



UNIVERSIDAD NACIONAL DE COLOMBIA

Zonificación de la Hoya del Río Suárez por propiedades físicas del suelo, para el cultivo de caña panelera.

Ruy Edeymar Vargas Díaz.

Universidad Nacional de Colombia.

Facultad de Ciencias.

Departamento de Estadística.

Bogotá, Colombia.

2019.

Zonificación de la Hoya del Río Suárez por propiedades físicas del suelo, para el cultivo de caña panelera.

Ruy Edeymar Vargas Díaz.

Trabajo de grado presentado como requisito parcial para optar al título de:
Magister en Ciencias, Bioestadística.

Director:
Ramón Giraldo Henao.
PhD en Estadística.
Profesor Titular.

Línea de Investigación:
Bioestadística.

Universidad Nacional de Colombia.
Facultad de Ciencias.
Departamento de Estadística.
Bogotá, Colombia.
2019.

Dedicatoria

A mi familia, mamá y hermanos, que son mi norte en todo lo que realizo.

Agradecimientos

A mi director Ramón Giraldo, profesor titular adscrito al Departamento de Estadística de la Universidad Nacional de Colombia, por su disposición y asesoría. A Julio Ricardo Galindo Pacheco y Jeannette Amparo Español Aragón por su apoyo y orientación. A la Corporación Colombiana de Investigación Agropecuaria, por suministrar la información analizada en el estudio. A mis compañeros de las maestrías en Estadística y Bioestadística con quienes tuve el placer de compartir y aprender.

Resumen

La degradación del suelo a nivel mundial ha venido empeorando y ahora es crítica (Scholes et al., 2018), poniendo en riesgo el bienestar de 3200 millones de personas (IPBES, 2018). El manejo no sostenible de las tierras de cultivo y pastoreo es el factor que más influencia tiene sobre la degradación del suelo.

En Colombia la degradación del suelo por efecto de la erosión se ve intensificada por la topografía de alta montaña donde se establecen algunas explotaciones agrícolas (Manrique et al., 2000). Este es el caso del cultivo de caña panelera, que tiene su principal zona productora en el país, en la denominada Hoya del Río Suárez (*HRS*) (localizada en la cordillera oriental entre los departamentos de Boyacá y Santander) (Osorio, 2007). Esta región tiene las características de relieve antes mencionada, lo que unido a labores de labranza como corte y quema de rastrojo y arado con azadón que la hace propensa a la erosión (García et al., 2007).

La identificación de zonas susceptibles a este problema es de gran importancia en aras de formular planes de manejo que lo mitiguen. De acuerdo con Scholes et al. (2018) implementar acciones para combatir la degradación del suelo puede transformar la vida de millones de personas.

Este estudio se desarrolló con el objetivo de contribuir en la detección de zonas susceptibles a erosión en la *HRS*. Para ello se emplearon análisis geoestadísticos (Waller y Gotway, 2004) de cada una de las propiedades del suelo consideradas, de ejes factoriales (obtenidos a través de un análisis de componentes principales (Anderson, 2003)) y de un índice de erodabilidad (Ramírez et al., 2009).

Se cubrió un área de 45.000 hectáreas. Específicamente fueron analizados datos registrados en 935 sitios distribuidos a lo largo de toda la zona de interés (CORPOICA, 2016).

Los mapas generados indican que las áreas al sur de San Benito (en el departamento de Santander) y la parte oriental de los municipios de San José de Pare y Chitaraque (ambos en el departamento de Boyacá) son las más propensas a la erosión. Así mismo permiten establecer que las regiones más estables se encuentran en Santander hacia el sur de Chipatá y oriente de San Benito.

En las zonas susceptibles a erosión (identificadas a través de los mapas construidos) deben proponerse estrategias para el manejo del suelo que permitan detener e incluso revertir su degradación y que hagan del cultivo de caña una práctica rentable y responsable con el ambiente.

Palabras clave: componentes principales, erodabilidad, interpolación y predicción espacial, zonificación.

Abstract

Soil degradation worldwide is increasing and is now critical (Scholes et al., 2018), putting at risk the welfare of 3.2 billion people (IPBES, 2018). The unsustainable management of farmland and grazing land is the most influential factor in soil degradation.

In Colombia, soil degradation caused by erosion is intensified by the high mountain topography where some farms are established (Manrique et al., 2000). This is the case of the sugarcane crop, which has its main production area in the so called Hoya del Río Suárez (*HRS*) (located in the eastern ranges between the departments of Boyacá and Santander of Colombia) (Osorio, 2007). This region has the topography characteristics mentioned above, which together with tillage tasks such as cutting and burning of stubble and plowing with hoe makes it prone to erosion (García et al., 2007).

The identification of susceptible zones with this problem is of great importance to formulate management plans that mitigate the problem. According to Scholes et al. (2018), implementing actions to combat soil degradation can transform the lives of millions of people.

This study was developed with the aim of contributing to the detection of zones susceptibilities to erosion in the *HRS*. For this purpose, geostatistical analyses of each of the soil properties considered were used (Waller y Gotway, 2004), factorial axes (obtained through an analysis of principal components (Anderson, 2003)) and an index of erodability (Ramírez et al., 2009).

An area of 45 hectares was covered. The data recorded in 935 sites distributed throughout the area of interest were analyzed (CORPOICA, 2016).

The maps generated indicate that the areas south of San Benito (in the department of Santander) and the eastern part of the municipalities of San José de Pare and Chitaraque (both in the department of Boyacá) are the most prone to erosion. They also allow establishing that the most stable regions are located in Santander towards the south of Chipatá and east of San Benito.

In susceptible zones to erosion (identified through built in maps), strategies for soil management should be put in place to stop and even reverse their degradation and make sugarcane crop a profitable and responsible practice with the environment.

Keywords: principal components, erodibility, interpolation and spatial prediction, zoning.

Contenido

Agradecimientos	VII
Resumen	IX
Lista de figuras	XV
Lista de tablas	XIX
1 Introducción	2
2 Marco teórico	5
2.1 El suelo y su importancia	5
2.2 Propiedades físicas del suelo	8
2.2.1 Textura del suelo	9
2.2.2 Estabilidad estructural	10
2.2.3 Densidad real y aparente	11
2.2.4 Capacidad de retención de agua disponible	12
2.2.5 Erodabilidad	14
2.3 Erosión del suelo	15
2.4 Métodos estadísticos	17
2.4.1 Interpolación determinística	17
2.4.2 Predicción espacial usando geoestadística	18
2.4.3 Diagnóstico de predicción e interpolación	31
2.4.4 Geoestadística multivariada a través de análisis de componentes principales.	32
3 Área de estudio e información disponible	36
3.1 Hoya del Río Suárez (HRS)	36
3.2 Base de datos	37

4 Metodología	39
4.1 Definición del método de manejo de las variables	39
4.1.1 Análisis individual de las variables	39
4.1.2 Análisis conjunto por componentes principales	39
4.1.3 Análisis del índice de erodabilidad K	40
4.2 Evaluación de los métodos de interpolación y predicción espacial	40
4.2.1 Establecimiento de la rejilla para interpolación y predicción	40
4.2.2 Interpolación determinística	40
4.2.3 Predicción estocástica	41
4.3 Evaluación de las alternativas resultantes	42
4.3.1 Zonificación para las variables resultantes	42
4.3.2 Selección final de métodos de zonificación	43
5 Resultados y discusión	44
5.1 Análisis individual de las variables	44
5.2 Análisis conjunto por componentes principales	60
5.3 Análisis del índice de erodabilidad (K)	68
6 Conclusiones	73
7 Recomendaciones	74
Bibliografía	75

Lista de Figuras

2-1.	Gráfico de clases de agua en el suelo (Burbano y Silva, 2010).	13
2-2.	Representación de un semivariograma típico acotado, con parámetros: pepita $\hat{\tau}^2$, meseta $\hat{\sigma}^2$ y rango $\hat{\phi}$ (Waller y Gotway, 2004).	23
2-3.	Representación de regiones de tolerancia a distancias h_1 y h_2 a un ángulo de dirección θ (Waller y Gotway, 2004).	24
2-4.	Semivariogramas a diferentes ángulos de dirección θ (0, 45, 90 y 135 grados), usando datos de contenido de materia orgánica en el suelo de una parcela de 20 x 20 metros (Gallardo, 2006).	24
2-5.	Modelo de semivariograma esférico (izquierda) y exponencial (derecha) (Waller y Gotway, 2004).	26
3-1.	Derecha. El recuadro rojo enmarca la zona de estudio en Colombia. Izquierda. Municipios en color naranja son los sitios de los departamentos de Santander (Barbosa, Chipatá, Güepsa, San Benito, Suaita, Vélez) y Boyacá (Chitaraque, Moniquirá, San José de Pare, Santana y Togüí) donde se midieron las variables.	37
3-2.	Esquema de muestreo. Los puntos rojos corresponden a los lugares de medición en cada uno de los municipios.	38
5-1.	Diagramas de caja y bigotes (datos estandarizados) de las variables porcentaje de arenas (A), diámetro medio ponderado (DMP), densidad aparente (DA), densidad real (DR) y capacidad de retención de agua disponible ($CRAD$). Calculados con datos obtenidos en la <i>HRS</i> en 2016.	46
5-2.	Gráfico de dispersión de valores de porcentaje de arenas (A) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la <i>HRS</i> en 2016.	47

