

PREDICCIÓN DE PROPIEDADES DEL SUELO CON KRIGING DE REGRESIÓN EN PAISAJES DE MONTAÑA DE LA CUENCA ALTA DEL RIO GUÁRICO, VENEZUELA

Ángel Rafael Valera^{1, 2*}, Jesús Arnaldo Viloria²; María Corina Pineda³

¹ Universidad Rómulo Gallegos, Centro de Investigación y Extensión en suelos y Aguas, San Juan de los Morros, Guárico. Venezuela, e-mail: angelvalera@unerg.edu.ve, <https://orcid.org/0000-0001-5500-1332e>

² Universidad Central de Venezuela, Facultad de Agronomía, Instituto de Edafología, Maracay, Aragua. Venezuela. viloriaj@agr.ucv.ve

³ Universidad Central de Venezuela, Facultad de Agronomía, Instituto de Edafología, Maracay, Aragua. Venezuela. maria.c.pineda@ucv.ve

* Autor de correspondencia

Recibido: 14 - 02 - 2023; **Aceptado:** 22 - 04 - 2023; **Publicado:** 30 - 06 - 2023

RESUMEN

La predicción de propiedades del suelo es importante para el estudio de las relaciones suelo-paisaje. En este estudio, se aplicó un método cuantitativo híbrido denominado kriging de regresión (RK), para generar modelos de la variación espacial de propiedades morfológicas, físicas y químicas de los suelos en áreas de montaña, en un sector de la cuenca alta del río Guárico, en el estado Aragua, Venezuela. En esta zona, las relaciones entre el suelo y el paisaje son complejas, por lo que no pueden ser caracterizadas completamente con modelos de regresión generales. El enfoque utilizado combina la técnica de regresión lineal múltiple de los datos de suelo y la información ambiental derivada de un modelo digital de elevación (MDE) y de una imagen satelital de 15 m de resolución espacial, con la interpolación de los errores obtenidos del modelo de regresión lineal múltiple. El método RK, aunado al desarrollo de los sistemas de información geográfica y al apoyo de información auxiliar de adecuada resolución espacial, ofrece nuevas oportunidades para superar las limitaciones de los levantamientos convencionales, con la finalidad de producir información edáfica de manera más eficiente, con adecuada precisión, exactitud y rapidez. Para el ajuste y desarrollo de los modelos RK se utilizó un conjunto de 133 perfiles de suelo, de los cuales el 75% se utilizó para la generación del modelo, y para la validación se empleó el 25% del total de perfiles. Las propiedades del suelo consideradas fueron espesor del horizonte A (Esp_A, cm), espesor del *solum* (Esp_AB, cm), profundidad efectiva (PEF, cm), contenido de esqueleto grueso (%EG), contenido de arena (%a), contenido de arcilla (%A), porcentaje de carbono orgánico (%CO), porcentaje de saturación con bases (PSB) y pH en agua 1:1, las cuales fueron seleccionadas por medio de un análisis de componentes principales. La confiabilidad de los modelos indicó una relación directa con la naturaleza de las variables, ya que el coeficiente de concordancia entre los valores estimados y observados fue de 61% para las propiedades morfológicas, 74% para las variables físicas, y 85% para las químicas.

Palabras clave: Predicción; kriging de regresión; modelo digital de elevación; imagen de satélite, variables auxiliares.

PREDICTION OF SOIL PROPERTIES WITH REGRESSION KRIGING IN MOUNTAIN LANDSCAPES OF HIGHT GUARICO RIVER BASIN, VENEZUELA

ABSTRACT

The prediction of soil properties is important for the study of soil-landscape relationships. In this study, a hybrid quantitative method called regression kriging (RK) was applied to generate models of the spatial variation of morphological, physical and chemical properties of soils in mountain areas, in a sector of the upper basin of the Guárico River, Aragua State, Venezuela. In this area, soil-landscape relationships are complex and cannot be fully characterized with general regression models. The approach used combines the multiple linear regression technique of soil data and environmental information derived from a digital elevation model (DEM) and a 15 m spatial resolution satellite image, with interpolation of the errors obtained from the multiple linear regression model. The RK method, together with the development of geographic information systems and the support of auxiliary information of adequate spatial resolution, offers new opportunities to overcome the limitations of conventional surveys in order to produce edaphic information more efficiently, with adequate precision, accuracy and speed. For the adjustment and development of the RK models, a set of 133 soil profiles were used, of which 72-75% were used for model generation, and 25% of the total number of profiles were used for validation. The soil properties considered were A horizon thickness (Esp_A, cm), solum thickness (Esp_AB, cm), effective depth (PEF, cm), coarse skeleton content (%CS), sand content (%a), clay content (%A), percent organic carbon (%OC), percent base saturation (PBS) and pH in 1:1 water, which were selected by principal component analysis. The reliability of the models indicated a direct relationship with the nature of the variables, as the coefficient of agreement between estimated and observed values was 61% for morphological properties, 74% for physical variables, and 85% for chemical variables.

Keywords: Prediction; regression kriging; digital elevation model; image of satellite; auxiliary variables.

INTRODUCCIÓN

Se han aplicado numerosos modelos estadísticos para la interpolación de propiedades del suelo, entre los que destacan los métodos geoestadísticos, los cuales son exigentes en cuanto al número de muestras y la reducida extensión geográfica que representan. Uno de los desarrollos metodológicos más significativos para la predicción de propiedades del suelo son los métodos que combinan la regresión lineal múltiple y la interpolación de los residuos (kriging de regresión o RK) (Hengl *et al.*, 2007; Sun *et al.*, 2012; Zhu *et al.*, 2022)). La técnica kriging de regresión constituye una amplia base científica para el fortalecimiento de la cartografía digital de suelos. El modelo RK puede desempeñar un papel importante en la geoestadística, debido a que muchas covariables están disponibles con el avance en la teledetección y tecnologías de posicionamiento (Sun *et al.* 2012a). Muchos estudios han demostrado que RK es fácil de usar y su precisión a menudo supera a la regresión lineal ordinaria, al *kriging* ordinario (OK) (Minasny y McBratney, 2007) y al *co-kriging*. Al respecto, Bishop y McBratney (2001) encontraron que RK es más preciso en la predicción de la CIC del suelo; ha sido de gran importancia en la predicción de la profundidad efectiva del suelo (Penížek y Borůvka. 2006), y es más apropiado cuando la información auxiliar puede explicar parte de la variación de la variable estimada (Hengl *et al.*, 2007).

El método de interpolación de la regresión (*regression kriging*, RK) se ha convertido en una herramienta importante de la geoestadística, debido a la amplia disponibilidad de covariables de alta resolución espacial con el avance de la teledetección y las

tecnologías de posicionamiento satelital (Sun *et al.*, 2012). Las técnicas RK utilizan un modelo determinístico para estimar el valor de la variable de interés. De esta manera, se emplean los valores medidos en campo o laboratorio para desarrollar la calibración del modelo, posteriormente se analiza la correlación espacial de los residuos, y finalmente se combina el ajuste estadístico y el modelo determinístico. Es decir, el RK utiliza la regresión de la información auxiliar y luego emplea OK para interpolar los residuales del modelo de regresión (Hengl *et al.*, 2007).

El modelo RK puede desempeñar un papel importante en la geoestadística, debido a que muchas variables auxiliares están disponibles con el avance en teledetección y tecnologías de posicionamiento global (Sun *et al.*, 2012; Zhu *et al.*, 2022). Muchos estudios han demostrado que RK es fácil de usar y su precisión a menudo supera a la regresión lineal ordinaria, al *kriging* ordinario (Minasny y McBratney, 2007) y al *co-kriging*. Al respecto, varios estudios han comparado los métodos RK y OK para distintas propiedades del suelo. Bishop y McBratney (2001) encontraron que RK fue más preciso en la predicción de la capacidad de intercambio catiónico del suelo y Hengl *et al.* (2007) considera que RK es mejor cuando la información auxiliar puede explicar parte de la variación de la variable estimada. Sin embargo, Kravchenko y Robertson (2007) consideran que RK no es capaz de superar al OK en la predicción de propiedades del suelo que presentan una fuerte dependencia espacial. En otros casos, el uso de datos auxiliares del terreno ha mejorado la predicción de la profundidad efectiva del suelo, aunque la mejora ha sido relativamente pequeña debido a la baja correlación de la variable estimada con los atributos del terreno utilizados (Penížek y Borůvka, 2006).

