



Terra. Nueva Etapa  
ISSN: 1012-7089  
ISSN: 2542-3266  
vidal.saezsaez@gmail.com  
Universidad Central de Venezuela  
Venezuela

# AGRUPAMIENTO DE SUELOS CON REDES NEURONALES DE MAPAS AUTOORGANIZADOS EN PAISAJES DE MONTAÑA EN LA REGIÓN CENTRO NORTE DE VENEZUELA

**Valera, Ángel R.; Vilorio, Jesús A.; Pineda, María C.; Ferrer, Jairo C.**  
AGRUPAMIENTO DE SUELOS CON REDES NEURONALES DE MAPAS AUTOORGANIZADOS EN PAISAJES DE MONTAÑA EN LA REGIÓN CENTRO NORTE DE VENEZUELA  
Terra. Nueva Etapa, vol. XXXV, núm. 58, 2019  
Universidad Central de Venezuela, Venezuela  
**Disponible en:** <http://www.redalyc.org/articulo.oa?id=72164777005>

Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional.

# AGRUPAMIENTO DE SUELOS CON REDES NEURONALES DE MAPAS AUTOORGANIZADOS EN PAISAJES DE MONTAÑA EN LA REGIÓN CENTRO NORTE DE VENEZUELA

Soil clustering by self-organizing maps neural networks in mountain landscapes of the north central region of Venezuela

*Ángel R. Valera*  
Universidad Rómulo Gallegos, Venezuela  
angelvalera@unerg.edu.ve

Redalyc: <http://www.redalyc.org/articulo.oa?id=72164777005>

*Jesús A. Viloría*  
Universidad Central de Venezuela, Venezuela  
jesus.viloria@ucv.ve.

*María C. Pineda*  
Universidad Central de Venezuela, Venezuela  
maria.c.pineda@ucv.ve

*Jairo C. Ferrer*  
Universidad Rómulo Gallegos, Venezuela  
ferrerjairo9@gmail.com

## RESUMEN:

La agrupación de suelos en clases contribuye a reducir la complejidad de la información, recordar las principales características de los grupos de suelos y comprender las relaciones entre estos grupos. En este estudio se aplicó la red neuronal artificial “mapas autoorganizados (Self-Organizing Maps) de Kohonen” para identificar clases locales de suelos en un sector de la cuenca alta del río Guárico, en la región Centro Norte de Venezuela. Los datos consistieron en variables medidas en 108 perfiles representativos de la variabilidad de suelos observada en el área de estudio. Las variables de entrada a la red neuronal incluyeron perfil de suelo, espesor del horizonte A, espesor del horizonte B, espesor del *solum*, profundidad efectiva, contenido de elementos gruesos, contenido de arcilla y arena, pH en agua, capacidad de intercambio catiónico, porcentaje de saturación con bases, calcio cambiante y carbono orgánico del suelo. La calidad de la clasificación obtenida se evaluó por medio de un análisis discriminante canónico para todas las variables de suelo consideradas y un análisis de varianza para cada variable por separado. El análisis discriminante canónico reveló una mayor variación entre las clases que la variación dentro de ellas y el análisis de varianza demostró que la clasificación permite predecir los valores de las variables consideradas, con excepción del espesor del horizonte A y de los porcentajes de arena, arcilla y carbono orgánico. La relación entre las clases de suelo obtenidas y las geoformas identificadas en un estudio previo se comprobó por medio de un coeficiente de contingencia. Finalmente se asignó cada clase local de suelo a una familia de la Taxonomía de Suelos del Departamento de Agricultura de los EUA.

**PALABRAS CLAVE:** Redes neuronales artificiales, Clasificación, Self-Organizing Maps (SOM), Suelos.

## ABSTRACT:

Grouping soils into classes helps to reduce the complexity of the information, remember the main characteristics of the soil groups, and understand the relationships between these groups. This study applied the artificial neural network “Kohonen Self-Organizing Maps” to identify local soil classes in a sector of the upper Guárico River basin, in the North Central region of Venezuela. The data consisted of soil variables measured in 108 soil profiles, representative of the soil variability observed in the study area. Input variables to the neural network included soil profile, thickness of the A horizon, thickness of the B horizon, thickness of solum, effective depth, coarse fragment content, clay and sand percent, pH in water, cation exchange capacity, percent base saturation, exchangeable calcium and organic soil organic carbon. The usefulness of the obtained soil classes was evaluated by means of a canonical discriminant analysis for the whole set of soil variables and an analysis of variance for each variable separately. The canonical discriminant analysis revealed a greater between-class variation than the within-class variation. The analysis of variance showed that the classification is useful to predict the values of the variables considered, except for thickness of the A horizon, sand percent, clay percent and soil organic carbon. The relationship between the obtained soil classes and the geoforms identified in

a previous study was verified by means of a contingency coefficient. Finally, each local soil class was assigned to a family of the Soil Taxonomy.

KEYWORDS: Artificial neural networks, Classification, SOM, Soil.

## INTRODUCCIÓN

La clasificación de suelos puede reducir eficazmente la complejidad de la información y contribuir al entendimiento de las principales características de los grupos de suelos. Aunque en este proceso siempre se pierde alguna información, también es importante considerar que proporciona un medio conveniente para la identificación de grupos naturales de individuos que tienen propiedades comunes, y para la transferencia sistemática de dicha información. El proceso de agrupamiento contribuye a entender la complejidad mediante la búsqueda de propiedades o comportamientos similares. En la ciencia del suelo se utilizan las clasificaciones jerárquicas, donde los objetos individuales como los perfiles de suelo se agrupan en clases (por ejemplo, las series de suelo). Estas clases entonces constituyen otros objetos, que a su vez pueden ser clasificados a un nivel jerárquico superior, como, por ejemplo, en grupos de suelos de referencia (Rossiter, 2001). La mayoría de los sistemas de clasificación de suelos subdividen al suelo en clases separadas entre sí por límites claros y precisos. De esa manera, el suelo es concebido como un continuo conformado por un conjunto de individuos discretos (Morales y Vilorio, 2006). En la actualidad, muchos estudios utilizan redes neuronales artificiales para solventar problemas de datos de alta complejidad. Entre estas aplicaciones, destacan las redes de mapas auto-organizados (SOM, *Self-Organizing Map*) desarrolladas por Kohonen (Kohonen, 1982). Los SOM son apropiados para la identificación de grupos comunes, ya que son capaces de reducir la dimensionalidad de datos complejos e identificar patrones, permitiendo una fácil visualización, el análisis y la comprensión de clases o categorías. La red autoorganizada es un método de aprendizaje no supervisado para la clasificación de datos multidimensionales, cuya principal característica es la creación de mapas bidimensionales. En estos mapas los contenidos que estén relacionados aparecerán juntos y, cuanto más similares sean los datos más próximos estarán en el mapa. Los SOM han sido de gran utilidad en la clasificación de imágenes de satélite (Gonçalves *et al.*, 2011; Petrov *et al.*, 2013), pero también se han aplicado en la visualización de datos agrícolas (Ruß *et al.* 2009), en la predicción de órdenes de suelo (Sarmiento *et al.*, 2012), en la categorización de la calidad del agua, suelo y sedimentos en regiones petroquímicas como indicadores de fuentes de contaminación ambiental (Olawoyin *et al.*, 2013), en el análisis de propiedades físicas y químicas del suelo (Standkiewicz y Kosiba, 2009; Merdun, 2011; Dhar y Cherkassky, 2011), en el análisis morfométrico del paisaje (Ehsany y Quiel, 2009a) y en la evaluación de la calidad de variedades de caña de azúcar (Ibrahim *et al.*, 2015).