- 5-3. Gráfico de dispersión de valores de diámetro medio ponderado (*DMP*) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016. 47
- 5-4. Gráfico de dispersión de valores de densidad aparente (*DA*) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016. 48
- 5-5. Gráfico de dispersión de valores de densidad real (*DR*) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016. 48
- 5-6. Gráfico de dispersión de valores de porcentaje de capacidad de retención de agua disponible (*CRAD*) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016. 49
- 5-7. Semivariogramas direccionales (0, 45, 90 y 135 grados), para los procesos de porcentaje de arenas (izquierda) y diámetro medio ponderado (derecha). 50
- 5-8. Semivariogramas direccionales (0, 45, 90 y 135 grados), para el proceso de densidad aparente (izquierda) y densidad real (derecha). 50
- 5-9. Semivariogramas direccionales (0, 45, 90 y 135 grados), para la capacidad de retención de agua disponible 51
- 5-10. Semivariogramas experimentales (puntos negros) calculados con datos muestrales y ajuste de un modelo exponencial ($\hat{\tau}^2 = 108,88$, $\hat{\sigma}^2 = 81,64$ y $\hat{\phi} = 2639,03$) y uno esférico ($\hat{\tau}^2 = 0,34$, $\hat{\sigma}^2 = 0,31$ y $\hat{\phi} = 1922,91$) para las variables porcentaje de arenas (*A*) (izquierda) y diámetro medio ponderado (*DMP*) (derecha), respectivamente. 52
- 5-11. Semivariogramas experimentales (puntos negros) calculados con datos muestrales y ajuste de modelos exponenciales ($\hat{\tau}^2 = 0,021$, $\hat{\sigma}^2 = 0,007$ y $\hat{\phi} = 2818,13$) y ($\hat{\tau}^2 = 0,011$, $\hat{\sigma}^2 = 0,005$ y $\hat{\phi} = 1637,79$) para las variables de densidad aparente (*DA*) (izquierda) y real (*DR*) (derecha), respectivamente. 52
- 5-12. Semivariograma experimental (puntos negros) calculados con datos muestrales y ajuste de modelo esférico ($\hat{\tau}^2 = 5,55$, $\hat{\sigma}^2 = 1,98$ y $\hat{\phi} = 4638,14$) para la variable capacidad de retención de agua disponible (*CRAD*) (línea continua). 53

5-13.	Distribución espacial para el porcentaje de arenas (A). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).	54
5-14.	Distribución espacial para el diámetro medio ponderado (DMP). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).	55
5-15.	Distribución espacial para la densidad aparente (DA). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging universal (derecha).	55
5-16.	Distribución espacial para la densidad real (DR). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).	56
5-17.	Distribución espacial de la capacidad de retención de agua disponible ($CRAD$). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).	57
5-18.	Distribución espacial de varianzas de la predicción del kriging del porcentaje de arenas (A) (izquierda) y diámetro medio ponderado (DMP) (derecha).	58
5-19.	Distribución espacial de varianzas de la predicción del kriging de la densidad aparente (DA) (izquierda) y real (DR) (derecha).	59
5-20.	Distribución espacial de varianzas de la predicción del kriging de la capacidad de retención de agua disponible ($CRAD$).	59
5-21.	Diagrama de caja y bigotes de los dos primeros componentes principales, generados a partir de información medida en la HRS en 2016.	62
5-22.	Gráfico de dispersión de valores “scores” del componente principal uno, respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en HRS en 2016.	63
5-23.	Gráfico de dispersión de valores “scores” del componente principal dos, respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la HRS en 2016.	63
5-24.	Semivariogramas direccionales (0, 45, 90 y 135 grados), para los componentes principales uno (izquierda) y dos (derecha).	64
5-25.	Semivariogramas experimentales (puntos negros) calculados con datos muestrales y ajuste de modelos esférico ($\hat{\tau}^2 = 1,30$, $\hat{\sigma}^2 = 0,31$ y $\hat{\phi} = 6308,48$) y exponencial ($\hat{\tau}^2 = 0,51$, $\hat{\sigma}^2 = 0,56$ y $\hat{\phi} = 837,02$) para las variables componente principal uno (izquierda) y dos (derecha), respectivamente.	64

5-26.	Distribución espacial del primer componente principal. Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).	66
5-27.	Distribución espacial del segundo componente principal. Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging universal (derecha).	66
5-28.	Distribución espacial de varianzas de la predicción del kriging del primer (izquierda) y segundo componente principal (derecha).	67
5-29.	Diagrama de caja y bigotes del índice de erodabilidad (K), generados a partir de información medida en la <i>HRS</i> en 2016.	69
5-30.	Gráfico de dispersión de los valores de erodabilidad (puntos negros), respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la <i>HRS</i> en 2016.	69
5-31.	Semivariogramas direccionales (0, 45, 90 y 135 grados) (izquierda), y experimental (puntos negros) calculados con datos muestrales con ajuste de modelo teórico exponencial ($\hat{\tau}^2 = 2,1e - 06$, $\hat{\sigma}^2 = 4,8e - 06$ y $\hat{\phi} = 609,23$) (derecha), para el índice de erodabilidad (K).	70
5-32.	Distribución espacial para el índice de erodabilidad (K). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).	71
5-33.	Distribución espacial de varianzas de la predicción de kriging ordinario del índice de erodabilidad (K).	72

Lista de Tablas

2-1.	Sistemas de clasificación de partículas según su tamaño propuestos por el Departamento de Agricultura de Norteamérica (USDA) y la Sociedad Internacional de Ciencia del Suelo (ISSS). Fuente (Giménez, 2004)	10
2-2.	Clasificación de diámetro medio ponderado (<i>DMP</i>) en milímetros (mm).	11
2-3.	Rangos para interpretar la <i>CRAD</i> , propuesta por Montenegro y Malagón (1990)	14
2-4.	Clasificación de erodabilidad, propuesta por Ramírez et al. (2009)	15
4-1.	Valores para definir condiciones favorables y susceptibles de las variables porcentaje de arenas (<i>A</i>), diámetro medio ponderado (<i>DMP</i>), densidad aparente (<i>DA</i>), densidad real (<i>DR</i>) y capacidad de retención de agua disponible (<i>CRAD</i>).	42
5-1.	Medidas de localización y variabilidad de las variables porcentaje de arenas (<i>A</i>), diámetro medio ponderado (<i>DMP</i>), densidad aparente (<i>DA</i>), densidad real (<i>DR</i>) y capacidad de retención de agua disponible (<i>CRAD</i>). Calculados con datos obtenidos en la <i>HRS</i> en 2016.	45
5-2.	Estadísticos de validación cruzada (Error medio de predicción <i>EMP</i> , raíz del error cuadrado medio de predicción <i>RECM</i> y coeficiente de determinación R^2) de los análisis de predicción (Kriging ordinario <i>KO</i> y universal <i>KU</i>) e interpolación (inverso de la distancia <i>DIP</i>), realizados con las variables porcentaje de arenas (<i>A</i>), diámetro medio ponderado (<i>DMP</i>), densidad aparente (<i>DA</i>), densidad real (<i>DR</i>) y capacidad de retención de agua disponible (<i>CRAD</i>).	58
5-3.	Coeficientes de correlación de Pearson (valor-p), calculados con la información de las variables físicas del suelo; arenas (<i>A</i>), diámetro medio ponderado (<i>DMP</i>), densidad aparente (<i>DA</i>), densidad real (<i>DR</i>) y capacidad de retención de agua disponible (<i>CRAD</i>), medidas en la <i>HRS</i> en 2016.	61

5-4.	Valores propios y porcentajes de varianza explicados por los componentes principales generados con la información de cuatro variables físicas del suelo, medidas en la <i>HRS</i> en 2016.	61
5-5.	Vectores propios de los dos primeros componentes principales.	62
5-6.	Estadísticos de validación cruzada (Error medio de predicción <i>EMP</i> , raíz del error cuadrado medio de predicción <i>RECM</i> y coeficiente de determinación R^2) de los análisis de predicción (Kriging ordinario <i>KO</i> y universal <i>KU</i>) e interpolación (inverso de la distancia <i>DIP</i>), realizados con los componentes principales uno y dos.	67
5-7.	Medidas de localización y variabilidad de la variable índice de erodabilidad (<i>K</i>). Calculados con datos obtenidos en la <i>HRS</i> en 2016.	68
5-8.	Estadísticos de validación cruzada (Error medio de predicción <i>EMP</i> , raíz del error cuadrado medio de predicción <i>RECM</i> y coeficiente de determinación R^2) de los análisis de predicción (Kriging ordinario <i>KO</i>) e interpolación (inverso de la distancia <i>DIP</i>), realizados con el índice de erodabilidad (<i>K</i>).	72

1 Introducción

El suelo es fundamental en la agricultura porque provee de agua y nutrientes a los cultivos (Silva y Correa, 2009). Este interviene en los ciclos del agua, carbono, nitrógeno, fósforo y otros elementos (Ferrerías et al., 2015). Su pérdida (debida a procesos naturales o desencadenada como consecuencia de sobre-explotación del medio) pone en riesgo el bienestar de aproximadamente el 40 % de la humanidad (Scholes et al., 2018). Entre otros aspectos causa la extinción de algunas especies, influye sobre el cambio climático y genera emigración (IPBES, 2018). El manejo no sostenible de tierras de cultivo y pastoreo puede propiciar detrimento de diversidad biológica, inseguridad alimentaria y carencia de agua y energía (Scholes et al., 2018). La industrialización, la globalización y el crecimiento demográfico han expandido la frontera agrícola, induciendo el deterioro de las propiedades físicas, químicas y biológicas del suelo (Novillo et al., 2018).(Bravo et al., 2011)

Un tipo de degradación del suelo es la erosión, definida como el desgaste de la capa superficial de la corteza terrestre por acción del agua y/o el viento (Burbano y Silva, 2010). Esta conlleva la carencia de nutrientes, materia orgánica y humedad (Montenegro y Malagón, 1990). La disminución tolerable de suelo está alrededor de 11 toneladas por hectárea (Johnson, 1987). En zonas donde las tasas de erosión son naturalmente altas, por ejemplo en áreas montañosas de gran precipitación (Peláez, 2003), se supera ampliamente este valor de referencia (Morgan, 1986). Todo esto en ocasiones redundo en pobreza, en desarraigo de la tierra e incluso en violencia (León, 2001).