En áreas de montaña como la cuenca del río Caramacate, las relaciones entre el desarrollo del suelo y las condiciones de formación son muy complejas, y en la mayoría de los casos son desconocidas para un área específica, por lo que dichas relaciones no pueden ser caracterizadas completamente con modelos de regresión generales. Alternativamente, los enfoques híbridos como el *kriging* de regresión, el cual combina un modelo de regresión lineal con el *kriging* ordinario de los residuos de la regresión, constituye un método no estacionario de gran utilidad para mejorar las predicciones de propiedades de los suelos sobre áreas extensas, considerando atributos de los paisajes.

El objetivo de esta investigación consistió en aplicar el algoritmo RK para la predicción espacial de propiedades morfológicas, físicas y químicas del suelo en áreas de montaña, donde existe una alta complejidad en las relaciones suelo-paisaje debido a la influencia del relieve, la susceptibilidad a movimientos en masa y el uso de la tierra, lo que ha generado una alta heterogeneidad edáfica.

MATERIALES Y MÉTODOS

Área de Estudio

La investigación se realizó en un sector de la cuenca alta del río Guárico, específicamente en la cuenca del río Caramacate, la cual está ubicada entre los

municipios Santos Michelena y San Sebastián de los Reyes del estado Aragua (Venezuela), entre las coordenadas geográficas 9,55 a 10,09° Norte y -67,12 a -67,03° Oeste, (Figura 1). La cuenca del río Caramacate representa el 8,5% de la cuenca alta del río Guárico, de la cual es tributaria. Para el ensayo de agrupamiento de los suelos se seleccionó un área muestra de 6.760 ha, donde el paisaje está dominado por laderas de montaña con pendientes del 40%. La geología está representada por rocas metavolcánicas y basaltos de la formación El Caño-El Chino, por metalavas máficas de la formación El Carmen (Urbani y Rodríguez 2004), sedimentos aluviales acarreados por el río Caramacate y los mantos coluvio-aluviales del Grupo Villa de Cura. La zona presenta una altitud comprendida entre 334 a 1.405 msnm, con una precipitación media anual que oscila entre 1.100 y 1.400 mm y una temperatura media anual varía entre 22 y 26°C. La vegetación herbácea ocupa más del 50% de la cobertura del sector, como producto de la deforestación y las quemas para el uso ganadero en condiciones extensivas. Los suelos en su mayoría son Entisoles, Inceptisoles y Alfisoles, cuya variabilidad se ha incrementado por el uso de la tierra dominante basado en ganadería extensiva y la incidencia de los movimientos en masa (Pineda *et al.* 2011a).

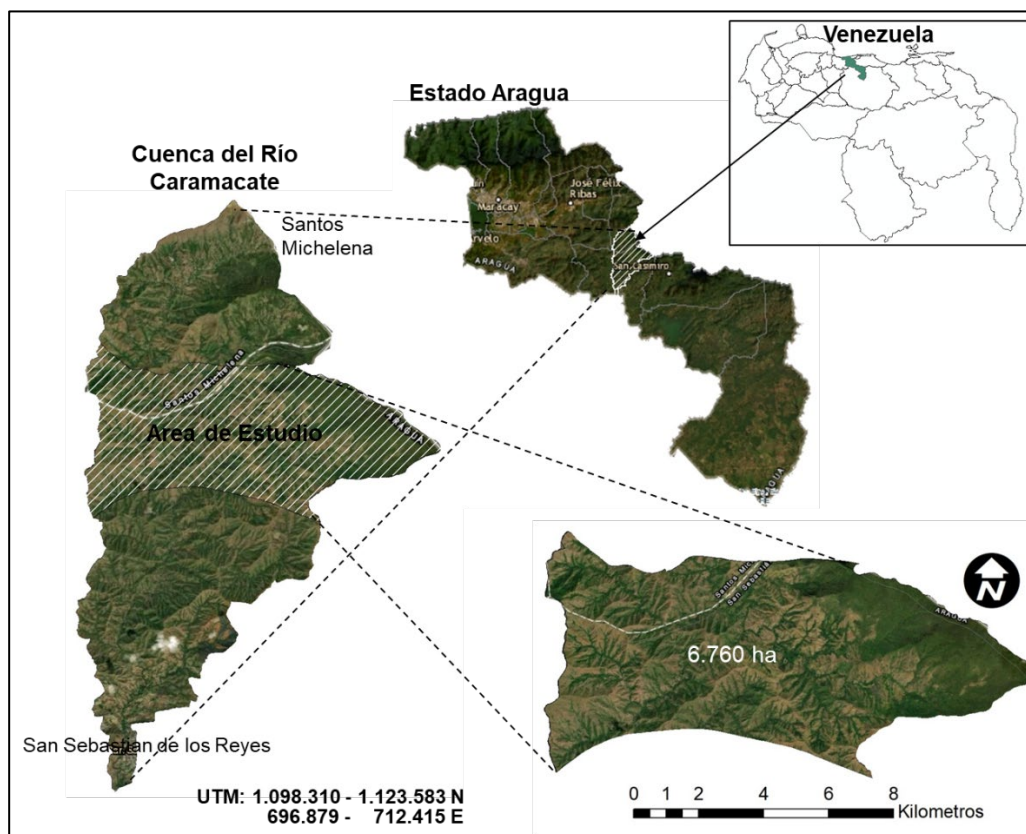


Figura 1. Ubicación relativa del área de estudio dentro de la cuenca del río Guárico, estado Aragua, Venezuela.

Datos de suelo

El conjunto de datos en la zona de estudio seleccionada está conformado por 133 perfiles de suelo ubicados en diferentes posiciones geomorfológicas de laderas de paisajes de montaña, en una superficie de 6.760 ha (Pineda *et al.* 2011b; Valera 2015). Los sitios de muestreo corresponden a observaciones de calicatas, cortes, y barrenos agrológicos, equivalentes a una densidad de muestreo de 2 observaciones/km² (1 observación/0,5 km² o 1 observaciones/2 cm² de mapa a escala 1:50.000). La distribución del conjunto de perfiles de suelo se presenta en la Figura 2, donde los círculos (color rojo) corresponden a los perfiles empleados en la obtención de los modelos y los rectángulos (color amarillo) se refieren a los perfiles utilizados en la validación de modelos. Asimismo, se puede observar el límite de la zona de estudio sobre las clases de terreno, y algunos perfiles de suelo con evaluaciones morfológicas (círculos color verde) dentro y fuera del área, utilizadas como chequeos complementarios.