En el campo edafológico, este tipo de análisis permite: i) la agrupación de un conjunto de datos de suelo en función de características similares, ii) la producción de categorías con el centro de las clases, para cada unidad de suelo generada, iii) el cálculo de medidas de similitud relativas entre cada grupo de suelos, y iv) la distribución de las clases de suelo con valores y variabilidad similar de las propiedades consideradas en el agrupamiento (Ehsani y Quiel, 2009a). La red auto-organizada debe descubrir rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, e incorporarlos a su estructura interna de conexiones, por lo que las neuronas deben auto-organizarse en función de los estímulos (datos) procedentes del exterior (Ehsany y Quiel, 2009a). El objetivo de este aprendizaje es categorizar los datos que se introducen en la red. Se clasifican valores similares en la misma categoría y, por tanto, deben activar la misma neurona de salida. Las clases o categorías son creadas por la propia red, debido a que se trata de un aprendizaje no supervisado a través de las similitudes entre los datos de entrada.

La cuenca alta del río Guárico, aunque abastece el 60% del agua que consume el área metropolitana de Caracas, está sometida a intensos procesos de degradación ambiental y carece de información de suelos para apoyar la implantación de planes de manejo. Los procesos morfodinámicos en áreas montañosas tienden a

generar patrones complejos de variación espacial del suelo. Un sector de esta cuenca ha sido utilizado como área de investigación sobre cómo modelar estos patrones en forma útil para el diseño y aplicación de planes de manejo de la cuenca. En particular, Vilorio *et al.* (2016) modelaron el paisaje del área piloto por medio de una red neuronal borrosa aplicada a parámetros derivados de un modelo digital de elevación y una imagen satelital. Asimismo, Ferrer *et al.* (2018) aplicaron la teoría de conjuntos borrosos para agrupar perfiles de suelos de esa área de estudio.

La presente investigación explora la aplicabilidad del algoritmo de mapas autoorganizados de Kohonen para agrupar en clases de suelos los datos medidos en puntos de muestreo, en un sector de la cuenca alta del río Guárico, en el Estado Aragua de Venezuela. Esta exploración comprende, además, la evaluación de la capacidad predictiva de las clases obtenidas y del grado de asociación de estas con clases taxonómicas de suelos y unidades geomorfológicas identificadas en el área de estudio.

## MATERIALES Y MÉTODOS

### Descripción del área de estudio

La investigación se realizó en un sector de la cuenca alta del río Guárico, específicamente en la cuenca del río Caramacate. Está ubicada entre los municipios Santos Michelena y San Sebastián de los Reyes, 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 6760 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, y por metalavas máficas de la formación El Carmen (Pineda *et al.*, 2011a). La zona presenta una altitud comprendida entre 334 a 1405 msnm., con una precipitación media anual que oscila entre 1100 y 1400 mm y la temperatura media anual varía entre 22 y 26 °C. La vegetación herbácea ocupa más del 50% del sector, como producto del uso ganadero en condiciones extensivas, acompañada de sobrepastoreo y quemadas anuales de vegetación. Los suelos en su mayoría son Entisoles, Inceptisoles y Alfisoles, cuya variabilidad se ha incrementado por el uso de la tierra basado en ganadería extensiva y la incidencia de movimientos en masa (Pineda *et al.*, 2011b). Esta situación ha llevado a que la cuenca esté sometida a intensos procesos de degradación, por lo que es necesaria la organización de la información de suelos para apoyar la implantación de planes de manejo.

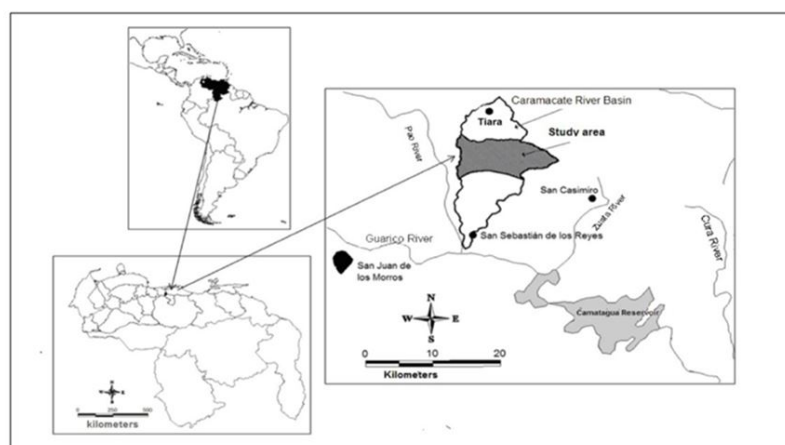


Figura 1. Ubicación relativa del área de estudio dentro de la cuenca del río Guárico estado Aragua, Venezuela.

### Datos de entrada y propiedades del suelo utilizadas en el agrupamiento

El conjunto de datos en la zona de estudio consistió en 108 perfiles de suelo ubicados en diferentes posiciones geomorfológicas de laderas de paisajes de montaña (Pineda *et al.*, 2011b; Valera, 2015). Los sitios de muestreo están conformados por observaciones de calicatas, cortes de vialidad, y sondeos con barreno agrológico. La densidad de muestreo fue equivalente a 1,5 observaciones/km., cumpliendo con las consideraciones y exigencias de los levantamientos de suelos a la escala 1:50.000. Los levantamientos, usualmente, consideran una observación en campo por 1-4 cm. de mapa o 0,5 observaciones/cm. de mapa, lo que significa que para la escala de trabajo se requiere una observación por cada 25-100 ha, o un promedio de una observación por cada 50 ha (1 observación/0,5 km.).

El conjunto de datos derivados de cada sitio de muestreo incluyen 25 propiedades edafológicas: i) variables físicas relacionadas con la granulometría (contenido de arena, limo, arcilla, esqueleto grueso), con la retención de humedad (agua retenida a capacidad de campo) e índices estructurales (conductividad hidráulica, densidad aparente, porosidad total); ii) variables químicas relacionadas con la disponibilidad y retención de nutrientes (capacidad de intercambio catiónico, Calcio, Magnesio, Potasio y Sodio intercambiables, relación Ca/Mg, porcentaje de saturación con bases, reacción del suelo, acidez intercambiable, fósforo disponible, carbono orgánico, conductividad eléctrica); iii) variables morfológicas relacionadas con la profundidad del suelo (espesor del horizonte A, secuencia de horizontes en el perfil, espesor del *solum*, espesor del horizonte B, profundidad efectiva del suelo) (Valera, 2015). En la Tabla 1, se señalan las propiedades consideradas y las metodologías empleadas en la determinación analítica de los suelos.

Los atributos utilizados en el agrupamiento neuronal fueron seleccionados por medio de un análisis de componentes principales (ACP). El ACP indicó que los primeros ocho componentes explican el 78% de la varianza que ocurre en los suelos de laderas de montaña, donde destacaron trece (13) variables edafológicas: perfil de suelo (secuencias de horizontes genéticos, 1: A/C, 2: A/AC/C, 3: A/Bw/C, 4: A/Bw1/Bw2/C, 5: A móllico/Bt1/C, 6: A ócrico/Bt1/C, 7: A/Bt1/Bt2/C); espesor del horizonte A (EspA, cm), espesor del horizonte B (EspB, cm), espesor del *solum*(EspAB, cm), profundidad efectiva (PEF, cm), contenido de arcilla (%A), arena (%a), pH en agua 1:1; capacidad de intercambio catiónico (CIC, cmol.kg<sup>-1</sup>), porcentaje de saturación con bases (PSB), calcio cambiante (Ca, cmol.kg<sup>-1</sup>), contenido de elementos gruesos (%EG) y carbono orgánico del suelo (%CO). Se realizó un análisis exploratorio para determinar la distribución de los datos, las medidas de tendencia central y dispersión, y la presencia de valores atípicos. Posteriormente, se efectuó la prueba de hipótesis de normalidad de la distribución de dichos datos, aunque estos supuestos no son requisitos indispensables para la aplicación del algoritmo SOM, cuya función es realizar los procesos de análisis, clasificación y predicción.