En Colombia las cifras sobre pérdida del suelo por erosión oscilan entre 17.000 y 20.000 hectáreas por año (Montenegro y Malagón, 1990). Estos son sin duda valores considerables que generan alerta. Como en otros campos de la ciencia, la identificación temprana del problema es esencial para su mitigación y/o solución. Por ello, la detección de zonas propensas a la erosión (erodabilidad) se ha convertido en un campo de investigación muy relevante. Generalmente para este fin se emplean técnicas experimentales (de campo y laboratorio) basadas en la medición de propiedades físicas del suelo (Montenegro y Malagón, 1990) o en el cálculo de índices de erodabilidad (Ramírez et al., 2009). Las muestras de suelo son

tomadas en muchos sitios de la región de interés. Usando su georreferenciación y empleando técnicas de interpolación y/o predicción (inverso de la distancia y Kriging), posteriormente se pueden generar mapas que permiten establecer aquellas zonas más propensas a erosión. Por ejemplo Addis y Klik (2015) y Ayoubi y Alizadeh (2005) emplean técnicas Kriging para modelar la distribución espacial de la susceptibilidad a la erosión en regiones de Etiopía e Irán, respectivamente.

En la literatura se han referenciado muchos estudios sobre erosión y su relación con las propiedades físicas y químicas del suelo (Villatoro et al., 2008; Mousavi et al., 2017; Henríquez et al., 2005; Moreno, 2011). Colombia no ha sido ajeno a este esfuerzo. Entre otros, en este contexto, se pueden citar los trabajos de Bravo et al. (2011), Jaramillo et al. (2011) y Vásquez et al. (2010). No obstante cabe resaltar que en el medio local, usualmente, no se ha considerado la variación espacial de las propiedades, particularmente en las regiones agrícolas más susceptibles (zonas de montaña).

Dado que la erosión generalmente obedece a la interacción de varias propiedades, en ocasiones se acude para su estudio al empleo de herramientas estadísticas multivariadas (por ejemplo análisis componentes principales) que permiten detectar el grado en el que estas, de manera conjunta, influyen sobre el fenómeno (Ghaemi et al., 2014; Kooch et al., 2008).

El cultivo para la producción de panela es la segunda agroindustria más importante de Colombia después del café (Ahumada, 2015). La zona de la Hoya del Río Suárez (*HRS*), localizada en la cordillera oriental entre los departamentos de Boyacá y Santander (Osorio, 2007), es la de mayor preponderancia para dicho cultivo. En la región, la plantas de caña se establecen en zonas de alta pendiente (catalogadas como media y altas) con inclinaciones entre el 10 % y el 40 %, que hacen los suelos propensos a la erosión (Manrique et al., 2000). Dicha situación se acentúa con las labores de labranza como corte y quema de rastrojo, uso de azadón, arado con bueyes y maquinaria (García et al., 2007). La detección de las zonas propensas a la erosión en esta región resulta de suma importancia para establecer planes de manejo sostenible del suelo y del cultivo.

Los estudios sobre zonificación de la erosión en Colombia son limitados. A un nivel de resolución alto (todo el territorio colombiano) el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) llevó a cabo una zonificación de acuerdo a la susceptibilidad de degradación por erosión. Para ello tomaron como insumos imágenes de satélite y encuestas puntuales (IDEAM, 2015). En la *HRS* se han realizado varios estudios encaminados a la

fertilización y al mejoramiento de las variedades de caña (Rodríguez y Gottret, 1998), pero sin tener en cuenta la variabilidad espacial de las propiedades del suelo.

Teniendo en cuenta lo anterior, en este trabajo se hace un estudio detallado de la susceptibilidad a la erosión en la *HRS*. Este puede ser tomado por las instituciones estatales como insumo para la elaboración de planes de manejo específicos para la región. Particularmente se usan mediciones de variables físicas del suelo en distintos puntos del área para establecer cuáles son las regiones que requieren planes de manejo focalizados.

Se plantean en el trabajo tres estrategias de modelación para identificar, de acuerdo a los valores de las variables físicas del suelo, las zonas de la región susceptibles a erosión. Inicialmente se construyen mapas de distribución espacial de cada una de las variables. Posteriormente, para identificar el efecto conjunto de las propiedades del suelo sobre la erosión se hace un análisis de variabilidad espacial de ejes factoriales (calculados a partir de un análisis de componentes principales). Finalmente, se calcula un índice de erodabilidad (propuesto por Ramírez et al. (2009)) y se obtiene nuevamente un mapa de distribución espacial. A partir del estudio global de los mapas calculados por los tres caminos, se identifican las zonas más propensas a erosión dentro de la región.

El documento está organizado de la siguiente manera: En el capítulo 2 se presenta un marco teórico en el que se definen conceptos relacionados con el suelo y sus características físicas (las que intervienen en la susceptibilidad a la erosión). También se hace en este mismo una revisión de los métodos estadísticos usados en el análisis. En el capítulo 3 se describe el área de estudio y la información disponible para el análisis. En el capítulo 4 se detalla la metodología de trabajo. Posteriormente, en el capítulo 5 se resaltan los resultados más relevantes. Al final del trabajo se dan conclusiones y recomendaciones prácticas.

2 Marco teórico

En este capítulo se presentan inicialmente las definiciones relacionadas con el estudio del suelo. Posteriormente, se revisa el concepto de erosión y como esta afecta la estabilidad de los suelos. Finalmente, se muestran las metodologías estadísticas empleadas para dar cumplimiento a los objetivos propuestos en el trabajo.

2.1. El suelo y su importancia

La información y conceptos presentados en esta sección son tomados de Burbano y Silva (2010), Montenegro y Malagón (1990), Hillel (1998), Buol et al. (1974), López (2005) y Ramírez (1997).

A través de la historia se han escrito numerosas definiciones del suelo. Algunos autores simplemente mencionan que es el resultado de las condiciones climáticas y de los organismos vivos sobre un material geológico original a lo largo del tiempo. Otros, utilizando nuevos argumentos, lo describen como un cuerpo natural involucrado en interacciones dinámicas con la atmósfera y con los estratos que están por debajo, que influye en el clima y en el ciclo hidrológico del planeta y que sirve como medio de crecimiento para una variada comunidad de seres vivientes. Además, desempeña un papel ambiental preponderante a manera de reactor bio-físico-químico que descompone materiales de desecho y recicla dentro de él nutrientes con el fin de generar continuidad de la vida en la tierra.

¿Cuáles son los factores que intervienen en la formación del suelo?. La literatura menciona que los organismos, el clima, el material parental, el relieve y el tiempo son importantes como agentes modificadores de la superficie terrestre.

Los organismos, macroorganismos, mesoorganismos y microorganismos, de origen vegetal o animal, afectan al suelo química y físicamente, aportando elementos orgánicos que liberan los nutrientes necesarios para favorecer la fertilidad de los terrenos. Las plantas constituyen

la principal fuente de materia orgánica (*MO*), además la vegetación posee un papel clave en la formación del suelo, por su capacidad de meteorización de la roca (líquenes, raíces, etc.). Los animales excavadores trituran y mezclan el material, lo que influye sobre la estructuración, la permeabilidad y la aireación. Los vertebrados y algunos invertebrados, como los artrópodos, son responsables de la bioturbación del suelo. Los invertebrados no artrópodos colaboran en la alteración de la *MO* y la mezcla con los minerales de la superficie.

El clima es considerado como el factor ambiental más activo y de mayor influencia en la evolución de los suelos. Se destaca la precipitación, la temperatura, los vientos y los fenómenos de evapotranspiración. Estos factores influyen sobre el pH, el contenido de *MO*, la formación de horizontes del suelo, la actividad microbiológica, entre muchos otros.

Por otra parte, el material parental es donde se deriva el suelo. Es el producto de la alteración de las rocas o sedimentos no consolidados, de cualquier procedencia y composición. El clima y los organismos actúan sobre este material generando un tipo de suelo determinado. De este proceso se da origen a las propiedades físicas como color, textura, profundidad, estructura, densidad, porosidad, entre otras y las características químicas (pH, capacidad de intercambio catiónico y salinidad).

El relieve que se relaciona con las formas del terreno y la pendiente, afecta la acumulación o acarreo de materiales por gravedad que determinan la profundidad del suelo, concentración de sales, erosión y la diferenciación de horizontes. Los relieves abruptos favorecen la degradación, originando suelos poco profundos y condicionan aspectos como el drenaje y procesos geomorfológicos. En los valles, por el contrario se favorece el desarrollo de la capa arable.

Respecto a la edad los suelos se clasifican como jóvenes, maduros y altamente evolucionados. Los más antiguos son los que muestran un alto desarrollo en profundidad del perfil y mayor diversificación de horizontes. La velocidad de formación de este recurso va desde 0,001 mm/año hasta 1 mm/año. Un centímetro de capa arable tarda cientos de años en formarse y puede perderse por arrastre con solo una lluvia o viento fuerte.

El suelo es un sistema polifásico (tres fases), heterogéneo (diversas sustancias minerales y orgánicas), particulado (compuesto de partículas de diversa forma y tamaño), disperso (los componentes son separables) y poroso (permite el paso del aire y los solutos) en el cual ocurren interacciones interfásicas, dando lugar a procesos y fenómenos tales como la cohesión-adhesión, agregación-dispersión, sorción-desorción, expansión-contracción, tensión

superficial, retención y flujo hídrico, intercambio iónico y gaseoso, entre otros aspectos.

El sistema polifásico mencionado anteriormente se refiere a que físicamente el suelo es un medio poroso que está conformado por tres fases. La primera es la sólida que constituye el esqueleto o la matriz del suelo. La segunda es la líquida compuesta por el agua que contiene sustancias disueltas (solución) que ocupa una parte o todos los espacios porosos entre las partículas sólidas y por último la gaseosa o de vapor que es la atmósfera que se establece en aquellos lugares que se encuentran vacíos, es decir, los poros que no son llenados por el agua.

En términos de volumen, la fase sólida es aproximadamente el 50 % del total, mientras que la gaseosa y líquida se reparten por igual el resto del espacio disponible. Esta es una proporción idealizada en un terreno cultivado para que pueda ofrecer un medio adecuado al crecimiento de las plantas. De acuerdo con la dinámica de los factores y procesos formadores se origina un suelo con propiedades físicas, químicas y biológicas particulares.