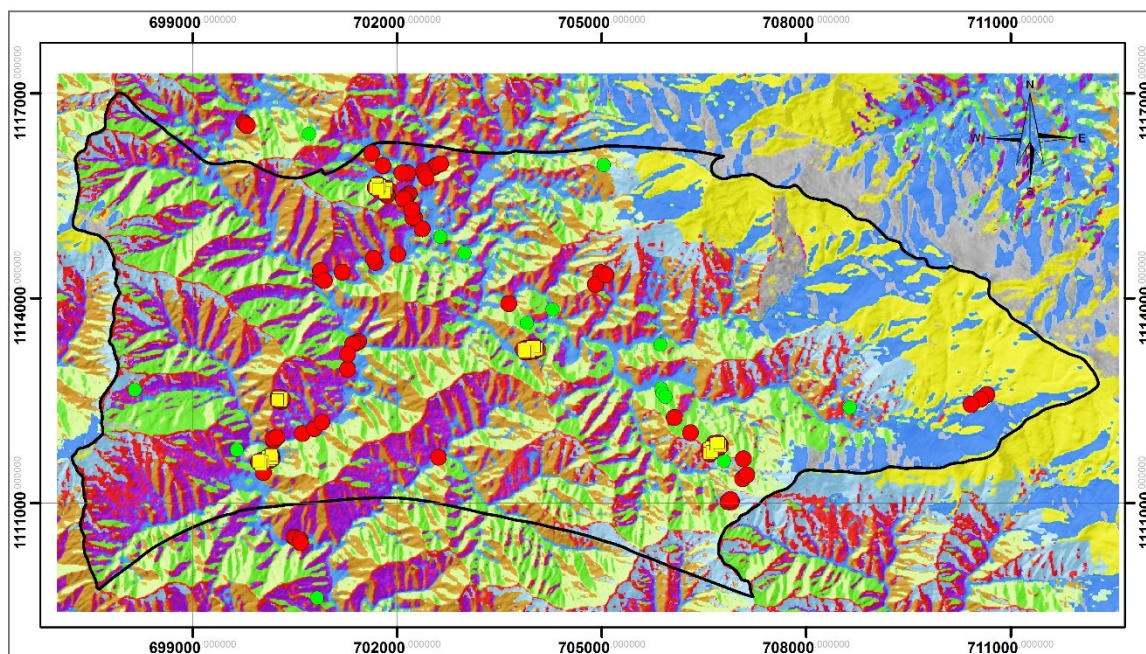


Figura 2. Distribución de los sitios de muestreo en el sector de estudio de la cuenca del río Caramacate.

Las propiedades del suelo que se consideraron en esta investigación fueron espesor del horizonte A (Esp_A, cm), espesor del *solum* (Esp_AB, cm), profundidad efectiva (PEF, cm), contenido de esqueleto grueso (%EG), contenido de arena (%a), contenido de arcilla (%A), porcentaje de carbono orgánico (%CO), porcentaje de saturación con bases (PSB) y pH en agua 1:1. Estas variables fueron seleccionadas por medio de un análisis de componentes principales (ACP), que indicó que ellas se relacionan con los primeros ocho componentes, los cuales explican el 78% de la varianza que ocurre en los suelos.

Variables auxiliares

Las covariables ambientales utilizadas se corresponden al modelo de factores de formación de suelos (Jenny, 1941) y al modelo geoespacial multivariado scorpan, formulado por McBratney et al. (2003). Se empleó un modelo digital de elevación (MDE) de 15 m de resolución espacial, a partir del cual se derivaron los diversos parámetros morfométricos (variables auxiliares o variables topográficas) tales como: altitud (Zevenbergen y Thorne, 1987; Burrough y McDonnell, 1998), gradiente y orientación de la pendiente (Moore et al., 1991), posición relativa (Verbrugge, 2006), perfil de curvatura (curvatura vertical o longitudinal), y plano de curvatura (curvatura horizontal o transversal) (Moore et al., 1993), área de captación (Tarboton et al., 1991) y el índice topográfico de humedad (Wilson y Gallant, 2000). Adicionalmente, se incorporó un mapa de precipitación realizado por kriging ordinario (Pineda et al., 2011a), y el índice de vegetación de diferencia normalizada (NDVI) calculado a partir de una imagen Spot 5 multiespectral 658-330 (08/03/2018). Estas variables auxiliares fueron empleadas como parámetros de correlación ambiental para la estimación de las propiedades específicas del suelo en áreas no muestreadas

Análisis estadístico

Método de predicción de propiedades del suelo

El método de predicción de las propiedades del suelo por kriging de regresión (RK) incluye tres etapas secuenciales: i) modelos de regresión lineal múltiple entre variables del suelo y variables auxiliares en formato *raster*, ii) modelos de interpolación de los residuos de la regresión con kriging ordinario, iii) modelos híbridos, constituidos por la sumatoria de los mapas de regresión lineal y los mapas de interpolación de residuos o errores de la regresión. Es decir, con RK la estimación se basa en la predicción en forma separada de la tendencia y los residuos, y luego se suman ambos componentes.

Modelo kriging de regresión

El kriging de regresión pertenece a los métodos geoestadísticos no estacionarios. Es una técnica adecuada para la predicción de una variable primaria cuando las variables auxiliares están disponibles en formato *raster* y se encuentran correlacionadas con la variable de interés (Hengl et al., 2004). La concepción fundamental del enfoque RK tiene sus raíces en los trabajos de Matheron (1969) quien propuso el modelo universal de variación espacial, donde las características de la varianza de ciertos elementos y procesos de un sistema son el resultado de componentes determinísticos y estocásticos. Estos componentes pueden ser representados por separado a través de la ecuación:

$$Z(s) = m(s) + \varepsilon'(s) + \varepsilon''(1) \quad (1)$$

donde Z es la variable a predecir en una localización determinada (s), $m(s)$ es el componente determinístico (tendencia espacial global), $\varepsilon'(s)$ es el componente aleatorio (tendencia local) correlacionado espacialmente (variación espacialmente dependiente) y ε'' es el ruido puro (variación espacialmente independiente), el cual

incluye las variaciones a nivel de microescala y el error de medición.

En RK, el valor de una propiedad en un lugar nuevo y no muestreado ($\mathbf{s}_0$) se puede predecir por combinación de los componentes determinísticos y estocásticos; es decir, sumando la tendencia espacial y los componentes aleatorios o residuales (Odeh *et al.*, 1994) mediante la siguiente ecuación:

$$\hat{Z}(\mathbf{s}_0) = \hat{m}(\mathbf{s}_0) + \hat{e}(\mathbf{s}_0) \quad (2)$$

donde ($\mathbf{m}$) es ajustada mediante el análisis de regresión lineal múltiple y los residuales ($\hat{\mathbf{e}}$) son interpolados utilizando kriging ordinario (Bishop y McBratney, 2001; Hengl *et al.*, 2007; Zhu y Lin, 2010), como se señala a continuación:

$$\hat{Z}(\mathbf{s}_0) = \sum_{k=0}^p \hat{B}_k \cdot q_k(\mathbf{s}_0) + \sum_{i=1}^n w_i \cdot e(\mathbf{s}_i); \quad q_0(\mathbf{s}_0) = 1 \quad (3)$$

donde $\mathbf{B}_k$ son los coeficientes estimados del modelo de dispersión o la tendencia, q_k es la variable auxiliar explicativa o predictiva en la ubicación $\mathbf{s}_0$; p es el número de variables auxiliares, $w_i(\mathbf{s}_0)$ son los pesos determinados por la función de covarianza, y $\mathbf{e}(\mathbf{s}_i)$ son los residuos de la regresión en una posición $\mathbf{s}_i$. Los coeficientes del modelo de tendencia (globales) pueden ser estimados de las muestras por algún método de ajuste (e.g. mínimos cuadrados ordinarios) para encontrar la correlación espacial de los residuales (Cressie, 1993).