Tabla 1. Propiedades del suelo analizadas para cada uno de los sitios de muestreo.

Propiedad del Suelo	Símbolo	Unidad	Método Utilizado
Perfil de suelo (secuencia de horizontes)	Perfil	-	<i>Soil Survey Staff</i> (2014)
Espesor del horizonte A	EspA	cm	Barreno manual, cinta métrica.
Espesor del horizonte B	EspB	cm	
Espesor del <i>Solum</i>	EspAB	cm	
Profundidad efectiva	PEF	cm	
Arcilla	A	%	Método de la pipeta (Gee y Or, 2002)
Arena	a	%	Sedimentación, tamiz de 2 mm.
Limo	L	%	
Elementos gruesos	EG	%	Contenido en base a peso
Agua retenida a -33 kPa	CC	%	Mesa de tensión
Conductividad hidráulica saturada	Ks	cm h <sup>-1</sup>	Permeámetro de carga constante
Porosidad total	PT	%	Mesa de tensión
Densidad Aparente	Da	Mg m <sup>-3</sup>	Método del cilindro (Grossman y Reinsch, 2002)
Capacidad de intercambio catiónico	CIC	cmolk <sup>-1</sup>	NH <sub>4</sub> OAc, pH 7.0
Calcio intercambiable	Ca	cmolk <sup>-1</sup>	
Magnesio intercambiable	Mg	cmolk <sup>-1</sup>	
Sodio intercambiable	Na	cmolk <sup>-1</sup>	
Potasio intercambiable	K	cmolk <sup>-1</sup>	
Porcentaje de saturación con bases	PSB	%	
Relación Calcio/Magnesio	Ca/Mg	-	Potenciómetro, suelo/agua 1:1
pH en agua (1:1)	pH	-	
Acidez intercambiable	Acidez	cmol kg <sup>-1</sup>	Hopkins <i>et al.</i> (1903) modificado
Conductividad eléctrica (1:1)	CE	dS m <sup>-1</sup>	Conductímetro, suelo/agua 1:1
Fósforo disponible	P	Mg kg <sup>-1</sup>	Olsen modificado, NaHCO <sub>3</sub> 0,5 M, pH 8,5
Carbono orgánico	CO	%	Walkley y Black modificado (Heanes, 1984)

Fuente: elaboración propia

### Red neuronal artificial de mapas autoorganizados (SOM)

Los mapas autoorganizados consisten en una red neuronal artificial para la agrupación y visualización de información. Este algoritmo, desarrollado por Kohonen (1982), descubre rasgos comunes y semejanzas en los datos de entrada, y agrupa las observaciones similares de manera automática y no supervisada. La red está compuesta por dos capas de neuronas; una de entrada y otra de salida. La capa de entrada consiste en vectores de pesos que reciben y transmiten a la capa de salida la información procedente del exterior. La magnitud de cada vector corresponde al número de variables en el conjunto de datos de entrada y los pesos dependen de los valores de estas variables. La capa de salida está formada por  $m$  neuronas o unidades de mapa, que se encuentran organizadas, normalmente, en forma de una cuadrícula o mapa bidimensional (Ehsani y Quiel, 2009b).

La red SOM, como la mayoría de las redes neuronales artificiales, opera a través de dos etapas: entrenamiento y mapeo. En la etapa de entrenamiento se construye el mapa de salida en función de los ejemplos de entrada contenidos en un conjunto de datos de aprendizaje. Culminada la fase de entrenamiento, puede comenzar la fase de mapeo en la cual la red clasifica automáticamente a cada nuevo vector de entrada.

Se debe destacar que esta investigación se focaliza solo en la fase de entrenamiento de la red SOM, como una manera de explorar relaciones en los datos y patrones de similitud que permitan agrupar las observaciones en clases. Los datos de entrada consisten en un conjunto de características del suelo medidas en un número determinado de puntos de muestreo (observaciones). Dado que las escalas de medición son diferentes para distintas variables de suelo, es necesario estandarizar los datos antes de entrenar la red, de manera que el promedio de cada variable sea igual a cero y la varianza igual a uno. Cada unidad del mapa de salida de la red es una clase que agrupa a puntos de muestreo con características similares de suelo.

### Entrenamiento de la red SOM

Al inicio del entrenamiento, se asigna un vector modelo de pesos a cada unidad del mapa de salida, en un espacio de datos multidimensional. Los valores iniciales de los pesos ( $w$ ) son seleccionados aleatoriamente. Cada individuo en los datos de entrada está representado por un vector con  $n$  pesos  $x_k$ , determinados por

los valores de las variables de entrada. El aprendizaje de SOM consiste en un procedimiento competitivo que compara las distancias euclidianas entre los pesos ( $x$ ) de cada nuevo vector de entrada y los pesos ( $w$ ) de los vectores modelo de todas las unidades del mapa de salida. La unidad con la menor distancia euclidiana se selecciona como la mejor unidad de correspondencia (BMU, por *best matching unit*) o neurona ganadora (Kohonen, 1990; Kohonen, 1991; Kalteh *et al.*, 2008; Ehsani y Quiel, 2009a; Ehsani y Quiel, 2009b). El procedimiento de autoaprendizaje actualiza los pesos no solo de la neurona ganadora sino también de las neuronas vecinas, para preservar la topología de manera que los vectores que están cerca en el espacio de entrada (observaciones similares) se asignan a unidades que están cerca en el mapa de salida (clases semejantes). De esta forma, la neurona ganadora se convierte en el centro de una vecindad de actualización de las unidades del mapa y sus pesos asociados, de modo que cada vector de peso converja con el patrón de entrada. La velocidad con la cual los nodos ganadores convergen hacia los vectores de entrada se denomina velocidad de aprendizaje. A lo largo del aprendizaje, la velocidad de aprendizaje y el tamaño de la vecindad de actualización (el radio de actualización) disminuyen, de modo que los patrones generalizados iniciales se refinan progresivamente. El objetivo principal del modelo es permitir que los pesos de los vectores de las unidades de salida aprendan de lo que se les presenta a los vectores de entrada ( $x$ ). Este proceso es iterativo y se repite secuencialmente para cada nueva entrada hasta alcanzar el número de iteración predefinido para el aprendizaje. Después de la fase de aprendizaje, la red SOM consta de varios vectores, con vectores similares cerca y vectores diferentes más separados (Kohonen, 1990; Kohonen, 1991; Kalteh *et al.*, 2008; Ehsani y Quiel, 2009 b).

#### Sistema Clasificador NeuroXL

Los SOM han demostrado su eficacia en algunos problemas reales en los que incluyen tareas de clasificación, reducción de dimensiones y extracción de rasgos. Su utilidad más importante se relaciona con la clasificación de información o el agrupamiento de patrones por tipos o clases. En la presente investigación se empleó el programa *NeuroXL*® versión 4.0.6 (*Neuro XL Clusterizer*) (OLSOFTE, 2016), el cual se integró en *Microsoft Excel*®. El clasificador *NeuroXL*® utiliza un algoritmo de autoorganización integrado en hojas de cálculo. El programa funciona con varios parámetros de entrenamiento que gobiernan el proceso de operación de la clasificación, donde destacan los siguientes: i) *número de clases*; ii) *tasa de aprendizaje*, corresponde a un valor entre 0 y 1 que afecta el ritmo al cual la RNA inicia el aprendizaje; iii) *ciclos*, referidos al número de pasos completos de la RNA través del conjunto total de datos; iv) *peso inicial de la sinapsis*, valor con el que son inicializados los pesos sinápticos de cada neurona; v) *función de activación*, incluye cuatro (4) funciones alternativas: umbral, tangente hiperbólica (sigmoideal bipolar), sigmoideal logística con base cero, y sigmoideal logarítmica; vi) *algoritmo de combinación convexa*, utilizado para el entrenamiento de la red; vii) *relación de aprendizaje*, es el número de neuronas que afectará el aprendizaje durante la etapa de entrenamiento, el cual decrece a cero al finalizar el proceso de clasificación; y viii) *valor de escala*, esta opción permite escalar los valores de entrada en rangos dependientes de la función de activación empleada (e.g. 0-1, -1-1). En esta investigación, el entrenamiento de la red neuronal se realizó con una tasa inicial de aprendizaje de 0,5; 1.000 iteraciones, un peso inicial de 0,3 con una relación de aprendizaje de 3 neuronas por ciclo. La función de activación empleada fue la *tangente hiperbólica* o *sigmoideal bipolar*. Las funciones sigmoideales bipolares de tipo tangente hiperbólica son ampliamente utilizadas en tareas de clasificación de patrones (Quintín y Paz, 2007). Esta función tiene una variación de -1 a 1, y es probablemente la función de activación más empleada en la actualidad.