La fase sólida proviene de la descomposición de las rocas y de los residuos vegetales, es relativamente estable en cuanto a su constitución y organización. Dicha estabilidad suele servir para la caracterización y está organizada por un conjunto de partículas minerales y orgánicas cuya disposición permite una cantidad determinada de poros. La fracción mineral que procede de la meteorización de la roca madre por acción de agentes climáticos se compone por elementos de varios tamaños como son la arena, el limo y la arcilla, que ocupan el 45 % del volumen total. El ordenamiento de estas fracciones posibilita la existencia de una proporción variable de espacio poroso. La mayoría de las propiedades físicas de un suelo depende de la distribución de dimensiones de sus componentes y la composición mineralógica que influye en la retención y suministro de nutrientes.

La *MO* ocupa el 5 % del volumen total e incluyen residuos vegetales y animales en descomposición como también organismos en vida activa. Los terrenos de cultivo contienen entre el 2 y el 10 % de *MO* que interviene en las propiedades físicas, químicas y biológicas del suelo. La *MO* se descompone por la acción de los microorganismos hasta transformarse en humus que da el color oscuro a los suelos e influye en la estructura, aireación y capacidad de retención de agua y nutrientes.

La disposición y acomodación de las partículas de la fase sólida, define una serie de características físicas, como estructura, porosidad, permeabilidad y densidad. Por tanto, esta fase

es la fuente de la mayoría de los nutrientes vegetales, es el almacén del agua requerida por las plantas y determina la eficiencia del suelo al desempeñar las funciones que permiten el desarrollo de la vida vegetal.

2.2. Propiedades físicas del suelo

Las definiciones a continuación son tomadas de Burbano y Silva (2010), García (2013), Luna (2006), Montenegro y Malagón (1990) y Rucks et al. (2004).

El comportamiento mecánico de la fase sólida determina a su vez las propiedades físicas del suelo, las cuales en asociación con las características químicas generan la habilidad para producir cosechas.

Las características físicas son en gran parte responsables del buen desarrollo de las plantas, pero pocas veces se les tiene en cuenta. Estas propiedades y algunas de ellas asociadas con la estructura, han sido llamadas, la clave de la productividad del suelo sin desconocer la estrecha relación entre las variables químicas y físicas.

Las propiedades físicas del suelo (textura, estructura, consistencia, densidad, porosidad, entre otras), inciden de manera considerable sobre el crecimiento de las plantas y la disponibilidad de nutrientes, ya que determinan la rigidez, la facilidad de penetración de las raíces, la aireación, capacidad de drenaje, almacenamiento de agua, la plasticidad y la retención de elementos nutricionales. Estos aspectos son fundamentales para la sostenibilidad y actividades asociadas al manejo agrícola.

Es necesario resaltar que el arreglo de la fase sólida, la naturaleza y distribución por tamaño de agregados y del espacio poroso entre sus partículas constituyentes (estructura), es determinante en el desarrollo de las propiedades físicas y de la fertilidad del suelo.

A continuación, se explican algunas características físicas:

2.2.1. Textura del suelo

Las definiciones a continuación son consideradas de Giménez (2004), Burbano y Silva (2010) y Montenegro y Malagón (1990).

La textura hace referencia a la composición granulométrica de la fracción inorgánica del suelo que se expresa en porcentaje. Se define como la proporción relativa en que se encuentran en una masa de suelo, varios grupos de partículas individuales de diámetro menor a 2 mm. Por su tamaño, los componentes del suelo se clasifican como arena, limo y arcilla. La porción de arena tiene diámetros de 2 a 0,05 mm, las de limo entre 0,05 y 0,002 mm, y las de arcilla son menores de 0,002 mm. Los sistemas referentes de clasificación de acuerdo con sus tamaños son propuestos por el Departamento de Agricultura de Norteamérica (USDA) y por la Sociedad Internacional de Ciencia del Suelo (ISSS) (Tabla **2-1**).

Las arenas, constituyen la fracción gruesa del suelo que le imprime menor capacidad de almacenamiento de humedad, drenaje alto, baja retención de nutrientes, excesiva aireación, susceptibilidad al encostramiento superficial, facilidad de laboreo mecánico y son fácilmente arrastradas por el agua y el viento. Los limos tienen mejor dinamismo químico e hidrodinámico que las arenas, pero inferior que las arcillas. Los territorios ricos en limos presentan buenas condiciones en términos generales para la actividad agropecuaria. Están asociados con zonas de valles y son fácilmente erosionables. La arcilla es la partícula más activa física, química e hidrodinámica de los suelos. Los terrenos con este componente tienden a presentar drenaje deficiente, dificultad en la preparación de la capa arable, mayor encharcamiento en la superficie, retención de agua, nutrientes, contenido de *MO* y resistencia a la erosión.

La textura es uno de los atributos más estables del suelo, aunque puede ser modificada por la remoción de horizontes superficiales por laboreo y el desarrollo de una nueva superficie diferente por la acumulación de materiales acarreados por el viento o el agua de riego. La fracción textural se expresa en porcentaje.

Tabla 2-1: Sistemas de clasificación de partículas según su tamaño propuestos por el Departamento de Agricultura de Norteamérica (USDA) y la Sociedad Internacional de Ciencia del Suelo (ISSS). Fuente (Giménez, 2004)

USDA		ISSS	
Fracción	Diámetros (<i>mm</i>)	Fracción	Diámetros (<i>mm</i>)
Arena muy gruesa	2,00-1,00	Arena	2,00-0,20
Arena gruesa	1,00-0,50	Arena fina	0,20-0,02
Arena media	0,50-0,25	Limo	0,02-0,002
Arena fina	0,25-0,10	Arcilla	< 0,002
Arena muy fina	0,10-0,05		
Limo	0,05-0,002		
Arcilla	< 0,002		

2.2.2. Estabilidad estructural

Los conceptos sobre esta propiedad física son tomados de Kohnke (1968), Narro (2004), Burbano y Silva (2010) y Montenegro y Malagón (1990).

Las partículas orgánicas y minerales que conforman el suelo no se encuentran aisladas unas de otras, sino que forman agregados que constituyen la estructura. El agregado (terrón) se forma teniendo en cuenta la fuerza mecánica producida por el movimiento de las raíces, animales, prácticas agrícolas, que producen un acercamiento entre los componentes del suelo o también por la acción de un agente cementante como la materia orgánica y las arcillas.

La estructura es considerada como una de las principales propiedades, ya que el arreglo que presente la fase sólida está determinando el espacio que queda disponible en las fases líquida y gaseosa. Los poros permiten el movimiento de gases y líquidos, ofreciendo un entorno favorable a la actividad de los microorganismos y facilitando el crecimiento radicular de las plantas. La resistencia que los componentes del suelo hacen para no ser destruidos se conoce como estabilidad estructural. Esto se define en la capacidad de los elementos en conservar su forma cuando se humedecen. Depende de la disposición de las partículas y de la adhesión de las mismas en la formación estructuras mayores. Se usa en la valoración de la estructura del suelo y se evalúa estableciendo el grado de agregación, la firmeza de los agregados y

la naturaleza de poros vacíos. Estas características cambian con las labores agrícolas y los diferentes sistemas de cultivo. Además, es una propiedad fundamental en la conservación de la estructura y del suelo mismo. Algunos se deterioran frente a la acción del agua lluvia o de la escorrentía, mientras otros muestran rigidez.

Existen diversos métodos para medir el grado de estabilidad de los agregados al agua, siendo el más utilizado el método de Yoder. Consiste en colocar una serie de tamices¹ en un aparato con agitación vertical. A partir del porcentaje de componentes acumulados en cada tamiz se obtiene el índice de agregación, que corresponde al diámetro medio ponderado (*DMP*). A continuación, se presenta la clasificación del *DMP* de acuerdo con Gómez y Osorio (2016):

Tabla 2-2: Clasificación de diámetro medio ponderado (*DMP*) en milímetros (mm).

<i>DMP</i> (mm)	Estabilidad estructural
< 0,5	Inestable
0,5-1,5	Ligeramente estable
1,5-3,0	Moderadamente estable
3,0-5,0	Estable
>5	Muy estable

2.2.3. Densidad real y aparente

La descripción de estas dos propiedades se realiza tomando como referencia a Burbano y Silva (2010), Giménez (2004), Lal y Shukla (2004), Montenegro y Malagón (1990) y Narro (2004).

La densidad es la relación entre la masa y el volumen del suelo y puede referirse a la densidad total del terreno (densidad aparente) o de las partículas (densidad real).

La densidad real de las partículas (*DR*) se define como la relación entre la masa total de los sólidos y el volumen de los mismos, sin incluir el espacio poroso. Las unidades de expresión son gramos sobre centímetros cúbicos (g/cm^3). La *DR* permite calcular la porosidad total del suelo, y también se utiliza para evaluar la concentración de los sólidos en suspensión,

¹Malla metálica constituida por barras tejidas y que dejan un espacio entre sí por donde se hace pasar el suelo previamente triturado.

determinar la velocidad de sedimentación de los componentes en líquidos o gases, realizar cálculos de erosión y en la predicción de pérdida y conservación de terrenos. Cuando no se presentan cantidades considerables de MO , fluctúa entre 2,5 y 2,6 g/cm^3 y alcanza en suelos arcillosos o arenosos con poca MO un valor de 2,65 g/cm^3 . En regiones con abundante contenido de hierro sobrepasa el 2,7 g/cm^3 .

La densidad aparente del suelo (DA), es la relación entre la masa de los sólidos y el volumen total que éstos ocupan, es decir, se incluye el espacio existente de las partículas sólidas. Es la característica que reviste gran importancia desde varios puntos de vista. En primer lugar permite calcular la porosidad y transformar la humedad gravimétrica² en volumétrica³. Se utiliza como parámetro en la clasificación de suelos orgánicos y derivados de cenizas volcánicas, en la determinación de la capa arable y en riegos se establece la lámina de agua. Al igual de la DR es también afectada por la MO , los valores más bajos se obtienen en suelos con alto contenido orgánico ($< 1 g/cm^3$), ya que a medida que disminuye la MO se alcanza por lo general un mayor valor de DR .