Modelo de regresión lineal múltiple

El propósito general de la regresión lineal múltiple (RLM) es encontrar una relación entre varias variables independientes o predictivas y una variable dependiente o criterio. Es posible construir una ecuación lineal que incluya todas las variables. En general, los procedimientos de regresión múltiple se estiman mediante una ecuación lineal de la forma:

$$Z = \beta_0 + \beta_1 \cdot q_1 + \beta_2 \cdot q_2 + \dots + \beta_k \cdot q_k + \varepsilon \quad (4)$$

donde: β_0 es la intersección o el término constante; $\beta_1, \beta_2, \dots, \beta_k$ son los coeficientes de regresión que representan la contribución de cada variable independiente o explicativa ($q_1, q_2, q_3 \dots$) en la predicción de la variable dependiente (z). El término ε se refiere al error asociado a la medición con media igual a cero y varianza constante (es equivalente a las perturbaciones aleatorias). Los coeficientes de regresión del modelo suelo-paisaje no son conocidos *a priori*, pero pueden ser estimados utilizando el método de los mínimos cuadrados ordinarios. La fórmula corresponde a la primera parte de la ecuación (3):

$$\check{z}(\mathbf{s}_0) = \sum_{k=0}^p \beta_k \cdot q_k(\mathbf{s}_0); \quad q_0(\mathbf{s}_0) = 1 \quad (5)$$

Para la estimación de las propiedades del suelo con RLM, se emplearon los mapas de las variables auxiliares continuas seleccionadas, y para la selección de las variables auxiliares predictoras se empleó el procedimiento “paso a paso” del

programa InfoStat (Di Rienzo et al., 2011), el cual utiliza la introducción y eliminación progresiva de las variables, reteniendo las más significativas. En el análisis de los datos con RLM se consideraron los siguientes supuestos: i) linealidad: la relación entre los valores de las variables es lineal; ii) homocedasticidad: todos los errores tienen la misma varianza; iii) independencia: los errores o las perturbaciones aleatorias de las variables explicativas son independientes entre sí; iv) normalidad: los errores tienen distribución normal, y v) las variables explicativas q_k se obtienen sin errores de medida, es decir los errores tienen una esperanza matemática igual a cero (Rojo, 2007; Tusell, 2011).

La normalidad se evaluó mediante el estadístico W^* de Shapiro-Wilks modificado por Mahibbur y Govindarajulu (1997). El estadístico también permitió probar el supuesto que evalúa si el valor medio de los residuos es igual a cero, para lo cual se utilizó la prueba de hipótesis que considera que la media de las observaciones tiene una distribución normal con media cero. Para la evaluación del supuesto de homogeneidad de varianzas, se realizaron gráficos de dispersión de los residuos tipificados versus los valores predichos. Para el logro de los supuestos mencionados, se consideró el estadístico basado en la Distancia de Cook (D_i), como prueba para la identificación de puntos influyentes o valores atípicos que presentan un valor D_i superior a la unidad. El cálculo de D_i se basó en el análisis de regresión, considerando inicialmente cada observación en el modelo RLM y el posterior retiro de cada elemento atípico de la muestra, y su influencia en los resultados (Cook, 1977).

Modelo kriging ordinario de los residuos

El kriging ordinario es uno de los métodos más elementales del kriging (Webster y Oliver, 1990). Proporciona la estimación en la ubicación no observada de la variable z , en base a la media ponderada de los sitios vecinos observados dentro de un área determinada. La teoría se deriva de las variables regionalizadas y puede ser descrita brevemente por considerar una función aleatoria intrínseca denotada por $z(s_i)$, donde s_i representa todas las ubicaciones de la muestra, $i=1, \dots, n$. Una estimación de la media ponderada propuesta para el predictor kriging ordinario en un sitio sin muestrear $z(s_0)$ se define por la fórmula que corresponde a la segunda parte de la ecuación (3):

$$z(s_0) = \sum_{i=1}^n w_i z(s_i) \quad (6)$$

dónde: w_i son los pesos asignados a cada uno de los sitios de muestreo observados, equivalentes a los pesos de la función de covarianza entre pares de puntos, los cuales se estiman mediante un variograma. El variograma es definido por la función de la semivarianza $[\gamma(h)]$, que está relacionada con la función de covarianza a través de $\gamma(h) = C_0 - C(h)$. C_0 (varianza nugget), C_1 (varianza estructural) son los parámetros del semivariograma, y $|h|$ es la distancia euclidiana entre los pares de puntos. La semivarianza se estima con la siguiente expresión, según Upchurch y Edmonds (1991):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{N(h)} [z(x_i) - z(x_{i+h})]^2 \quad (7)$$

donde **N** es el número de pares de puntos separados por una determinada distancia, **h**; **z (x_i)** es el valor de la variable en una localidad, **x**; **z (x_{i+h})** es el valor que toma la variable en otra localidad ubicada a una distancia **h** de **x**.

Para la estimación de las propiedades que caracterizan a los variogramas de los residuos y el ajuste a modelos matemáticos, se empleó el programa Vesper 1.6 (Minasny *et al.*, 2002). La varianza relativa (NR), basada en la relación de la varianza nugget y el umbral, se utilizó para cuantificar el grado de dependencia espacial (Cambardella *et al.*, 2004). El criterio de información de Akaike (AIC) se utilizó para decidir acerca del ajuste de los puntos de observación al modelo matemático más apropiado sin pérdida de información (Hengl *et al.*, 2007).

Evaluación de los modelos de propiedades del suelo

Para el ajuste y desarrollo de los modelos se utilizó un conjunto de datos equivalentes al 75% del total de perfiles de suelo ($n=133$), y para la validación se empleó un conjunto de datos independientes correspondientes al 25% del total de perfiles ($n=33$).

Bondad de ajuste de los modelos de predicción (Calibración)

Para la calibración de los modelos de regresión lineal múltiple y los mapas de propiedades del suelo obtenidos por kriging de regresión, se utilizó el coeficiente de determinación múltiple ajustado (R^2_a) por el número de variables predictivas (Hengl *et al.*, 2004) y la suma de cuadrado de los errores (**SSE**, *sum of squared error*), equivalente a la variación residual o variación no explicada por el modelo. El R^2_a se determina con la siguiente fórmula:

$$R^2_a = 1 - \left(\frac{n-1}{n-p} \right) \cdot \frac{SSE}{SST} = \left(\frac{n-1}{n-p} \right) \cdot (1 - R^2) \quad (8)$$

donde **SSE** es la suma de cuadrados de los errores (residuos), **SST** es la suma de cuadrados totales; **n** es el número de datos y **p** es equivalente al número de variables regresoras consideradas en el modelo.

Confiabilidad de los modelos de predicción (Validación)

Para la evaluación de la exactitud de los modelos de predicción de las propiedades de los suelos generados, se utilizaron varios índices, tales como: el error medio absoluto (**MAE**, *mean absolute error*), el error medio (**ME**, *mean error*), raíz del error cuadrático medio (**RMSE**, *root mean squared error*) y el coeficiente de concordancia (**AC**, *agreement coefficient*). Los índices ME, MAE y RMSE contribuyen al análisis de los resultados indicando el error en los valores de la propiedad del suelo de interés (Hengl *et al.*, 2004). El ME evalúa el error sistemático e indica la presencia de subestimación o sobrestimación del modelo, el MAE garantiza que el resultado del

error sea estrictamente positivo, y el RMSE evalúa la precisión de la predicción. Los valores de RMSE, ME y MAE iguales a 0 indican un ajuste perfecto. Cuando los valores de estos índices son inferiores a la *desviación estándar (S)* (Marcheti et al., 2010) o a la mitad de la desviación de los datos medidos (Singh et al., 2004), pueden ser considerados bajos, e indican que los modelos son adecuados para evaluar las variables. Las ecuaciones que definen estos índices son las siguientes:

$$ME = \frac{1}{n} \sum_{i=1}^n (O_i - E_i) \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |O_i - E_i| \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - E_i)^2} \quad (11)$$

donde n es el número de observaciones; O_i es el valor observado o medido, y E_i es el valor estimado o inferido.

El coeficiente de concordancia (AC) es una medida normalizada del grado de error de los modelos predicción y varía entre 0 y 1. Un valor calculado de 1 indica un acuerdo perfecto entre los valores medidos y predichos, y un valor de 0 indica que no hay acuerdo o existe un desacuerdo total entre los valores estimados y los observados (Willmott, 1981). El índice representa la relación entre el error de la raíz cuadrática media y el error potencial (PE) (Willmott, 1984). El PE es definido como la suma de los cuadrados de los valores absolutos de las distancias de los valores predichos y el valor de la media observada, y la distancia entre los valores observados y el valor de la media observada; tal como se indica en las siguientes fórmulas:

$$AC = 1 - \frac{n(RMSE)^2}{PE} \quad (12)$$

$$PE = \sum_{j=1}^n (|E_i - \bar{O}| + |O_i - \bar{O}|)^2 \quad (13)$$

donde n es el número de observaciones y PE es el error potencial de la varianza. $\bar{O}$ es el valor medio observado y E_i y O_i son los valores estimados y observados, respectivamente. El índice AC tiene la ventaja de escalar la magnitud de las variables reteniendo el significado de la información, y no amplifica los valores extremos.