#### Número de clases neuronales SOM

Se desconoce de antemano cuál es el número más adecuado de clases para agrupar los puntos de muestreo disponibles en el área de estudio. Una aproximación a este número se obtuvo por medio de un procedimiento que consistió en agrupar los datos repetidamente en un número creciente de clases desde 2 hasta 18, por medio del algoritmo *k-means*. El procedimiento aplicado supone que si en los datos existen relaciones de

similitud que permitan agrupar las observaciones en clases, estas relaciones también serán detectadas por el algoritmo *k-means*.

Se calculó la suma de cuadrados intraclase de cada agrupamiento y se dibujó un diagrama de dispersión para observar la variación de la suma de cuadrados con relación al número de clases. Este diagrama (Figura 2) revela que la suma de cuadrados intraclase disminuye y, por ende, la homogeneidad interna de las clases aumenta a medida que crece el número de clases. El diagrama de dispersión también revela que la suma de cuadrados tiende a hacerse asintótica a partir de diez (10) clases. En consecuencia, se utilizaron diez unidades de salida de la red SOM, lo cual corresponde a un mapa bidimensional de 2 x 5 neuronas. Este número de clases es parecido al utilizado por Ferrer *et al.* (2018), quienes agruparon los perfiles de suelo del área de estudio en once (11) clases, por medio del algoritmo *fuzzy c-means*.

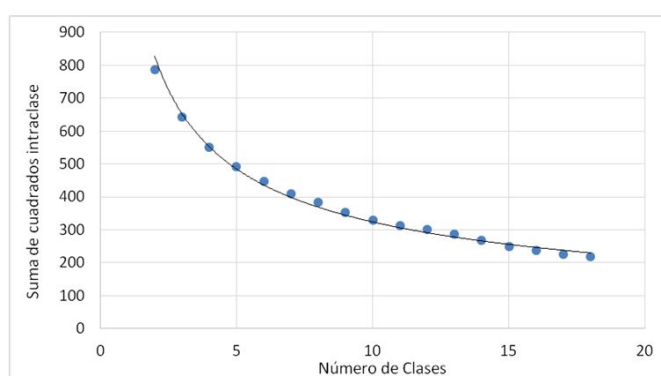


Figura 2. Variación de la suma de cuadrados intraclase con el número de clases de suelo en un sector de la Cuenca del río Caramacate.

### Evaluación de la capacidad predictiva de las clases neuronales SOM

Para evaluar la capacidad predictiva de las clases, se aplicaron métodos estadísticos multivariados y univariados. El método multivariados se realizó mediante el análisis discriminante canónico y el análisis de correlación canónica entre las variables edáficas y las clases locales de suelos. Los estadísticos descriptivos multivariados para la diferenciación entre clases neuronales fueron: i) *la M de Box* para probar la igualdad de las matrices de covarianza de las variables independientes entre los grupos que forman la variable dependiente, ii) el estadístico *Lambda de Wilk* ( $\lambda$ ) para evaluar si las funciones discriminantes canónicas contribuyen significativamente en la separación de las clases, iii) *Chi-cuadrado de Bartlett* ( $\chi^2$ ), el cual es el valor transformado del estadístico *Lambda*, y iv) *distancia de Mahalanobis* ( $D$ ) como criterio de selección de variables, con clasificaciones basadas en validaciones cruzadas.

El otro método utilizado para evaluar la capacidad predictiva de los agrupamientos fue el análisis de varianza para la clasificación de una vía, mediante el coeficiente de correlación intra-clase ( $\rho$ ) y el complemento de la varianza relativa ( $1-rv$ ) (Beckett y Burrough, 1971; Webster, 2008), con la finalidad de verificar el efecto de las propiedades edáficas en la diferenciación de las clases locales de suelos del sector estudiado dentro de la cuenca del río Guárico.

### Asociación entre clases neuronales de suelo y unidades de paisaje

Para evaluar la asociación entre las clases de suelo generadas a partir de sus propiedades y las unidades de paisaje se utilizó el estadístico *Chi-cuadrado de Pearson* ( $\chi^2$ ) y el coeficiente de contingencia (CC).

Para evaluar la asociación entre las clases de suelo neuronales y las unidades de paisaje se utilizó el estadístico *Chi-cuadrado de Pearson* ( $\chi^2$ ) en tablas de contingencia. También se utilizó el coeficiente de contingencia, el cual se emplea para evaluar la intensidad de la relación entre variables nominales. Este coeficiente es una medida normalizada para tablas mayores a 2x2, que oscila entre 0 para el caso de no asociación y el valor del coeficiente que nunca alcanza el valor 1. En el supuesto caso que el estadístico manifieste relación y/o



asociación entre las variables, es necesario valorar su significación estadística. Esta se realizó con la prueba  $\chi^2$  para contrastar la hipótesis nula;  $H_0$  = las variables X y Y son independientes. La base de cálculo de la Chi-cuadrado son las diferencias entre las frecuencias observadas y esperadas, y cuando la significación asociada a este estadístico es menor o igual a 0,05 se rechaza la hipótesis de independencia entre las clases de suelo y las unidades de paisaje, y se acepta que existe una relación de dependencia (Sánchez, 1989).

En tal sentido, se consideraron las clases morfométricas del terreno o geoformas obtenidas mediante un análisis geomorfométrico realizado en la zona de estudio, las cuales representan las unidades de relieve del paisaje de montaña a escala 1: 50.000 (Valera, 2015; Vilorio *et al.*, 2016). Estas clases digitales corresponden a la integración de variables ambientales derivadas de un modelo digital de elevación y una imagen de satélite multiespectral *Spot*, remuestreada a 15 m de resolución espacial.

## RESULTADOS Y DISCUSIÓN

### Clases de suelo neuronales y centroides de clase

Para las clases de suelo autoorganizadas, el entrenamiento de la red neuronal SOM indicó que 10 clases son suficientes para agrupar los perfiles de suelos de ladera de paisajes de montaña del área de estudio. El entrenamiento de la red SOM para las clases de suelo permitió la obtención de los centros de cada una de las clases SOM que se indican en la Tabla 2. En la tabla, el peso de las variables de la clasificación está dado por el promedio ponderado del valor de los atributos del suelo en cada clase.