2.2.4. Capacidad de retención de agua disponible

La bibliografía usada en esta sección es tomada de Burbano y Silva (2010), Montenegro y Malagón (1990), Narro (2004) y Valero (1999).

La capacidad de retención de agua disponible ($CRAD$) es una de las características que está relacionada con la cantidad de agua disponible para el crecimiento de las plantas. Afecta especialmente las reacciones físicas y químicas que suceden en el suelo.

La $CRAD$ se define como la cantidad de agua en un perfil del suelo que puede teóricamente extraerse por la planta, después de que el terreno drena libremente por gravedad. Se define como la diferencia entre la capacidad de campo (CC) y el punto de marchitez permanente (Pmp). Se expresa como:

$$CRAD = CC - Pmp, \quad (2-1)$$

donde, la CC es la capacidad de campo (%), la cual se determina mediante el contenido de humedad del suelo entre 0,1 y 0,3 bares y la Pmp es el punto de marchitamiento permanente

²Relación entre la masa del agua y la seca de los sólidos del suelo

³Relación entre el volumen de agua y el total del suelo

(%) (contenido de humedad a 15 bares).

El término *CC* fue introducido por Veihmeyer y Hendrickson (1931), y se define como la cantidad máxima de agua que un terreno puede retener o almacenar, bajo condiciones de humedecimiento total seguido de drenaje libre o gravitacional (Figura 2-1). Generalmente se determina en laboratorio, sometiendo una muestra de suelo inalterada a una energía de retención entre 0,10 y 0,33 bares. El concepto de *Pmp* fue propuesto por Briggs y Shantz (1911). Es el estado particular en que el agua es retenida por el suelo con fuerza superior a la succión de las plantas (próximo a 15 atmósferas) donde ellas pueden satisfacer algunas de sus necesidades hídricas, pero no todas. La *CRAD* se interpreta según la tabla (2-3).

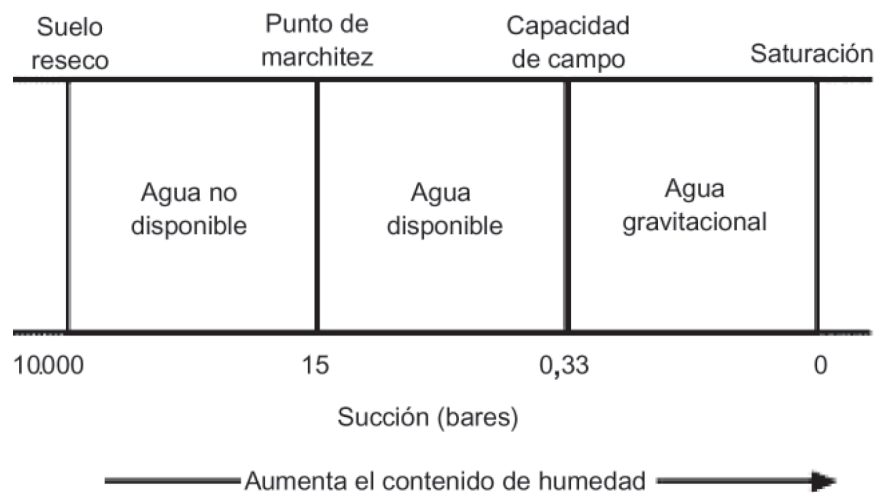


Figura 2-1: Gráfico de clases de agua en el suelo (Burbano y Silva, 2010).

Tabla 2-3: Rangos para interpretar la *CRAD*, propuesta por Montenegro y Malagón (1990)

Contenido de humedad %	Interpretación
Menor de 5	Muy baja
5 a 15	Baja
15 a 25	Media
25 a 35	Alta
Mayor a 35	Muy alta

2.2.5. Erodabilidad

Las siguientes definiciones son tomadas de Alarcón y Reyes (2013) y Montenegro y Malagón (1990).

El término erodabilidad fue sugerido por Cook (1937) para indicar la susceptibilidad de un suelo, en particular, a ser erosionado. Puede ser calculado mediante condiciones de pérdida de suelo en campo (parcelas de escorrentía), pero este método es costoso y requiere de períodos grandes de tiempo. Por ello, existe la tendencia a analizar muestras en laboratorio y calcular una medida cuantitativa de la erodabilidad basada en propiedades físicas de los terrenos. Simuladores de lluvia y algunos modelos matemáticos se han implementado como una alternativa en el estudio y determinación de la erosión.

Por medio de un modelo matemático basado en la aplicación de los procesos que intervienen en la erosión, Wischmeier y Smith (1978) desarrollaron la ecuación USLE (Universal Soil Loss Equation), considerando la erodabilidad (K) que permite conocer la susceptibilidad de los suelos al desprendimiento de macroagregados, microagregados y partículas, por efecto de los agentes erosivos.

Un modelo de regresión múltiple (2-2) desarrollado por Ramírez et al. (2009), incorpora tres variables físicas del suelo (contenido de arenas, diámetro medio ponderado y densidad aparente), para estimar la erodabilidad de un terreno.

$$K = -0,0007 + 0,0001A - 0,0025DMP + 0,0045DA, \quad (2-2)$$

donde, K es el índice de erodabilidad. A es el contenido de arenas en %. DMP es el diámetro medio ponderado en milímetros (mm). DA es la densidad aparente en g/cm^3 .

Los valores de K , se pueden clasificar de acuerdo con el rango propuesto por Ramírez et al. (2009), desde natural (baja susceptibilidad a erosión) hasta extremadamente severo (muy propenso a erosión) (Tabla 2-4).

Tabla 2-4: Clasificación de erodabilidad, propuesta por Ramírez et al. (2009)

ERODABILIDAD (K)	CLASIFICACIÓN
Menor de 0,01	Natural
0,01 a 0,02	Muy bajo
0,02 a 0,04	Bajo
0,04 a 0,06	Moderado
0,06 a 0,08	Alto
0,08 a 0,10	Muy alto
0,10 a 0,15	Severo
Mayor de 0,15	Extremadamente severo

2.3. Erosión del suelo

Las definiciones mostradas en esta sección son tomadas de Duque y Escobar (2002), Hincapié y Ramirez (2013), Prado y da Vega (1992), Montenegro y Malagón (1990) y Scholes et al. (2018).

La erosión es la remoción y el transporte de partículas superficiales por acción del viento (erosión eólica) y el agua (erosión hídrica). Ambas adquieren energía cuando se presentan en forma de precipitación, escorrentía y de vientos, que al estar en contacto con el suelo vencen la resistencia del mismo generando la erosión.

A nivel mundial la pérdida de suelo agrícola por efectos de la erosión se estima entre 6 y 7 millones de toneladas por año. En 2014 más de 1500 millones de hectáreas de ecosistemas naturales habían sido convertidas en territorios de cultivo. Menos del 25 % de la superficie

terrestre ha escapado de los impactos sustanciales de la actividad humana. Estimaciones del Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services IP-BES (2018) calculan que en el 2050 el porcentaje se habrá reducido a menos del 10 %. Las tierras dedicadas a agricultura y pastoreo en la actualidad cubren más de un tercio de la Tierra, con desmonte reciente de hábitats nativos que incluye bosques, pastizales y humedales, concentrados en algunos de los ecosistemas ricos en especies animales y vegetales del planeta. Además, la creciente demanda de alimentos y biocombustibles probablemente llevará al aumento continuo de aportes químicos y de nutrientes y a un cambio hacia los sistemas industrializados de producción de ganado. En Colombia no hay cifras actuales que demuestren la magnitud del problema, reportes afirman que alrededor de 2,3 millones de hectáreas han dejado de ser aptas para la explotación agrícola debido a este fenómeno.

La erosión puede ser natural, ocurre lentamente sin la intervención del hombre y se intensifica con las actividades humanas en la agricultura y la construcción. Son múltiples los agentes que favorecen e incrementan la degradación en los suelos, como el grado de inclinación, la longitud de la pendiente (topografía), el clima, factores del suelo y vegetación presente. Los territorios arenosos y limosos son más susceptibles a erosionarse que los arcillosos y aquellos con alto porcentaje de *MO*.

Los procesos anteriormente descritos implican el detrimento de suelo en mayor o menor proporción. Los efectos que puede causar son variados, entre ellos la disminución de la productividad debido a la pérdida de *MO*, de nutrientes o a la compactación en la zona de raíces, lo que reduce la infiltración y aumenta la escorrentía. El creciente disturbio de las cuencas hidrográficas, la tala y el sobrepastoreo son, quizá, las principales causas de las inundaciones que en la actualidad son consideradas como desastres naturales. Los sedimentos sepultan tierras productivas y destruyen cosechas y al acumularse en los lechos de los ríos reducen su caudal, la posibilidad de navegación y la vida útil de las represas. Además, la salinización de los regiones irrigadas, el uso de plaguicidas, la falta de rotación de cultivos y el exceso de mecanización agotan progresivamente los terrenos.

2.4. Métodos estadísticos

En esta sección inicialmente se hace una presentación general de métodos de interpolación y predicción. Luego se muestran los estadísticos que se usan para comparar resultados de validación cruzada y los elementos básicos del análisis de componentes principales.

2.4.1. Interpolación determinística

Los conceptos de esta subsección son tomados de Waller y Gotway (2004), Bivand et al. (2013), Hengl (2007) y Soto y López (2005).

Bosque (1992), define la interpolación como “un procedimiento que permite calcular el valor de una variable en una posición del espacio (punto no muestral, con un valor estimado conocido) conociendo los valores de esa variable en otras posiciones del espacio (puntos muestrales, con valores reales)”.

La interpolación determinística se basa en métodos que usan una función predefinida para hacer las predicciones. El ejemplo más común es el de interpolación del inverso de la distancia (*DIP*). Este es un promedio ponderado de los valores vecinos al sitio de predicción. Se define como:

$$\hat{Z}(s_0) = \sum_{i=1}^n \lambda_i Z(s_i), \quad (2-3)$$

donde, $\hat{Z}(s_0)$ es la predicción en el sitio (s_0) de interés, λ_i es el peso o ponderación de la variable $Z(s_i)$ en un sitio cercano s_i . El predictor (interpolador) es insesgado, es decir $E(\hat{Z}(s_0)) = E(Z(s_0))$ si $\sum_{i=1}^n \lambda_i = 1$. En términos matriciales el predictor es dado por

$$\hat{Z}(s_0) = \lambda^T \mathbf{Z},$$

donde, $\lambda^T = (\lambda_1, \dots, \lambda_n)$ y $\mathbf{Z} = (Z(s_1), \dots, Z(s_n))^T$.

