8895EN/>

- Osorio, G. (2007). *Manual técnico buenas prácticas agrícolas y buenas prácticas de manufactura en la producción de caña y panela*. Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO); MANA; Corporación Colombiana de Investigación Agropecuaria (Corpoica).
- Poggio, L., & Gimona, A. (2017). 3D mapping of soil texture in Scotland. *Geoderma Regional*, 9, 5-16. <https://doi.org/10.1016/j.geodrs.2016.11.003>
- Soil Survey Staff. (2014). *Keys to soil taxonomy* (12.^a ed.). Natural Resources Conservation Service, USDA.
- Tarazona Parra, G. A. (2011). *Manejo fitosanitario del cultivo de la caña panelera: medidas para la temporada invernal*. Instituto Colombiano Agropecuario (ICA). <https://www.ica.gov.co/getattachment/6a54658e-1723-488d-a7ab-2f4baad793cb/Manejo-fitosanitario-del-cultivo-de-la-cana-panele.aspx>
- Vargas Díaz, R. E., Galindo Pacheco, J. R., Giraldo Henao, R., Varón Ramírez, V. M., Wilches Ortiz, W. A., & Franco Florez, C. V. (2023). Metodologías de interpolación y predicción espacial para el análisis de las propiedades físicas del suelo en la Hoya del Río Suárez (Colombia). *Siembra*, 10(1), artículo e4118. <https://doi.org/10.29166/siembra.v10i1.4118>
- Varón-Ramírez, V. M., Araujo-Carrillo, G. A., & Guevara Santamaría, M. A. (2022). Colombian soil texture: Building a spatial ensemble model. *Earth System Science Data*, 14(10), 4.719-4.741. <https://doi.org/10.5194/essd-14-4719-2022>
- Varón-Ramírez, V. M., Vargas Díaz, R. E., & Murcia Contreras, G. A. (2022). Spatial soil properties distribution in “Hoya del río Suárez” region in Colombia. *Environmental Data Initiative*. <https://doi.org/10.6073/pasta/4fd9a405bc7bbe2997f549d68989779d>
- Wadoux, A. M. J.-C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews*, 210, artículo 103359. <https://doi.org/10.1016/j.earscirev.2020.103359>
- Warrick, A. W., & Nielsen, D. R. (1980). Spatial variability of soil physical properties in the field. En D. Hillel (Ed.), *Applications of soil physics* (pp. 319-344). Academic Press. <https://doi.org/10.1016/B978-0-12-348580-9.50018-3>

- Willmott, C. J., Robeson, S. M., & Matsuura, K. (2012). A refined index of model performance. *International Journal of Climatology*, 32(13), 2.088-2.094. <https://doi.org/10.1002/joc.2419>
- Yang, S., Zhao, W., Liu, Y., Cherubini, F., Fu, B., & Pereira, P. (2020). Prioritizing sustainable development goals and linking them to ecosystem services: A global expert's knowledge evaluation. *Geography and Sustainability*, 1(4), 321-330. <https://doi.org/10.1016/j.geosus.2020.09.004>
- Zomer, R. J., Bossio, D. A., Sommer, R., & Verchot, L. V. (2017). Global sequestration potential of increased organic carbon in cropland soils. *Scientific Reports*, 7, artículo 15554. <https://doi.org/10.1038/s41598-017-15794-8>