**Tabla 2.** Centroides del modelo generado con 10 clases neuronales autoorganizadas de suelos.

Clase	EspA	EspB	EspAB	PEF	Perfil	%A	%a	%EG	CIC	Ca	PSB	pH	%CO
<b>A</b>	29	47	76	110	7	14,2	18,7	12	17	10	86	5,58	1,71
<b>B</b>	18	26	44	64	6	24,1	30,7	50	19	8	83	5,67	1,57
<b>C</b>	10	0	10	25	1	20,2	37,8	56	20	4	45	4,89	1,57
<b>D</b>	13	0	13	29	1	17,4	38,6	44	27	12	77	5,68	1,69
<b>E</b>	12	0	12	31	1	16,9	29,4	50	33	10	53	5,43	1,92
<b>F</b>	14	0	15	28	1	19,4	24,6	23	27	9	69	5,52	1,74
<b>G</b>	13	0	13	31	1	16,6	24,3	15	41	14	63	5,47	1,83
<b>H</b>	13	23	36	54	3	19,9	21,5	11	41	14	64	5,60	1,73
<b>I</b>	16	23	39	52	3	18,0	25,6	16	26	10	72	5,52	1,68
<b>J</b>	26	55	81	110	7	23,2	22,3	17	35	7	37	4,74	2,10

Fuente: elaboración propia

**EspA:** Espesor del horizonte A; **EspB:** Espesor del horizonte B; **EspAB:** Espesor del *Solum*; **PEF:** Profundidad efectiva; **Perfil:** Perfil de suelo; **%A:** Arcilla; **%a:** Arena; **%EG:** Elementos gruesos; **CIC:** Capacidad de intercambio catiónico; **Ca:** Calcio intercambiable; **PSB:** Porcentaje de saturación con bases; **pH:** pH en agua (1:1); **%CO:** Porcentaje de Carbono orgánico.

De la Tabla 2 mencionada se pueden deducir tres grandes categorías con base en los atributos morfológicos: i) clases de suelos muy profundos a profundos (A, B, J), ii) clases de suelos moderadamente profundos (H, I) y iii) clases de suelos superficiales (C, D, E, F, G). Todas las clases presentan diferencias entre sí, cuyos contrastes se complementan con el aporte de los centros de los atributos físicos y químicos. Los atributos físicos separan claramente las clases (C, D, E) de la clase G, la cual presenta el valor más bajo de %EG. La influencia de los atributos químicos permite contrastar las clases C y G por la influencia de la CIC, el pH y el PSB admite la diferenciación entre las clases A y J, y permite contrastar las diferencias entre A, C y J por la influencia del PSB y el %CO. Los resultados indican que en todas las clases obtenidas, se mantiene la estructura topológica de los datos de entrada a la red neuronal (Baçao *et al.*, 2005).

### Evaluación de la capacidad predictiva de las clases neuronales SOM

#### *Análisis multivariado discriminante canónico*

Los estadísticos multivariados obtenidos con la aplicación del ADC se resumen en la Tabla 3. Lo primero que se destaca es que la prueba  $F$  de Box y el estadístico  $\Lambda$ , permitieron rechazar la hipótesis de igualdad de varianzas-covarianzas ( $p < 0,05$ ), ya que las clases de suelo obtenidas son distintas unas de otras. Los resultados

del estadístico Lambda ( $\lambda$ ) relacionados con los autovalores, en las situaciones planteadas presentan valores muy cercanos a cero, lo cual indica que existen grandes diferencias entre las clases de suelo. De igual manera, los valores transformados de  $\lambda$  equivalentes a *c.* ( $p < 0,05$ ) indicaron que las clases comparadas tienen promedios diferentes en sus variables discriminantes, lo que permitió corroborar la existencia de diferencias entre ellas.

**Tabla 3.** Estadísticos obtenidos con el análisis discriminante canónico en el agrupamiento de redes neuronales autoorganizativas.

Parámetros estadísticos	Valor
M de Box (M)	341
Prueba F	1,53
$p < 0,05$	0,000
Valor Propio	16,2
% Varianza	70,3
Correlación Canónica	0,970
Lambda Wilks ( $\lambda$ )	0,002
$\chi^2$ de Barlett (B)	623
$p < 0,05$	0,000

Fuente: elaboración propia

Las primeras correlaciones canónicas del agrupamiento SOM presentan autovalores de 16,2 para las primeras funciones discriminantes canónicas, indicando que las variaciones entre las clases de suelo son dieciséis veces mayores que las variaciones internas de dichas clases. Es decir, la variabilidad entre las clases es superior a la variabilidad dentro de las clases. En cuanto a la valoración de la capacidad predictiva de las funciones discriminantes (EspAB, Perfil, %EG, CIC, Ca, PSB, pH), los resultados de la clasificación realizada indicaron que el 95,4% fue clasificado correctamente según los casos agrupados originales (Tabla 4). Además, el 88,9% fue clasificado correctamente según los casos agrupados corroborados mediante validación cruzada y solamente se encontró un 4,6% de error. La distancia de Mahalanobis (*D.*), basada en las funciones canónicas para los datos originales, arrojó valores superiores al 95% con bajos errores. La *D.*, basada en las observaciones para los datos validados mediante validación cruzada, reflejó resultados menores que las clasificaciones originales, cercanos al 89% para los perfiles de suelo.

**Tabla 4.** Resultados de la clasificación basada en el tamaño de las 10 clases de suelo neuronales.

Clases	Variables discriminantes	Clasificación (%) <sup>1</sup>	Error (%)	Validación Cruzada (%) <sup>2</sup>
10	EspAB, Perfil, %EG, CIC, Ca, PSB, pH.	95,4	4,6	88,9

Fuente: elaboración propia <sup>1</sup> Clasificados correctamente según los casos agrupados originales. <sup>2</sup> Clasificados correctamente según los casos agrupados validados mediante validación cruzada.

Aunque los resultados de discriminación de variables para la clasificación edafológica son aceptables, estos son ligeramente superiores a los valores obtenidos en investigaciones con la aplicación de la teoría de conjuntos difusos (Ferrer *et al.*, 2018).

#### *Análisis de correlación canónica de agrupaciones neuronales autoorganizadas*

En la Tabla 5 se indica que la primera correlación canónica es alta ( $r = 0,97$ ), al igual que las dos siguientes, cuyo conjunto explica el 96% de la interrelación entre las propiedades del suelo y las clases de suelo autoorganizadas. Las últimas tres correlaciones restantes son bajas y solamente explican menos del 3% de la variabilidad presente entre las variables edáficas y las clases de suelo. El valor propio de la primera correlación

es de 16,2 lo cual indicó que la variación entre las clases es dieciséis veces mayor que la variación interna de dichas clases.

**Tabla 5.** Correlación canónica entre las clases neuronales y las propiedades del suelo.

Función	Autovalor <sup>1</sup>	% de varianza	% acumulado	Correlación canónica (r)	R <sup>2</sup>
1	16,2	70,3	70,3	0,970	0,942
2	3,9	16,8	87,0	0,891	0,794
3	1,9	8,6	95,6	0,815	0,664
4	0,7	3,2	98,9	0,653	0,426
5	0,1	0,5	99,3	0,317	0,100
6	0,1	0,5	99,8	0,311	0,097
7	0,05	0,2	100,0	0,210	0,044

Fuente: elaboración propia <sup>1</sup> En el análisis se emplearon las 7 primeras funciones discriminantes canónicas.

Las correlaciones indicadas en la Tabla 6, señalan la evidente contribución de la secuencia de horizontes en el perfil y la profundidad de los suelos a la primera correlación canónica.