En el método del inverso de la distancia los pesos λ_i , $i = 1, \dots, n$, se calculan como:

$$\lambda_i = \frac{\frac{1}{d_{i0}}}{\sum_{j=1}^n \frac{1}{d_{j0}}},$$

con $d_{i0} = \|s_i - s_0\|$, $d_{j0} = \|s_j - s_0\|$, $s_i = (x_i, y_i)$, $s_j = (x_j, y_j)$, $s_0 = (x_0, y_0)$, (x_i, y_i) $i = 1, \dots, n$, son las coordenadas de ubicación geográfica.

Una vez calculadas las interpolaciones en varios puntos se pueden representar gráficamente mediante mapas de isolíneas. Estas son líneas que unen puntos de igual valor. Una revisión detallada sobre la construcción de mapas de isolíneas se da en Soto y López (2005).

Los interpoladores *DIP* son populares en muchas disciplinas ya que son relativamente simples conceptualmente y son de bajo costo computacional. Sin embargo, las superficies mapeadas tienden a tener picos planos y valles de fondo plano que dan un patrón de ojo de buey, producido por contornos concéntricos alrededor de puntos de datos que pueden restar valor a la interpretación visual del mapa. Además, dado que no existe un modelo estadístico subyacente, no es fácil medir la incertidumbre asociada con el valor interpolado.

2.4.2. Predicción espacial usando geoestadística

Los conceptos mencionados a continuación son tomados en su mayoría de Schabenberger y Gotway (2005) y Bowman y Azzalini (1997).

Aunque interpolación y predicción son términos usados indistintamente, hay diferencias entre estos. Cuando se hace predicción se asume que hay una medida de la incertidumbre que acarrea esta misma. Al usar interpolación (por ejemplo, con el método del inverso de la distancia) solo se obtiene un valor en el sitio de interés, pero se desconoce la variabilidad asociada al método, es decir este es un método determinístico. En un contexto más general pueden emplearse métodos de predicción espacial, por ejemplo, usando técnicas kriging, con las que además de la predicción se puede obtener una estimación de la varianza de predicción (consecuentemente del error estándar de predicción). Estas herramientas son enmarcadas en un campo de los procesos estocásticos y de la estadística espacial denominado geoestadística.

Un análisis geoestadístico conlleva generalmente tres etapas. Estudio de la estacionariedad, evaluación de correlación espacial e isotropía y predicción kriging. A continuación, se dan los elementos básicos de cada una de estas.

Estacionariedad

Un campo aleatorio $\{Z(s) : s \in D \subset R^d\}$, se denomina estacionario fuerte si la distribución espacial es invariante bajo la traslación de las coordenadas. Existen dos conceptos de estacionariedad de importancia:

1. Estacionariedad de segundo orden: un campo aleatorio definido en un dominio D contenido en R^d es estacionario de segundo orden si:
 - $E[Z(s)] = \mu$.
 - $V[Z(s)] = \sigma^2$
 - $COV[Z(s), Z(s+h)] = C(h)$

La función $C(h)$ se denomina función de covariancia del proceso espacial y juega un papel importante en el modelamiento estadístico de datos espaciales.

2. Estacionariedad Débil o Intrínseca: el proceso es estacionario débil si se cumple:
 - $E[Z(s)] = \mu$
 - $V[Z(s) - Z(s+h)] = 2\gamma(h)$,

es decir, si la media es constante y la varianza de los incrementos depende solo de la distancia h . La función $2\gamma(h)$ se denomina variograma y $\gamma(h)$ se llama semivariograma. Puede demostrarse que $C(h) = \sigma^2 - \gamma(h)$.

Estudio de tendencia espacial

Usualmente, para estudiar la tendencia de una variable aleatoria con respecto a la latitud o longitud, se hacen gráficos de regresión simple. También pueden usarse técnicas no paramétricas que no están sujetas al supuesto de linealidad. A continuación, se presentan conceptos básicos de regresión no paramétrica tomados de Wasserman (2007) y Zamora (2010).

La regresión no paramétrica a diferencia de la lineal tiene como objeto central estimar la función de regresión y no los coeficientes de regresión. Proporciona un procedimiento versátil para estudiar la relación entre dos variables. Suponga un conjunto de n pares de datos (x_i, y_i) . Se plantea la ecuación:

$$y_i = m(x_i) + \varepsilon_i, i = 1, \dots, n, \quad (2-4)$$

donde, y es la variable respuesta aleatoria, $m(x)$ una función desconocida y $\varepsilon_1, \dots, \varepsilon_n$ variables aleatorias independientes con $E(\varepsilon_i) = 0$ y $V(\varepsilon_i) = \sigma^2$, es decir asume homocedasticidad. Se pueden distinguir dos tipos de modelos de regresión atendiendo a los supuestos que se establecen sobre la función de regresión m :

- El modelo de regresión paramétrica que asume la función de regresión tiene forma predeterminada.
- El modelo de regresión no paramétrico supone hipótesis de suavidad en cuanto a continuidad y diferenciabilidad sobre la función de regresión m y no asume ninguna forma predefinida.

Uno de los métodos de regresión no paramétrica más usados es la regresión Kernel (núcleo). Los estimadores usados son medias ponderadas de las observaciones y_i . Se definen como aquellos estimadores $\hat{m}_h(x)$ de m en el punto x y se expresa de la forma:

$$\hat{m}_h(x) = \sum_{i=1}^n w_{hi}(x)y_i, \quad (2-5)$$

donde w_{hi} denota una sucesión de pesos que pueden depender del vector completo x , h es el parámetro de suavizado o ancho de banda. Dependiendo de la función peso se tienen distintos subtipos de la clase de estimadores. La función

$$w_{hi}(x) = \frac{K(x - x_i)}{\sum_{i=1}^n K(x - x_i)}, \quad (2-6)$$

determina el estimador Kernel de Nadaraya-Watson, que está dado por:

$$\hat{m}_h(x) = \frac{\sum_{i=1}^n K(x - x_i)y_i}{\sum_{j=1}^n K(x - x_j)}, \quad (2-7)$$

donde, $K(\cdot) = \frac{1}{h}K(\frac{\cdot}{h})$. La función K se denomina Kernel y determina que peso se asigna a cada observación dependiendo de su distancia a x . K es una función de densidad simétrica que cumple:

- $K(x) \geq 0$ para todo $x \in \mathbb{R}$, $\int K_e(x)dx = 1$, $\int xK_e(x)dx = 0$.
- $V(K) = \int x^2K(x)dx$.

Algunas funciones Kernel usadas en la práctica son:

- Uniforme: $K(x) = \frac{1}{2}I(x)$
- Gaussiano $K(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$
- Epanechnikov $K(x) = \frac{3}{4}1 - x^2I(x)$,

con,

$$I(x) = \begin{cases} 1 & \text{si } |x| \leq 1 \\ 0 & \text{si } |x| > 1, \end{cases}$$

Para profundizar en lo referente a K y a regresión no paramétrica se pueden revisar los capítulos 4 y 5 de Wasserman (2007), capítulo 4 de Eubank (1999) y capítulo 2 de Wand y Jones (1994).

Contraste de ausencia de efectos

Las definiciones presentadas a continuación se toman de Bowman y Azzalini (1997) y Delicado (2008).

Suponga el modelo de regresión no paramétrico homocedástico,

$$y_i = m(x_i) + \varepsilon_i, i = 1, \dots, n,$$

La hipótesis nula de ausencia de efectos se plantea como:

$$H_0 : m(x) \text{ es constante e igual a } \mu_y = E(y),$$

$$H_1 : m(x) \text{ no es constante.}$$

El estadístico de contraste es:

$$F = \frac{(SCR_0 - SCR_1)/(gl_0 - gl_1)}{SCR_1/gl_1},$$

donde, las sumas de cuadrados de los residuos (SCR) y los correspondientes grados de libertad (gl) están dados por:

$$SCR_0 = \sum_{i=1}^n (y_i - \bar{y})^2, \text{ y } gl_0 = n - 1,$$

$$SCR_1 = \sum_{i=1}^n (y_i - \hat{m}(x_i))^2,$$

con $\hat{m}(x)$ un estimador no paramétrico de $m(x)$, que tiene gl_1 grados de libertad efectivos ($gl_1 = \text{Traza}((I - S)^T(I - S))$, con S como la matriz de suavizado). La tabulación de la distribución de F bajo la hipótesis nula se hace como un test de permutaciones. Bajo H_0 , cualquier permutación de y_1, \dots, y_n es igualmente probable para x_1, \dots, x_n fijo. Se realiza el siguiente proceso:

1. Se permuta y_1, \dots, y_n aleatoriamente y obtiene y_{i1}, \dots, y_{in} . Se construye la muestra permutada (x_j, y_{ij}) , $j = 1, \dots, n$.
2. Se calcula el valor del estadístico F en la muestra permutada (F_p).
3. Se repiten B veces los pasos 1 y 2: F_p^1, \dots, F_p^B .
4. Se compara el valor de F observado en la muestra original (F_{ob}), con F_p^1, \dots, F_p^B , y se obtiene el valor p del test:

$$\text{valor } p = \frac{\#(F_p^b > F_{obs})}{B}$$

En el ítem 2 se estima la función de regresión para cada muestra permutada. Si se representan las B funciones estimadas simultáneamente se obtiene una banda de referencia del modelo sin efectos, que permite contrastar gráficamente de forma aproximada la hipótesis nula de ausencia de efectos. Si la función estimada en la muestra original sale fuera de la banda de referencia, se rechaza la hipótesis nula.