El índice no es una medida de correlación, sino una medida del grado en el cual las predicciones del modelo están libres de errores (Willmott et al., 1985). Además, es más apropiado para la investigación en casos de validación, donde se requiere comparar los valores observados y predichos por un determinado modelo (Ji y Gallo, 2006).

RESULTADOS Y DISCUSIÓN

Análisis de regresión lineal múltiple

Los resultados de la tendencia espacial general a través de los modelos de regresión lineal múltiple, muestran las relaciones entre las variables edáficas y las covariables ambientales, obteniéndose los coeficientes de regresión estimados mediante el método de los mínimos cuadrados ordinarios (Cuadro 1). Estos resultados coinciden con los obtenidos en otras investigaciones (Hengl *et al.*, 2004; Zhu y Lin, 2010; Marchetti *et al.*, 2010; Sun *et al.*, 2012a; Olmedo *et al.*, 2012), donde los valores del índice R^2 se encuentran en rangos que explican entre el 20 y 40% de la variabilidad de los suelos.

Todos los modelos de variación espacial generados con RLM manifestaron una bondad de ajuste no satisfactoria para los datos empleados, e indicaron la necesidad del análisis de los residuos, que representa la tendencia espacial local observada en las propiedades de los suelos.

Cuadro 1. Modelos de predicción de propiedades del suelo basados en regresión lineal múltiple.

Variable	Ecuación de regresión lineal	R ² a
Esp A	$-11 + (2,31 * ITH) + (0,01 * Altitud) + (49,87 * Plano C)$	0,20
Esp AB	$-82,65 + (8,79 * ITH) + (0,09 * Altitud) - (48,06 * Pendiente) + (177,37 * Plano C)$	0,42
PEF	$-38,51 + (7,69 * ITH) + (0,09 * Altitud) - (75,86 * Pendiente)$	0,26
%EG	$25,79 + (0,03 * Altitud) - (2,4 * [ITH])$	0,16
%a	$13,64 + (1,2 * Orientación) + (16,72 * PR)$	0,14
%A	$6,27 + (0,02 * Altitud) - (9,86 * [Pendiente]) + (14,08 * [NDVI])$	0,21
%CO	$1,65 + (1,02 * NDVI)$	0,10
PSB	$66,67 - (142,04 * Perfil C)$	0,14
pH	$6,25 - (0,0013 * Altitud)$	0,41

Esp_A: Espesor horizonte A (cm), Esp_AB: Espesor del *solum* (cm), PEF: Profundidad efectiva (cm), EG: Esqueleto grueso (%), a: Porcentaje arena, A: Porcentaje Arcilla, CO: Porcentaje Carbono orgánico, PSB: Porcentaje de saturación con bases. Pendiente (rad), Orientación (rad), PP: Precipitación (mm), ITH: Índice Topográfico de Humedad, PR: Posición Relativa, Perfil C: Perfil de curvatura (m/m²), Plano C: Plano de curvatura (m/m²), NDVI: Índice de vegetación de diferencia normalizada.

Análisis geoestadístico de los residuos

En cuanto a la interpolación de los residuos de la RLM, el análisis estructural de los puntos de observación proporcionó los resultados indicados en el Cuadro 2. Los semivariogramas fueron ajustados principalmente con modelos exponenciales y esféricos, con un alcance promedio de 315 m para las propiedades del suelo con alta y moderada dependencia espacial. Los residuales de las propiedades morfológicas presentan una alta dependencia espacial con una varianza relativa (NR) de 7,3% en promedio, las propiedades químicas tienen una NR de 26,8% y las físicas tienen una moderada dependencia espacial, con una NR promedio de 55,3%.

Cuadro 2. Parámetros geoestadísticos obtenidos de los variogramas de los residuales de las variables edáficas en el área estudiada.

Propiedad	Modelo	C ₀	C ₁	C ₀ + C ₁	A ₁ (m)	AIC	NR (%)
Esp A (cm)	Exponencial	2,5	32,4	34,9	137	74	7,3
Esp AB (cm)	Exponencial	59,1	420,6	479,7	208	149	12,3
PEF (cm)	Esférico	27,2	1117,1	1144,3	302	135	2,4
%EG	Exponencial	193,7	132,7	326,4	363	118	59,3
%a	Exponencial	78,7	119,9	198,6	293	113	39,6
%A	Esférico	46,7	23,1	69,7	227	99	66,9
%CO	Esférico	0,01	0,4	0,4	718	21	1,8
PSB	Exponencial	82,0	242,5	324,5	272	148	25,3
pH	Esférico	0,05	0,04	0,08	709	58	53,3

Esp_A: Espesor A, Esp_AB: Espesor del *solum*, PEF: Profundidad efectiva, EG: Esqueleto grueso, a: arena, A: Arcilla, CO: Carbono orgánico, PSB: Porcentaje de saturación con bases. C₀: Varianza Nugget, C₁: Varianza Estructural, C₀+C₁: Umbral o Sill, A₁: Alcance o rango, NR (Nugget relativa): (C₀/C₀+C₁)*100, AIC: Criterio de Información de Akaike.

Los semivariogramas de los residuos de las propiedades de los suelos tienen valores de efecto pepita (varianza residual) que representan menos de la mitad de la varianza total, a excepción de %EG, %A y pH, las cuales superan ligeramente el 50% de la varianza relativa.

Evaluación de la bondad de ajuste de los modelos RK

La integración de los modelos estimados con RLM y la interpolación de residuos dieron como resultado un mapa final para cada propiedad del suelo, con evidentes expresiones geográficas de las variables utilizadas como predictoras. En cuanto a los resultados de la calibración de los modelos obtenidos con RK (Cuadro 3), la estimación del R²_a, indicó que las predicciones dependen fuertemente de la naturaleza de las propiedades, ya que las variables morfológicas explican el 62% de la variabilidad presente, las físicas explican en promedio el 46%, y las propiedades químicas explican el 56% de la variación de los suelos, lográndose una aceptable bondad de ajuste de los modelos obtenidos con el número de perfiles de suelos empleados en el estudio.

Los modelos obtenidos con el conjunto de datos de calibración, explicaron más del 55% de la variación edáfica para el sector considerado en la cuenca del río Caramacate. También, la variación residual (SSE) o las variaciones no explicadas por el modelo, presentan grandes diferencias debido a la magnitud de los valores de la propiedad considerada. Las varianzas residuales de las variables morfológicas son más grandes, indicando que el grado de dependencia entre los valores observados y estimados por los modelos presenta un menor ajuste. Por otra parte, la SSE es más pequeña en la predicción de propiedades químicas (error cuadrático medio cercano a cero), donde la varianza explicada es mayor, y por consiguiente la calibración de los modelos fue más favorable en la predicción.

Cuadro 3. Indicadores de la bondad de ajuste de los modelos de predicción de propiedades del suelo con RK.

Propiedad	N	R ² a	SSE
Esp A (cm)	96	0,39	3.132
Esp AB (cm)	96	0,65	28.394
PEF (cm)	96	0,76	48.718
EG (%)	90	0,47	14.551
a (%)	80	0,46	2.491
A (%)	97	0,41	2.652
CO (%)	92	0,41	6,7
PSB (%)	62	0,65	12.330
pH	78	0,71	3,9

Esp_A: Espesor A, Esp_AB: Espesor del *solum*, PEF: Profundidad efectiva, EG: Esqueleto grueso, a: arena, A: Arcilla, CO: Carbono orgánico, PSB: Porcentaje de saturación con bases. R²a: ajustado con las variables regresoras. SSE: Suma de cuadrado de los errores, Esp_A: Espesor A, Esp_AB: Espesor del *solum*, PEF: Profundidad efectiva, EG: Esqueleto grueso, a: arena, A: Arcilla, CO: Carbono orgánico, PSB: Porcentaje de saturación con bases.