**Tabla 6.** Correlación entre las variables canónicas y las propiedades del suelo.

Atributos	Matriz de estructura						
	Variables Canónicas (Función Discriminante) <sup>1</sup>						
	1	2	3	4	5	6	7
Perfil	0,609 <sup>3</sup>	0,304	-0,046	0,149	-0,561	0,378	-0,234
PEF <sup>2</sup>	0,401 <sup>3</sup>	0,062	-0,181	0,078	0,252	-0,349	-0,055
%A <sup>2</sup>	-0,119 <sup>3</sup>	-0,037	-0,034	0,098	0,045	-0,098	-0,107
CIC	-0,011	-0,615 <sup>3</sup>	-0,480	0,533	0,193	0,143	-0,221
%CO <sup>2</sup>	-0,161	-0,234 <sup>3</sup>	-0,201	-0,113	-0,089	0,093	-0,053
PSB	0,099	-0,069	0,731 <sup>3</sup>	0,052	0,334	-0,026	0,580
pH	0,039	-0,246	0,699 <sup>3</sup>	0,345	0,295	0,033	-0,493
%EG	-0,274	0,612	0,000	0,683*	0,278	0,087	-0,001
Ca	0,073	-0,526	0,216	0,560 <sup>3</sup>	-0,121	-0,329	0,484
%a <sup>2</sup>	0,008	0,036	0,012	0,123	-0,230 <sup>3</sup>	-0,083	-0,062
EspAB	0,528	0,253	-0,260	-0,093	0,237	-0,713 <sup>3</sup>	-0,132
EspB <sup>2</sup>	0,522	0,158	-0,294	0,017	0,221	-0,550 <sup>3</sup>	-0,169
EspA <sup>2</sup>	0,112	0,214	0,008	-0,210	0,072	-0,420 <sup>3</sup>	0,039

Fuente: elaboración propia <sup>1</sup>Correlaciones intraclases combinadas entre las variables discriminantes y las funciones discriminantes canónicas tipificadas. <sup>2</sup>Esta(s) variable(s) no se emplea(n) en el análisis. <sup>3</sup>Mayor correlación absoluta entre cada variable y cualquier función discriminante.

La segunda función discriminante está negativamente correlacionada con la CIC, y latercera correlación canónica con el PSB y pH. Solamente siete atributos contribuyen a la discriminación entre las clases de suelo, los cuales están relacionados con el espesor del *solum*, secuencia de horizontes, contenido de esqueleto grueso, cationes intercambiables, reacción del suelo, granulometría y pH, tal como se señala en la Tabla 7.

**Tabla 7.** Coeficientes estandarizados de las funciones discriminantes canónicas de los datos de suelos.

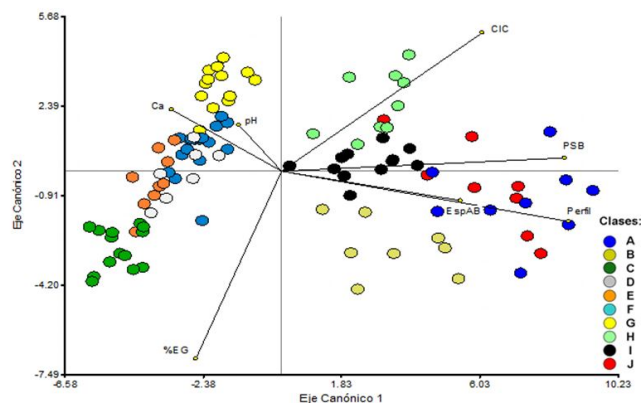
Atributo	Función Discriminante						
	1	2	3	4	5	6	7
EspAB	0,646 <sup>1</sup>	0,181	-0,206	-0,005	0,444	-0,617 <sup>1</sup>	-0,111
Perfil	0,855 <sup>1</sup>	0,181	0,034	0,255	-0,347	0,531	0,033
%EG	-0,312	0,674 <sup>1</sup>	-0,087	0,667 <sup>1</sup>	0,172	-0,005	0,085
CIC	0,598 <sup>1</sup>	-0,512 <sup>1</sup>	-0,656 <sup>1</sup>	0,222	0,936 <sup>1</sup>	0,757 <sup>1</sup>	0,139
Ca	-0,318	-0,243	0,240	0,607 <sup>1</sup>	-1,000 <sup>1</sup>	-0,859 <sup>1</sup>	0,205
PSB	0,884 <sup>1</sup>	-0,024	0,118	-0,178	1,103 <sup>1</sup>	0,836 <sup>1</sup>	0,840 <sup>1</sup>
pH	-0,119	-0,173	0,708 <sup>1</sup>	0,166	-0,059	-0,285	-0,886 <sup>1</sup>

Fuente: elaboración propia <sup>1</sup>Mayor correlación entre cada variable y cualquier función discriminante.

El diagrama de dispersión de los puntos de observación representados en la Figura 3 indica que las clases H e I se distribuyen en el primer cuadrante, en el que el PSB y la CIC tienen coeficientes de moderada a alta magnitud en el eje canónico 1. Las clases A, B y J ocupan el segundo cuadrante del diagrama; las dos primeras se identifican con suelos de horizonte A superior a 25 cm, con un gran espesor del *solum* y la mayor profundidad efectiva, lo que está relacionado con un mayor grado de desarrollo de los suelos.

Los perfiles de suelo de las clases C, D y E, ocupan el tercer cuadrante, con valores negativos de ambos ejes canónicos, definidos por la presencia de suelos de escaso grado de evolución, con horizonte A muy delgado, poca profundidad efectiva, y con altos contenidos de esqueleto grueso; además, estas clases incluyen suelos menos saturados en el complejo de cambio y con dominancia de clases texturales francas. Los suelos de la clase F y G se distribuyen en el cuarto cuadrante, los cuales también están asociados a suelos poco profundos, de escaso grado de evolución pedogenética, pero con menores contenidos de esqueleto grueso que las clases que ocupan la parte superior del eje canónico 2.

La distribución de los datos refleja un ligero solapamiento entre las clases A y J, y entre D, E y F. En las primeras, cuyos suelos muestran cierto grado de evolución, la mayoría de los atributos morfológicos son similares entre ambas clases, pero existen grandes diferencias en cuanto a CIC, PSB, %CO y %A. En las segundas, donde los suelos son de escaso desarrollo, las diferencias se deben a variaciones en el complejo de cambio y a la granulometría, originándose suelos de texturas francas a franco-limosas.



**Figura 3.** Diagrama de dispersión de los puntos de observación con relación a los ejes canónicos 1 y 2 de las clases locales autoorganizadas.

### *Análisis de varianza de una vía*

El análisis de varianza para la clasificación de una vía permitió la obtención del coeficiente de correlación intraclase (*ri*) de los atributos del suelo por efecto de las clases neuronales, así como también el complemento de la varianza relativa, los cuales se observan en la Tabla 8. Los resultados indican que en las situaciones descritas la varianza intraclases promedio presenta valores menores que la varianza total (Webster y Oliver,

1990), lo cual es un indicativo de que las clasificaciones realizadas son altamente meritorias para las variables consideradas.

**Tabla 8.** Componentes de la varianza de las clases neuronales, correlación intraclase y complemento de la varianza relativa para los atributos seleccionados.

Atributos	Media	S <sup>2</sup> <sub>B</sub>	S <sup>2</sup> <sub>w</sub>	S <sup>2</sup> <sub>T</sub>	ρ <sub>i</sub>	r <sub>v</sub>	1-r <sub>v</sub>
EspA	15,6	32,1	34,2	411,4	0,484	0,083	0,917
EspB	14,6	388,4	103,7	4359,8	0,789	0,024	0,976
EspAB	30,2	632,4	127,9	7017,3	0,832	0,018	0,982
PEF	49,3	930,8	332,8	10616,6	0,737	0,031	0,969
Perfil	2,6	4,2	0,5	46,2	0,900	0,010	0,990
%EG	29,1	320,1	107,6	3637,3	0,748	0,030	0,970
%a	27,3	36,4	58,8	507,0	0,383	0,116	0,884
%A	18,9	2,7	53,3	135,6	0,049	0,393	0,607
%CO	1,7	0,01	0,13	0,38	0,069	0,358	0,642
CIC	28,6	67,5	32,4	786,5	0,676	0,041	0,959
Ca	9,6	10,4	7,2	125,3	0,589	0,058	0,942
PSB	64,3	211,6	177,6	2618,0	0,544	0,068	0,932
pH	5,4	0,09	0,07	1,12	0,558	0,065	0,935

Fuente: elaboración propia S<sup>2</sup><sub>B</sub>: Varianza intraclases, S<sup>2</sup><sub>w</sub>: Varianza dentro de clases, S<sup>2</sup><sub>T</sub>: Varianza total, ρ<sub>i</sub>: Coeficiente de correlación intraclase, r<sub>v</sub>: Varianza relativa, 1-r<sub>v</sub>: Complemento de la varianza relativa.