Correlación espacial

La siguiente etapa en el desarrollo de un análisis geoestadístico es la determinación de la dependencia espacial. Para identificarla se pueden usar diferentes funciones (semivariograma, covariograma y el correlograma). La más utilizada es el semivariograma. Las definiciones mostradas a continuación son tomadas de Schabenberger y Gotway (2005) y Waller y Gotway (2004).

El semivariograma se define como:

$$\frac{1}{2} \text{Var}[Z(s) - Z(s + h)] = \gamma(h). \quad (2-8)$$

Su estimador de momentos (semivariograma experimental) es:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} [Z(s_i) - Z(s_j)]^2, \quad (2-9)$$

donde, $N(h)$ es el número de parejas que se encuentran separadas por la distancia h , $Z(s_i)$ es la magnitud de la variable en un sitio (s_i), $Z(s_j)$ es otro valor muestral separado del anterior por una distancia h . Los parámetros de un semivariograma son el efecto pepita $\hat{\tau}^2$, la meseta $\hat{\sigma}^2$ y el rango $\hat{\phi}$. Se representan en el gráfico (2-2).

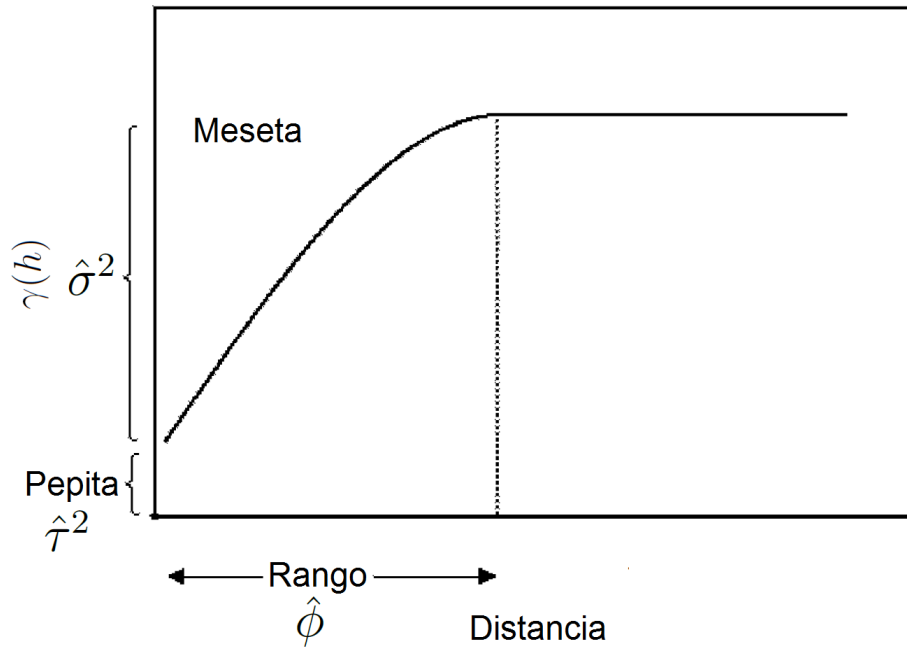


Figura 2-2: Representación de un semivariograma típico acotado, con parámetros: pepita $\hat{\tau}^2$, meseta $\hat{\sigma}^2$ y rango $\hat{\phi}$ (Waller y Gotway, 2004).

Isotropía y Anisotropía

Cressie (1993), describe una clase de proceso anisotrópico como aquel en el que la dependencia espacial es una función tanto de la magnitud como de la dirección, es decir, que hay diferencias entre los semivariogramas (2-9), en distintas direcciones (Oliver y Webster, 2015).

El estudio de anisotropía de un proceso espacial se puede realizar a través de la construcción de semivariogramas direccionales (diferentes ángulos θ). Si el semivariograma en los diferentes θ es similar, se puede decir que el proceso es isotrópico. En la práctica se pueden elegir cuatro direcciones principales a 0, 45, 90 y 135 grados (Figura 2-4)

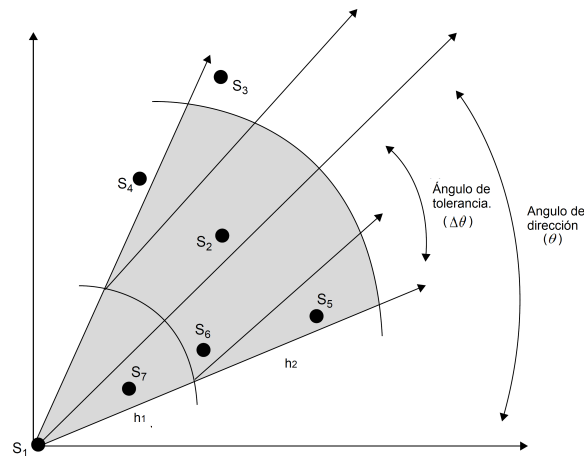


Figura 2-3: Representación de regiones de tolerancia a distancias h_1 y h_2 a un ángulo de dirección θ (Waller y Gotway, 2004).

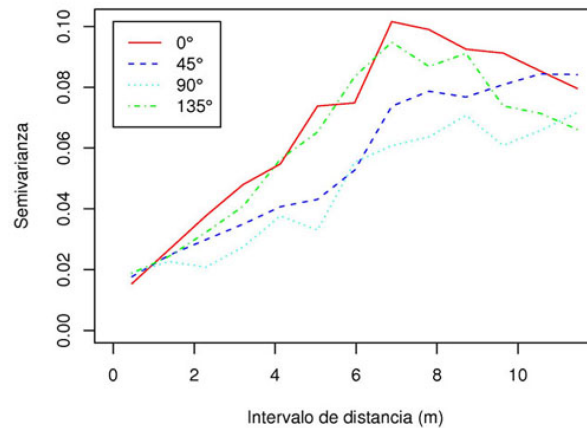


Figura 2-4: Semivariogramas a diferentes ángulos de dirección θ (0, 45, 90 y 135 grados), usando datos de contenido de materia orgánica en el suelo de una parcela de 20 x 20 metros (Gallardo, 2006).

Modelos teóricos de semivarianza

Para describir el comportamiento del semivariograma experimental (ecuación 2-9), se utilizan modelos teóricos que generalizan la correlación espacial. Estos se pueden dividir en no acotados (lineal, logarítmico, potencial) y acotados (esférico, exponencial, Gaussiano). Los del segundo grupo garantizan que la covarianza de los incrementos es finita, y son ampliamente usados en la práctica. Las expresiones de los modelos esférico y exponencial son (Waller y Gotway, 2004):

- Modelo Esférico:

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \tau^2 + \sigma^2 \left[\frac{3}{2} \frac{h}{\phi} - \frac{1}{2} \left(\frac{h}{\phi} \right)^3 \right], & 0 < h \leq \phi \\ \tau^2 + \sigma^2, & h > \phi, \end{cases} \quad (2-10)$$

donde, $\tau^2 \geq 0$, $\sigma^2 \geq 0$, $\phi > 0$. El semivariograma esférico se puede aproximar por una recta cerca del origen (Figura 2-5). El parámetro τ^2 mide el efecto pepita, σ^2 es parte de la meseta ($\tau^2 + \sigma^2$ es la meseta) y ϕ el rango.

- Modelo Exponencial:

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \tau^2 + \sigma^2 [1 - e^{-\frac{h}{\phi}}], & h > 0, \end{cases} \quad (2-11)$$

donde, $\tau^2 \geq 0$, $\sigma^2 \geq 0$, $\phi > 0$. El semivariograma exponencial (dependiendo de los parámetros) se eleva más lentamente desde el origen que el esférico (Figura 2-5). τ^2 es el efecto pepita y σ^2 es parte de la meseta ($\tau^2 + \sigma^2$ es la meseta). Este modelo se acerca a la meseta asintóticamente, por lo que el rango no es ϕ , ya que $\gamma(\phi) \neq \tau^2 + \sigma^2$.

Warrick (2002) especifica los parámetros de los semivariogramas de la siguiente forma. El τ^2 es una medida de la cantidad de varianza debida a errores en el muestreo, la medición y otras fuentes de varianza no explicadas. La suma de los parámetros $\tau^2 + \sigma^2$ es teóricamente igual a la varianza de la población muestreada a grandes distancias de separación si los datos no tienen tendencia. Si el τ^2 es aproximadamente igual a la varianza de la muestra, indica que

la propiedad muestreada tiene muy poca estructura espacial o varía aleatoriamente. El ϕ en los modelos lineales y esféricos corresponde a la distancia en que las variables son espacialmente independientes. Cressie (1993) menciona que los parámetros anteriormente descritos pueden ser estimados por una combinación de momentos y regresión no lineal o por máxima verosimilitud.

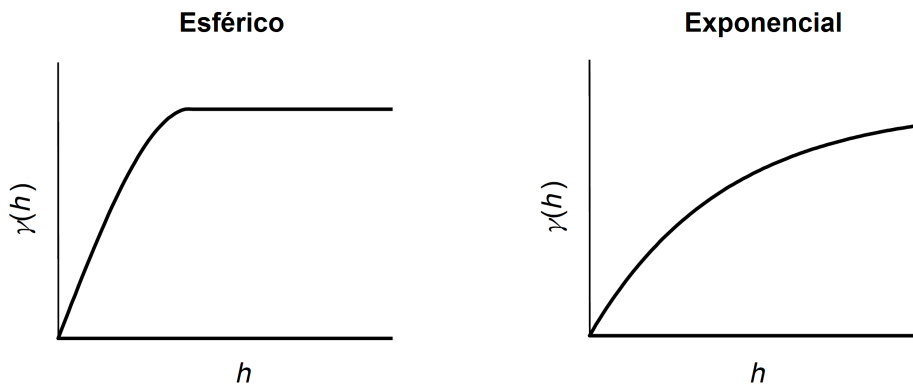


Figura 2-5: Modelo de semivariograma esférico (izquierda) y exponencial (derecha) (Waller y Gotway, 2004).