La integración de los modelos estimados con RLM y la interpolación de residuos dieron como resultado un mapa final para cada propiedad del suelo, con notables expresiones de las variables utilizadas como predictoras, que también expresaron cierto grado de incertidumbre. La Figura 3 ilustra gráficamente la aplicación de RK a la cartografía digital de las propiedades del suelo, utilizando como ejemplo el pH del horizonte superficial. RK modela, en primer lugar, la tendencia espacial global de la propiedad del suelo por medio del análisis de RLM, para producir un mapa del componente determinístico de la variación (Figura 3a). En segundo lugar, muestra el resultado de la interpolación por kriging ordinario de los residuos del modelo de RLM, para modelar la variación aleatoria local, remanente (Figura 3b). Finalmente, combina los dos modelos anteriores en un mapa final, obtenido a través de la suma de la tendencia espacial global y el componente aleatorio local (Figura 3c).

Evaluación de la confiabilidad de los modelos RK

El Cuadro 4 muestra los resultados de la evaluación de los modelos obtenidos con RK, realizada con base en el conjunto de datos de validación. En el caso del MAE y el ME, la mayoría de los valores son cercanos a cero, lo que indica una predicción imparcial no sesgada o con muy poco sesgo, aunque la excepción está dada por %EG y PEF, donde el error sistemático es mayor. También se destaca que los valores de RMSE son más pequeños que la desviación estándar de los valores observados de las muestras (S), lo cual sugiere que la evaluación de la predicción de las propiedades del suelo en la zona de interés es confiable. La excepción está dada por el Esp A y %A, cuyo valor supera ligeramente los valores de la desviación de los datos observados.

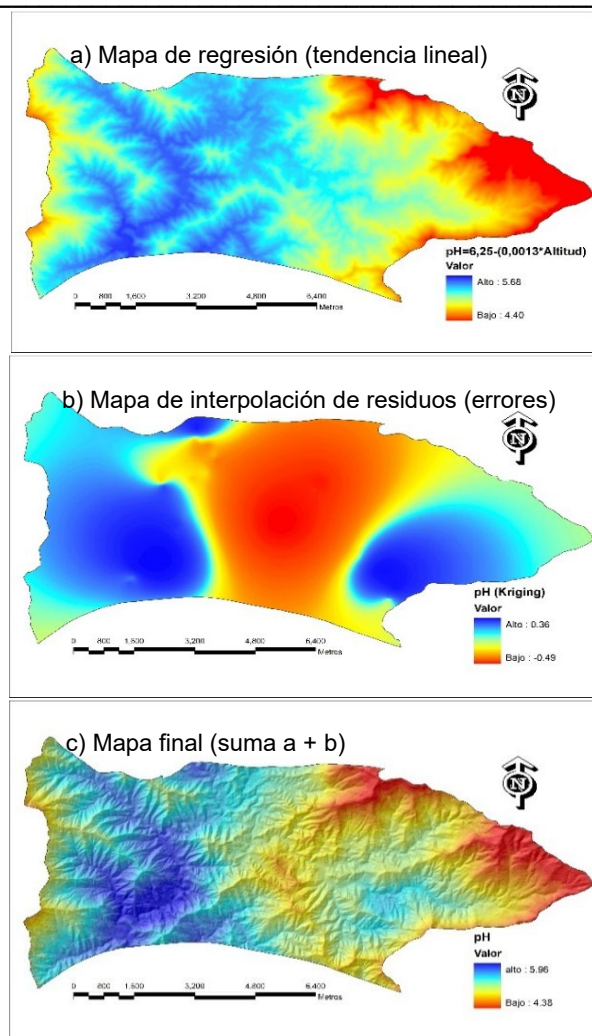


Figura 3. Representación del procedimiento *kriging* de regresión en la estimación del pH de los suelos en el área estudiada.

Cuadro 4. Evaluación de la exactitud de los modelos de predicción de propiedades del suelo con RK.

Propiedad	MAE	RMSE	AC	S	ME
Esp A (cm)	2,53	3,15	0,51	2,69	-0,04
Esp AB (cm)	8,29	12,2	0,73	14,1	1,59
PEF (cm)	12,5	16,1	0,60	14,6	3,99
EG (%)	10,9	14,2	0,82	18,7	3,28
a (%)	4,34	6,11	0,86	9,07	-1,72
A (%)	3,02	3,86	0,55	4,22	0,28
CO (%)	0,09	0,11	0,89	0,17	0,03
PSB (%)	8,50	11,6	0,75	12,0	0,87
pH	0,11	0,17	0,92	0,30	-0,05

Esp_A: Espesor A, Esp_AB: Espesor del *solum*, PEF: Profundidad efectiva, EG: Esqueleto grueso, a: arena, A: Arcilla, CO: Carbono orgánico, PSB: Porcentaje de saturación con bases. MAE: Error medio absoluto, RMSE: Error de la raíz de la media cuadrática, AC: Coeficiente de concordancia, ME: error medio, S: Desviación estándar de los datos observados; n= 33 perfiles de suelo.

El índice de concordancia entre los valores estimados de las propiedades de los suelos y los valores observados también expresan relaciones aceptables. La mayoría de las propiedades están por encima del 73% de acuerdos, a excepción de %A, PEF y Esp A, las cuales se caracterizan por expresar una alta variabilidad de corto alcance (Valera, 2018). Al considerar la naturaleza de las propiedades edáficas, el grado de concordancia es de 61% para el conjunto de variables morfológicas, 74% para las variables físicas, y 85% para las químicas. Esto significa que, mediante la información de los datos edáficos y auxiliares obtenidos por regresión lineal múltiple, y la correlación espacial de los residuos, se puede obtener una mejor estimación que utilizando solamente el valor de las observaciones como predicción.

Los resultados mostraron que los modelos RK en el conjunto de validación fueron ligeramente superiores a los obtenidos por Zhu et al (2022), los cuales indicaron explicaciones de la variación del carbono orgánico del suelo entre 63 y 72%. De igual manera el método ha sido capaz de mostrar una mayor fortaleza ante el kriging ordinario, como lo han señalado Gia et al (2019) y Li et al (2023), lo cual es aceptable cuando existen áreas heterogéneas.

Descripción de los modelos RK

El mapa final de cada propiedad del suelo muestra un patrón de distribución espacial determinado por los modelos de regresión correspondientes, indicados en el Cuadro 1. Por ejemplo, los mapas de espesor del horizonte A y del *solum* y el de profundidad efectiva muestran un incremento de la magnitud de estos atributos en dirección Oeste-Este (Figuras 4a, 4b y 4c). Pero en los mapas se destaca claramente que el horizonte A y el *solum* son más gruesos y los suelos son más profundos en los vallecitos intramontanos y las partes bajas de las laderas. Este comportamiento responde a la participación del índice topográfico de humedad (ITH) con coeficiente positivo y la pendiente del terreno con coeficiente negativo, en los modelos de regresión que predicen los valores de estas propiedades del suelo. En otro ejemplo, el mapa de % EG muestra que es menor en las vegas de los valles intramontanos donde el ITH es elevado, y el efecto contrario ocurre en general en los suelos de laderas de montañas altas (Figura 4d).