Los valores de r<sub>i</sub> señalan que la clasificación explica más de las dos terceras partes de la varianza en los atributos morfológicos y químicos. Sin embargo, para los atributos %a, %A y %CO, la clasificación neuronal presenta una menor utilidad práctica en la separación de unidades de suelo. En cuanto a los resultados del complemento de la varianza relativa, la mayoría de los atributos edáficos explican más del 92% de la varianza de las clases de suelo locales, a excepción de la granulometría (%A, %a) y el contenido de carbono orgánico, cuyo aporte a las clasificaciones son ligeramente inferiores. En general, para el conjunto de perfiles de suelos de ladera evaluados, los valores r<sub>i</sub> .1-r<sub>v</sub> señalan que la clasificación empleada en este estudio es considerada efectiva en la separación de los conjuntos de datos edáficos. Esto indica que existe cierto grado de homogeneidad dentro de las clases de suelo, lo que permite aseverar que las predicciones que pueden ser realizadas a partir de estas, sean más precisas.

Los resultados de la aplicación del estadístico c. (238,1) y el CC (0,798), indican que existe una relación entre las clases de suelo con las unidades geomorfológicas derivadas del análisis morfométrico del terreno en investigaciones previas (Viloria *et al.*, 2016). Es evidente la asociación entre clases, ya que los valores del CC son equivalentes al 80%, lo cual tiene una significación asociada a un estadístico c. menor a 0,05 (0,000). Estos resultados permiten rechazar la hipótesis de independencia entre las clases de suelo neuronales y las geoformas, y admitir que ambos grupos están ampliamente relacionados, lo cual facilitó las posibilidades de predicción y representación cartográfica de dichas clases. Debido a la asociación entre clases y a las relaciones suelo-paisaje, se realizó la correlación entre los suelos de ladera y las unidades geomorfométricas. La representación cartográfica de la variación de las clases de suelo autoorganizadas se presenta en la Figura 4, donde además de las clases de suelos de ladera se incluyen dos clases adicionales que representan las geoformas compuestas por las crestas y vigas de montaña (clase K) y las vegas de los valles intramontanos (clase L) (Valera, 2015).

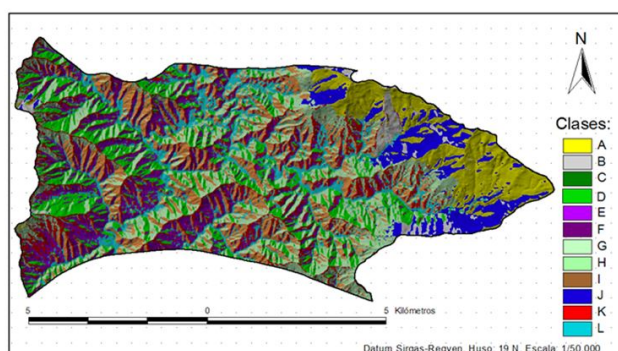


Figura 4. Representación espacial de las clases de suelo autoorganizadas en un sector de Caramacate, cuenca alta del río Guárico.

Estas últimas unidades de paisaje se caracterizan por presentar condiciones de mayor homogeneidad edáfica, razón por la cual no se incluyeron en la evaluación del presente estudio. La representación geoespacial de las clases de suelo, obtenidas con el método de agrupamiento neuronal, deja en evidencia la influencia aportada por la discretización de las unidades espaciales mínimas (celdas), cuya expresión geográfica está dada por los límites de las clases morfométricas del terreno, utilizadas como base para la representación espacial de las clases de suelo. Bajo este enfoque, la estructura de variación de las clases de suelo neuronales facilitó la comprensión de las relaciones suelo-paisaje en la zona de estudio, y permitió la correlación con las categorías taxonómicas a nivel de familia de clases de tamaño de partícula (Soil Survey Staff, 2014), tal como se indica en la Tabla 9.

Tabla 9. Correlación entre las clases de suelo autoorganizadas y las categorías taxonómicas

Clase	Familia(s) correlacionada(s) *	Superficie (%)
A	<i>TypicHaplustalfs, francafina</i>	9,9
B	<i>UlticHaplustalfs, franca fina</i>	2,2
C	<i>LithicUstorthents, franca fina</i>	9,5
D	<i>LithicUstorthents, franca gruesa</i>	4,9
E	<i>LithicUstorthents, franca fina</i>	4,8
F	<i>LithicUstorthents, franca fina</i>	8,4
G	<i>LithicUstorthents, franca gruesa</i>	7,7
H	<i>Lithic Haplustepts, franca fina; TypicHaplustepts, franca fina</i>	13,0
I	<i>TypicHaplustepts, limosafina; Lithic Haplustept, franca fina</i>	13,4
J	<i>Ultic Haplustalfs, fina</i>	6,4
k	<i>Lithic Haplustepts, francosafina; LithicUstorthents, franca fina</i>	10,0
L	<i>TypicHaplustepts, franca fina</i>	9,8

Fuente: elaboración propia \* Clasificación según la taxonomía de suelos (Soil Survey Staff, 2014).

En tal sentido, es posible ratificar que nivel de laderas, la poca estabilidad y la susceptibilidad a los movimientos en masa promueven la ocurrencia de suelos de escaso a moderado desarrollo pedogenético. Estas

clases ocupan una superficie equivalente al 71,5% del área evaluada en la zona. Asimismo, en la zona oriental de la cuenca del río Caramacate, la distribución de las clases de suelo se caracteriza por la dominancia de los subgrupos taxonómicos *Typic Haplustalfs-UlticHaplustalfs* (correspondientes al 19% del área estudiada), y en el resto del área dominan *Typic Haplustepts* entremezclados con *Lithic Ustorthents*. En las crestas también existen suelos de incipiente desarrollo con dominancia de *Lithic Haplustepts*, y en las vigas de laderas ocurren suelos menos profundos con predominio del gran grupo *Ustorthents*. En las vegas de los valles intramontanos ocurren procesos de acumulación localizada, creando una superficie estable con suficiente tiempo para que existan procesos de desarrollo de un endopedón *cámbico*, encontrándose suelos del subgrupo *Typic Haplustepts* que ocupan el 10% de la superficie.

## CONCLUSIONES

El número de clases neuronales seleccionadas en este estudio fue capaz de discriminar la variación existente en los suelos, lo que destaca la importancia de la aplicación del sistema de agrupamiento de suelos en áreas de alta complejidad para la obtención de clases homogéneas internamente.

Las clases de suelo autoorganizativas expresaron un alto grado de homogeneidad y fueron capaces de predecir las propiedades edáficas consideradas. Los grupos neuronales generados por el algoritmo SOM permitieron predecir con suficiente precisión las propiedades pertenecientes a las clases típicas de los suelos. Adicionalmente los parámetros estadísticos del análisis discriminante canónico determinaron que la variabilidad entre las clases de suelos es superior a la variabilidad dentro de las clases, y el complemento de la varianza relativa indicó que las propiedades morfológicas, químicas y físicas, son útiles al separar las clases de suelo.

La aplicación de las redes de mapas autoorganizadas constituye una alternativa para el agrupamiento de suelos en áreas de alta complejidad como los tipos de relieve de ladera de montañas, con alto impacto de intervención antrópica y con procesos erosivos que reflejan una gran variabilidad de los suelos.

## REFERENCIAS BIBLIOGRÁFICAS

- BAÇAO, F., LOBO, V., PAINHO, M. (2005). The self-organizing map, the Geo-SOM, and relevant variants for geosciences. *Computers & Geosciences*, 31: 155-163.
- BECKETT, PHT. y PA. BURROUGH. (1971). The relation between cost and utility in soil survey. IV. Comparison of the utilities of soil maps produced by different survey procedures, and to different scales. *J. Soil Sci.*, 22: 466-480.
- BEZDEK, JC., EHRLICH, R. y W. FULL. (1984). FCM: the fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10: 191-203.
- BURROUGH, PA., VAN GAANS, PFM., y RA. MACMILLAN. (2000). High-resolution landform classification using fuzzy k-means. *Fuzzy Sets and Systems*, 113: 37-52.
- DHAR, S. y V. CHERKASSKY. (2011). Application of SOM to analysis of Minnesota soil survey data. En: *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, 633-639.
- EHSANI, AH, y F. QUIEL. (2009a). A semi-automatic method for analysis of landscape elements using shuttle radar topography mission and landsat ETM+ data. *Computers & Geosciences*, 35: 373-389.
- EHSANI, AH, y F. QUIEL. (2009b). Self-organizing maps for multi-scale morphometric feature identification using shuttle radar topography mission data. *Geocarto International*, 24(5): 335-355.
- FERRER, J.C., VALERA, A.R., y B.Y. FLORES. (2018). Aplicación de la teoría de conjuntos borrosos en el agrupamiento de suelos de ladera en la cuenca del río Caramacate, estado Aragua. *Terra. Nueva Etapa*, 34(55), 97-117.

- GONÇALVES, ML., COSTA, JAF. y MLA. NETTO. (2011). Land-Cover Classification Using Self-Organizing Maps Clustered with Spectral and Spatial Information. En: *Self Organizing Maps - Applications and Novel Algorithm Design*. p. 299-322.
- GEE, GW. y D. OR. (2002). Particle-size analysis. En: Dane JH, Topp GC (Ed.) *Methods of soil analysis*. Part 4. SSSA Book series N° 5, SSSA, Madison, WI. p. 255-293.
- GROSSMAN, R.B. y T.G. REINSCH. (2003). Bulk density and linear extensibility. En: *Methods of soil analysis*. Part 1. 2nd ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI. p. 201-254.
- HEANES, DL. (1984). Determination of total organic-C in soils by an improved chromic acid digestion and spectrophotometric procedure. *Communications in Soil Science and Plant Analysis*, 15: 1191-1213.
- HOPKINS, C.; KNOX, W. y J. PETITT. (1903). A quantitative method of determining the acidity of soils. *U.S Bur. Chem. Bull.* 73: 114-126.
- IBRAHIM, OM., GAAFAR, AA., WALLI, AM. y M.M. TAWFIK. (2015). Assessing the performance and variability of some Sugar beet varieties using Self-organizing map artificial neural network and Cluster analysis. *International Journal of ChemTech Research*, 8(9):12-19.
- KALTEH, AM., HJORTH, P. y R. BERNDTSSON. (2008). Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application. *Environ. Model. Softw.*, 23: 835-845.
- KOHONEN, T. (1982). Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics*, 43 (1): 59-69
- KOHONEN, T. (1990). The self-organizing map. *Proceedings of the IEEE*. 78(9):1464-1480.
- KOHONEN, T. (1991). Self-organizing maps: Optimization approaches. En *Proceedings of the International Conference on Artificial Neural Networks*, Espoo, Finland, p. 981-990.
- KOHONEN, T., HYNINEN, J., KANGAS, J. y J. LAAKSONEN. (1996). SOM\_PAK: The Self-Organizing Map Program Package, Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science. 9 p.
- MERDUN, H. (2011). Self-organizing map artificial neural network application in multidimensional soil data analysis. *Neural Computing y Applications*, 20 (8): 1295-1303.
- MINASNY, B. y AB. MCBRATNEY. 2002. *FuzME v. 3.5*. Australian Centre for Precision Agriculture. The University of Sydney, Australia. Available from: <http://www.usyd.edu.au/su/agric/acpa/>
- MORALES, A. y J. VILORIA. (2006). Aplicabilidad del enfoque de conjuntos borrosos a la clasificación de suelos de la depresión del Lago de Valencia, Venezuela. *Interciencia*, 31(8): 598-604.
- ODEH, IOA., MCBRATNEY, AB. y DJ. CHITTLEBOROUGH. (1992). Soil Pattern Recognition with Fuzzy-c-means: Application to Classification and Soil-Landform Interrelationships. *Soil Sci. Soc. Am. J.*, 56: 505-516.
- OLAWOYIN, R., NIETO, A., GRAYSON, RL., HARDISTY, F. y S. OYEWOLE. (2013). Application of artificial neural network (ANN)-self-organizing map SOM) for the categorization of water, soil and sediment quality in petrochemical regions. *Expert Systems with Applications*, 40(9): 3634-3648.
- OLSOFT. (2016). NeuroXLClusterizer. V.4.0.6. Neural network software add-ins for Microsoft Excel. Olsoft LLC. USA.
- PETROV, N., ANTONIYA, G. y I. JORDANOV. (2013). Self-organizing maps for texture classification. *Neural Computing y Applications*, 22 (7-8): 1499-1508.
- PINEDA, MC., ELIZALDE, G. y J. VILORIA. (2011a). Relación suelo-paisaje en un sector de la cuenca del río Caramacate, Aragua, Venezuela. *Revista de la Facultad de Agronomía*, 37: 27-37.
- PINEDA, MC., ELIZALDE, G. Y J. VILORIA. (2011b). Determinación de áreas susceptibles a deslizamientos en un sector de la Cordillera de la Costa Central de Venezuela. *Interciencia*, 36: 370-377.
- QUINTÍN, MM. y Y.R. PAZ. (2007). *Aplicación de las Redes Neuronales Artificiales a la Regresión*. Editorial: La Muralla. Madrid. 136 p. ISBN: 9788471337672
- ROSSITER, D.G. 2001. Assessing the thematic accuracy of area-class soil maps. Soil Science Division, ITC. Enschede The Netherlands. 31-August-2001. Available from: <http://www.itc.nl/personal/rossiter/>



- RUß, G., KRUSE, R., SCHNEIDER, M. y P. WAGNER. (2009). Visualization of Agriculture Data Using Self-Organizing Maps. En: *Applications and Innovations in Intelligent Systems XVI*, p.47-60.
- SÁNCHEZ, J.J. (1989). *Análisis de tablas de contingencia*. Centro de Investigaciones Sociológicas, nº 105. Madrid. 171p.
- SARMENTO, E., GIASSON, E., WEBER, E., FLORES, CA. y H. HASENACK. (2012). Prediction of soil orders with high spatial resolution: response of different classifiers to sampling density. *Pesq. agropec. bras.*, 47(9): 1395-1403.
- SOIL SURVEY STAFF. (2014). *Keys to Soil Taxonomy*. Twelfth Edition. United States Department of Agriculture. Natural Resources Conservation Service.NRCS-USDA. Washington DC. 372 p.
- STANKIEWICZ, A. y P. KOSIBA. (2009). Advances in ecological modelling of soil properties by self-organizing feature maps of natural environment of lower Silesia (Poland). *Acta societatis Botanicorum Poloniae*, 78(2): 167-174.
- 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
- VILORIA, JA., VILORIA-BOTELLO, A., PINEDA, MC. y A. VALERA. (2016). Digital modelling of landscape and soil in a mountainous region: A neuro-fuzzy approach. *Geomorphology*, 253: 199-207.
- WEBSTER, R. 2008. *El muestreo en los estudios del suelo*. 1ra edición. Trad. Christine Siebe. Universidad Nacional Autónoma de México. Sociedad Mexicana de la Ciencia del Suelo. México, DF. 98 p.
- WEBSTER, R. y MA. OLIVER. (1990). *Statistical methods in soil and land resource survey*. Oxford: Oxford University Press. 316 p.

CC BY-NC-ND