Los contenidos de arena en la superficie del terreno están directamente relacionados con la posición relativa, cuyos valores son menores en las zonas bajas donde existen acumulaciones de materiales finos que han sido removidos de la superficie del suelo en las zonas altas (Figura 4e). Además, la orientación de la pendiente guarda cierta relación con los materiales arenosos, ya que existe la tendencia de que los valores sean mayores cuando las laderas están orientadas hacia el N y NO, y un efecto contrario hacia el S y SE. También el % arcilla muestra que los valores de esta variable tienden a ser mayores en los extremos oriental y occidental, y menores hacia la zona central del área de estudio (Figura 4f). Esto refleja el efecto de los coeficientes positivo del NDVI y negativo de la pendiente del terreno en el modelo de regresión respectivo.

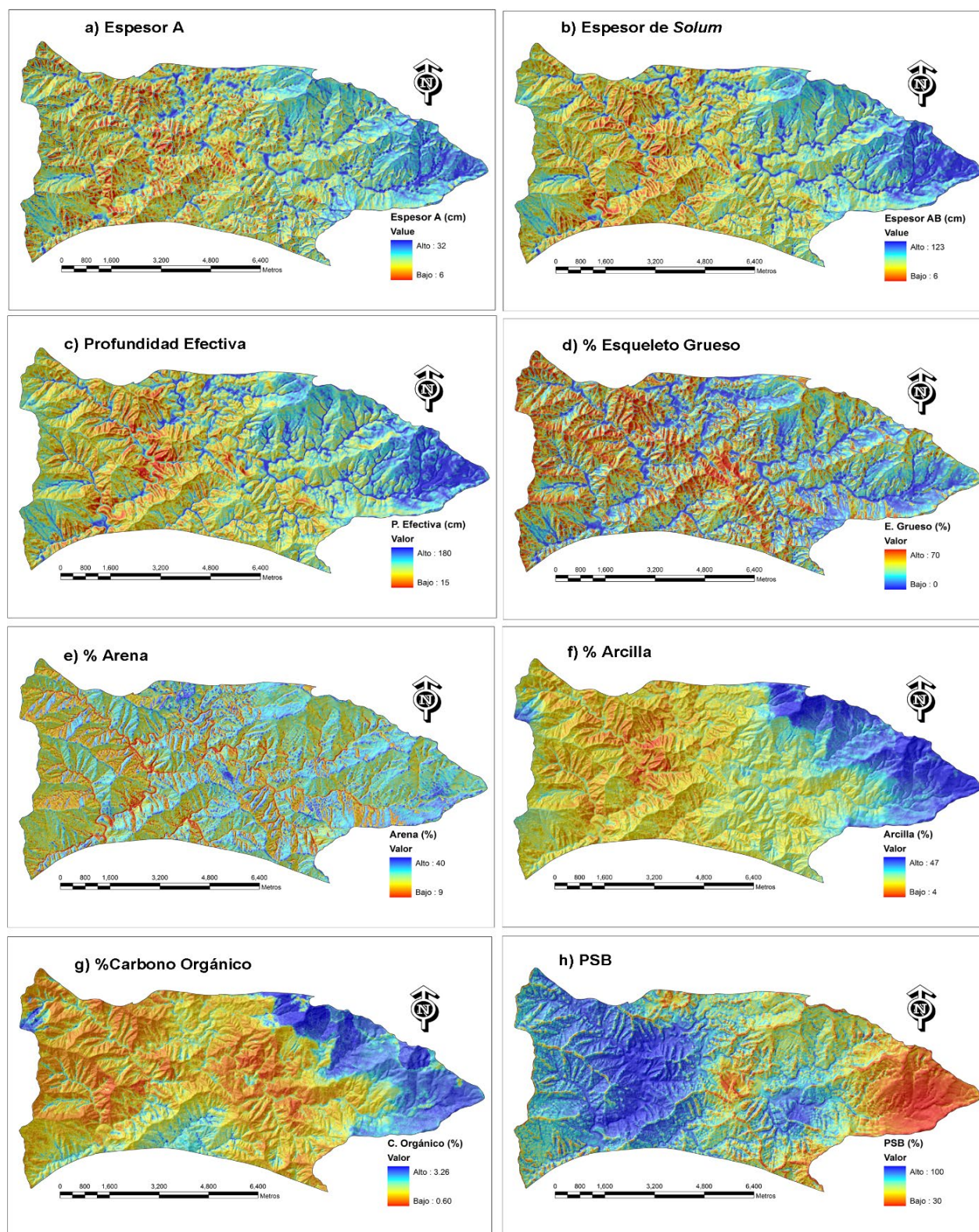


Figura 4. Modelo digital de propiedades del suelo en un sector de Caramacate, cuenca alta del río Guárico.

Los mapas de las propiedades químicas indican que el %CO es mayor en las áreas donde existen coberturas boscosas (mayor NDVI), ya sea en las zonas ribereñas o en las montañas altas hacia el Este de la cuenca (Figura 4g). En este sector, el microclima local caracterizado por una mayor precipitación, influye en los bajos contenidos bases cambiables, pH bajos con una reacción de moderada a fuertemente ácida, tal como se observa en la figura anterior (Figura 3c). Aunque la precipitación es el elemento climático que más afecta el lavado de bases, el PSB presenta mayores concentraciones en las zonas con formas de terreno convergentes con curvatura vertical de alta concavidad (Figura 3h).

CONCLUSIONES

La aplicación del enfoque RK en la cuenca del río Caramacate permitió la obtención de modelos de propiedades del suelo de diferente naturaleza, demostrando la utilidad de la técnica de cartografía digital de suelos en áreas de alta complejidad como los paisajes de montaña.

El enfoque combinado de ecuaciones de regresión lineal múltiple con técnicas de interpolación permitió la caracterización y diferenciación de la heterogeneidad edáfica causada por la influencia del relieve, el uso de la tierra y la susceptibilidad a los movimientos en masa, en los paisajes de montaña dominantes en el área de estudio.

Los modelos de predicción espacial de las propiedades morfológicas, físicas y químicas del suelo, fueron capaces de capturar la variación local de la información edáfica, así como también la variación continua de dichas propiedades, lo cual posibilitó el análisis de las relaciones suelo-paisaje.

La confiabilidad de los modelos indicó una relación directa de las propiedades investigadas con la naturaleza de las variables, ya que el grado de concordancia fue de 61% para las propiedades morfológicas, 74% para las variables físicas, y 85% para las químicas.

AGRADECIMIENTOS

Los autores agradecen al Fondo Nacional de Ciencia, Tecnología e Innovación (FONACIT) por el financiamiento parcial de la investigación, a través del programa “Misión Ciencia”, al Laboratorio de Agrología del Instituto de Edafología de la Facultad de Agronomía de la Universidad Central de Venezuela, y al Centro de Investigación y Extensión en Suelos y Aguas de la Universidad Nacional Experimental “Rómulo Gallegos” (CIESA-UNERG).

REFERENCIAS BIBLIOGRAFICAS

- Bishop, T.F.A. y McBratney, A.B. (2001). A comparison of prediction methods for the creation of field-extent soil property maps. *Geoderma*. Vol. 103 (1-2): 149-160.
- Burrough, P.A. y McDonnell, R.A.. (1998). *Principles of geographical information systems*. Oxford, Oxford University Press New York. 333 p.
- Cambardella, C., Moorman, T., Novak, J., Parkin, T., Karlen, D., Turco, R. y Konopka, E. (1994). Field scale variability of soil properties in Central Iowa Soils. *Soil Sci. Soc. Am. J.*

Vol. 58:1501-1511.

- Cook, D. (1977). Influential Observations in Linear Regression. *Journal of the American Statistical Association*. 74 (365): 169-174.
- Cressie, N. 1993. *Statistics for Spatial Data*, revised ed. Wiley, New York, 900 p.
- Di Rienzo J.A., Casanoves, F., Balzarini, M.G., González, L., Tablada M. y Robledo, C.W. (2011). InfoStat versión 2011. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. URL <http://www.infostat.com.ar>.
- Gia, T; Kappas M; Van Huynh C. y Hoang, L. (2019). Application of Ordinary Kriging and Regression Kriging Method for Soil Properties Mapping in Hilly Region of Central Vietnam. *ISPRS International Journal of Geo-Information*, 8(3):147. <https://doi.org/10.3390/ijgi8030147>
- Hengl, T., Heuvelink, G.B.M y Rossiter, D.G. (2007). About regression-kriging: from theory to interpretation of results. *Computers & Geosciences*. Vol. 33(10): 1301-1315.
- Hengl, T., Heuvelink, G.B.M. y Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*. Vol. 120: 75-93.
- Jenny, H. (1941). *Factors of Soil Formation: A System of Quantitative Pedology*. New York: McGraw-Hill. 281 p. On line: <http://www.soilandhealth.org/01aglibrary/010159.Jenny.pdf>
- Ji, L. y Gallo, K. (2006). An Agreement Coefficient for Image Comparison. *Photogrammetric Engineering & Remote Sensing*. Vol. 72 (7): 823-833.
- Kravchenko, A.N. y Robertson, G.P. (2007). Can topographical and yield data substantially improve total soil carbon mapping by regression kriging? *Agronomy Journal* 99: 12-17.
- Li, Y., Wang, X., Chen, Y. Gong, X., Yao, C., Cao, W. y Lian, J. (2023). Application of predictor variables to support regression kriging for the spatial distribution of soil organic carbon stocks in native temperate grasslands. *J Soils Sediments* 23, 700–717 (2023). <https://doi.org/10.1007/s11368-022-03370-1>
- Mahibbur, R.M. y Govindarajulu, Z. (1997). A modification of the test of Shapiro and Wilks for normality. *Journal of Applied Statistics*. Vol. 24(2): 219-235.
- Marchetti, A.; C. Piccini, R. Francaviglia, S. Santucci y Chiuchiarelli, I. (2010). Estimating Soil Organic Matter Content by Regression Kriging. In: *Digital Soil Mapping. Bridging Research, Environmental Application and Operation*. (Ed. A.B. McBratney and A.E. Hartemung). Chapter 20. New York. 241 p. ISBN 978-90-481-8862-8.
- Matheron, G. (1969). Part 1 of *Cahiers du Centre de morphologie mathématique de Fontainebleau. Le krigeage universel*. École nationale supérieure des mines de Paris.
- McBratney, A.B., Mendonça, M.L. y Minasny, B. 2003. On digital soil mapping. *Geoderma*. Vol. 117: 3-52.
- Minasny, B. y A.B. McBratney. (2007). Spatial prediction of soil properties using EBLUP with the Matérn covariance function. *Geoderma*. Vol. 140 (4): 324-336.
- Minasny, B., Mcbratney, A.B. y Whelan, M. (2002). VESPER. Version 1.6. Australian Centre for Precision Agriculture. McMilan Building. The University of Sidney, NSW 2006.
- Moore, I.D., Gessler, P.E., Nielsen, G.A. y Peterson, G.A. (1993). Soils attribute prediction using terrain analysis. *Soil Sci. Soc. Am. J.* Vol. 57: 443-452.
- Moore, I.D., Grayson, R.B. y Ladson, A.R. (1991). Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Process*. Vol. 5: 3-30.
- Odeh, I.O.A., McBratney, A.B. y Chittleborough, D.J. (1994). Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma*. Vol.

63: 197-214.

- Olmedo, G.F., Angelini, M.E., Vallone, R.C. y Moretti, L.M. (2012). Estimación de variables edáficas en el oasis productivo de Tupungato, Mendoza. XIX Congreso Latinoamericano de la Ciencia del Suelo. XXIII Congreso Argentino de la Ciencia del Suelo. Mar del Plata, Argentina - 16 al 20 de abril de 2012. 6p.
- Penížek, V. y Borůvka, L. (2006). Soil depth prediction supported by primary terrain attributes: a comparison of methods. *Plant Soil Environ.* Vol. 52 (9): 424-430.
- Pineda, M.C., Elizalde, G. y Vilorio, J. (2011a). Determinación de áreas susceptibles a deslizamientos en un sector de la Cordillera de la Costa Central de Venezuela. *Interciencia.* Vol. 36(5): 370-377.
- Pineda, M.C., Elizalde, G. y Vilorio, J. (2011b). Relación suelo-paisaje en un sector de la cuenca del río Caramacate, Aragua, Venezuela. *Revista de la Facultad de Agronomía.* Vol. 37(1): 27-37.
- Rojo, J.M. (2007). Regresión lineal múltiple. Instituto de Economía y Geografía. Laboratorio de estadística. Madrid. 32 p.
- Sun, W., Minasny, B. y McBratney, A.B. (2012). Analysis and prediction of soil properties using local regression-kriging. *Geoderma.* Vol. 171-172: 23-30.
- Tarboton, D.G., Bras, R.L. y Rodriguez-iturbe, I. (1991). On the extraction of channel networks from digital elevation data. *Hydrological Processes.* Vol. 5: 81-100.
- Tussel, F. 2011. Análisis de Regresión. Introducción Teórica y Práctica basada en R. Bilbao. 258p.
- Upchurch, D. y Edmonds, W.J. (1991). Statistical procedures for specific objectives. In: *Spatial variabilities of soils and landforms.* SSSA Special publication No. 28. 2ª Ed. SSSA. Madison. pp. 49-71.
- Urbani, F. y Rodríguez, J.A. (2004). Atlas geológico de la Cordillera de la Costa, Venezuela. Mapas a escala 1:25.000. Versión Digital. Edic. Fundación Geos, UCV. Caracas.
- Valera, A. (2015). Inventario de suelos y paisajes con apoyo de técnicas de cartografía digital en áreas montañosas. Caso Cuenca del Río Caramacate, Estado Aragua. Tesis de doctorado en Ciencias del Suelo. Universidad Central de Venezuela. Postgrado en Ciencias del Suelo. Maracay, Estado Aragua, Venezuela. 263 p. DOI: 10.13140/RG.2.1.1714.3920
- Valera, A. (2018). Geomorfometría y Edafometría. Cartografía Digital de Paisajes y Suelos con Técnicas de Inteligencia Artificial. Editorial Académica Española. Mauritius. ISBN: 978-620-2-12102-6. 317p.
- Verbrugge, L. (2006). Depth of Soil in the Goss-Gasconade-Rock outcrop complex in Callaway County, Missouri Using the Soil Land Inference Model (SoLIM). A thesis presented to the department of geology and geography in candidacy for the degree of Master of Science. Northwest Missouri State University. 76 p.
- Webster, R. y Oliver, M.A. (1990). *Statistical Methods in Soil and Land Resource Survey.* Oxford University Press. Oxford, RU. 316p.
- Willmott, C.J. (1981). On the validation of models. *Physical Geography.* Vol. 2: 184-194.
- Willmott, C.J. (1984). On the evaluation of model performance in physical geography. In: *Spatial Statistics and Models,* 443-460. Gaile, G.L. y C. J. Willmott, (eds). Norwell, Mass.: D. Reidel.
- Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., J. O'Donnell, J. y Rowe, C.M. (1985). Statistics for the evaluation and comparison of

- models. Journal of Geophysical Research. Vol. 90(5): 8995-9005.
- Wilson, J.P. y Gallant, J.C. (2000). Digital Terrain analysis. Principles and applications. New York, NY. John Wiley. 127 p.
- Zevenbergen, L.W. y Thorne, C.R. (1987). Quantitative analysis of land surface topography. Earth Surface Processes and Landforms. Vol. 12: 47-56.
- Zhu C, Wei Y., Zhu F., Lu W., Fang, Z., Li Z, y Pan, J. (2022). Digital Mapping of Soil Organic Carbon Based on Machine Learning and Regression Kriging. Sensors, 22(22):8997. <https://doi.org/10.3390/s22228997>
- Zhu, Q. y Lin, H.S. (2010). Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. Pedosphere. Vol. 20(5): 594-606.