Para establecer la proporción de la variabilidad explicada por el modelo teórico de semivarianza, se puede calcular el coeficiente de determinación R^2 (Weisberg, 2014):

$$R^2 = 1 - \frac{SCE}{SCT}, \quad (2-12)$$

donde, SCE es la suma de cuadrados del error y SCT es la suma de cuadrados totales del modelo usado. Montgomery et al. (2006) indica que a mayor valor de R^2 el modelo presenta un mejor ajuste.

Predicción Espacial por Kriging

A diferencia de la interpolación determinística (sección 2.4.1) se asume un modelo estadístico. Los predictores tienen errores estándar que cuantifican la incertidumbre asociada con los datos predichos. Los siguientes conceptos son tomados de Waller y Gotway (2004) y Warrick

(2002).

La predicción espacial kriging, fue desarrollada para la industria minera por Matheron (1963) y se implementó por primera vez en esa industria por Krige (1966). Se usó en la ciencia del suelo por Burgess y Webster (1980) en Europa y en los Estados Unidos por Vauclin et al. (1982).

El uso de esta metodología se ha extendido y existen diferentes tipos de kriging, que difieren según las suposiciones subyacentes y los objetivos analíticos. A continuación, se hace mención de los dos tipos más comunes (Waller y Gotway, 2004):

1.) Kriging ordinario.

Sea $Z(\cdot)$ estacionario (ver página 19) con semivarianza $\gamma(h)$. Dado el vector $\mathbf{Z} = [Z(s_1), \dots, Z(s_N)]$, se desea predecir $Z(\cdot)$ en una ubicación no observada (s_0), $s_0 \in D$. El predictor kriging ordinario (KO) se define como:

$$\hat{Z}_{KO}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i). \quad (2-13)$$

Los pesos λ_i se obtienen de tal manera que el predictor (2-13) sea insesgado y de mínima varianza, en otras palabras que sea el mejor predictor lineal insesgado ($MPLI$). El predictor es insesgado si

$$E[\hat{Z}_{KO}(s_0)] = E[Z(s_0)] = \mu.$$

Para cumplir esta restricción se debe cumplir que $\sum_{i=1}^N \lambda_i = 1$ (ver Waller y Gotway (2004)). El predictor (2-13) es de mínima varianza si $E[\hat{Z}(s_0) - Z(s_0)]^2$ es mínima.

A continuación se deduce el sistema de ecuaciones que permite estimar los parámetros λ_i del predictor (2-13) .

Debe minimizarse $E[\hat{Z}(s_0) - Z(s_0)]^2$ sujeto a la restricción $\sum_{i=1}^N \lambda_i = 1$, es decir, la función objetivo es:

$$E \left[\left(\sum_{i=1}^N \lambda_i Z(s_i) - Z(s_0) \right)^2 \right] - 2m \left(\sum_{i=1}^N \lambda_i - 1 \right).$$

Se tiene que:

$$\left[\sum_{i=1}^N \lambda_i Z(s_i) - Z(s_0) \right]^2 = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j [Z(s_i) - Z(s_j)]^2 + \sum_{i=1}^N \lambda_i [Z(s_0) - Z(s_i)]^2. \quad (2-14)$$

Tomando la esperanza a ambos lados de la ecuación 2-14 se obtiene:

$$-\sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(s_i - s_j) + 2 \sum_{i=1}^N \lambda_i \gamma(s_0 - s_i) - 2m \left(\sum_{i=1}^N \lambda_i - 1 \right). \quad (2-15)$$

Al minimizar (2-15), se tiene un sistema conocido como ecuaciones del kriging ordinario,

$$\sum_{j=1}^N \lambda_j \gamma(s_i - s_j) + m = \gamma(s_0 - s_i), \quad i = 1, \dots, N.$$

$$\sum_{j=1}^N \lambda_j = 1. \quad (2-16)$$

Resolviendo las ecuaciones para $\lambda_1, \dots, \lambda_N$ y m (multiplicador de Lagrange), se obtienen los pesos del predictor kriging ordinario. Se puede escribir el sistema de ecuaciones (2-16) en forma de matricial como:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_N \\ m \end{bmatrix} = \begin{bmatrix} \gamma(s_1 - s_1) & \dots & \gamma(s_1 - s_N) & 1 \\ \gamma(s_2 - s_1) & \dots & \gamma(s_2 - s_N) & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \gamma(s_N - s_1) & \dots & \gamma(s_N - s_N) & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(s_0 - s_1) \\ \gamma(s_0 - s_2) \\ \cdot \\ \cdot \\ \cdot \\ \gamma(s_0 - s_N) \\ 1 \end{bmatrix}$$

donde, $\gamma(s_i - s_j) = \hat{\gamma}(h)$, $h = \|s_i - s_j\|$ y $\gamma(s_i - s_0) = \hat{\gamma}(h)$, $h = \|s_i - s_0\|$.

La varianza de predicción del kriging ordinario está dada por:

$$\begin{aligned}
 \sigma_k^2(s_0) &= \boldsymbol{\lambda}'_0 \boldsymbol{\gamma}_0 \\
 &= \sum_{i=1}^N \lambda_i \gamma(s_0 - s_i) + m \\
 &= 2 \sum_{i=1}^N \lambda_i \gamma(s_0 - s_i) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(s_i - s_j). \tag{2-17}
 \end{aligned}$$

2.) Kriging universal.

El predictor de kriging universal se define como:

$$\hat{Z}(s_0)_{ku} = \sum_{i=1}^N \lambda_i Z(s_i), \tag{2-18}$$

A diferencia del caso anterior, se considera que el proceso no es estacionario y puede ajustarse el modelo de regresión:

$$\mathbf{Z}(\mathbf{s}) = X(s)\boldsymbol{\beta} + \epsilon, \tag{2-19}$$

donde, $\mathbf{Z}(\mathbf{s}) = [Z(s_1), Z(s_2), \dots, Z(s_N)]'$ corresponde a la variable en las ubicaciones s_1, \dots, s_N , $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})'$ es el vector de parámetros de regresión y $X(s)$ es una matriz de covariables:

Cuando se asume un modelo de regresión lineal (2-19), el semivariograma es $1/2V[\epsilon(s_i) - \epsilon(s_j)] = \gamma(s_i - s_j), \forall i, j$. El predictor lineal es de la forma $\hat{Z}(s_0) = \sum_{i=1}^N \lambda_i Z(s_i)$. Para obtener las estimaciones de λ_i se minimiza la varianza del error de predicción $E[\hat{Z}(s_0) - Z(s_0)]^2$. Sujeto a las restricciones de insesgadez, es decir se minimiza:

$$\left[\sum_{i=1}^N \lambda_i Z(s_i) - Z(s_0) \right]^2 - 2 \sum_{j=0}^{p-1} m_j \left\{ \sum_{i=1}^N \lambda_i x_j(s_i) - x_j(s_0) \right\}, \tag{2-20}$$

con p multiplicadores de Lagrange m_0, \dots, m_{p-1} . Usando el modelo lineal se obtiene:

$$\left[\sum_{i=1}^N \lambda_i Z(s_i) - Z(s_0) \right]^2 = \left[\boldsymbol{\lambda}' X \boldsymbol{\beta} + \sum_{i=1}^N \lambda_i \epsilon(s_i) - \boldsymbol{x}'(s_0) \boldsymbol{\beta} - \epsilon(s_0) \right]^2$$

$$\begin{aligned}
&= \left[\sum_{i=1}^N \lambda_i \epsilon(s_i) - \epsilon(s_0) \right]^2 \\
&= - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j [\epsilon(s_i) - \epsilon(s_j)]^2 / 2 + 2 \sum_{i=1}^N \lambda_i [\epsilon(s_0) - \epsilon(s_i)]^2 / 2
\end{aligned}$$

Entonces 2-20 es:

$$= - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(s_i - s_j) + 2 \sum_{i=1}^N \lambda_i \gamma(s_0 - s_i) - 2 \sum_{j=0}^{p-1} m_j \left[\sum_{i=1}^N \lambda_i x_j(s_i) - x_j(s_0) \right].$$

Matricialmente el sistema de ecuaciones del kriging universal en función de coordenadas x, y se expresa como:

$$\begin{bmatrix} \lambda_1 \\ \cdot \\ \cdot \\ \cdot \\ \lambda_N \\ m_0 \\ m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} \gamma(s_1 - s_1) & \dots & \gamma(s_1 - s_N) & x_1 & y_1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \gamma(s_N - s_1) & \dots & \gamma(s_N - s_N) & x_N & y_N & 1 \\ 1 & \dots & 1 & 0 & 0 & 0 \\ x_1 & \dots & x_N & 0 & 0 & 0 \\ y_1 & \dots & y_N & 0 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(s_0 - s_1) \\ \cdot \\ \cdot \\ \cdot \\ \gamma(s_0 - s_N) \\ 1 \\ x_0 \\ y_0 \end{bmatrix}$$

donde, $\gamma(s_i - s_j) = \hat{\gamma}(h)$, $h = \|s_i - s_j\|$ y $\gamma(s_i - s_0) = \hat{\gamma}(h)$, $h = \|s_i - s_0\|$.

La varianza del kriging universal es:

$$\begin{aligned}
\sigma_{uk}^2(s_0) &= \lambda'_{ku} \gamma_u & (2-21) \\
&= 2 \sum_{i=1}^N \lambda_i \gamma(s_0 - s_i) - \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \gamma(s_i - s_j).
\end{aligned}$$

Warrick (2002) expresa que los métodos de kriging son superiores al determinístico cuando hay dependencia espacial. Se pueden utilizar técnicas como la validación cruzada para calcular algunos estadísticos que permiten comparar los métodos de predicción.

Para profundizar en el tema de predicción espacial por kriging se recomiendan los capítulos 3 de Cressie (1993) y 5 de Schabenberger y Gotway (2005).

2.4.3. Diagnóstico de predicción e interpolación

A continuación, se contextualizará sobre el concepto de validación cruzada y algunos estadísticos de comparación (James et al., 2013).

La validación cruzada es una técnica que consiste en excluir la observación de uno de los n puntos muestrales y usando los $(n - 1)$ valores restantes y el método de interpolación o predicción considerados, se predice el valor de la variable de estudio en la ubicación del punto que se excluyó. Si el modelo elegido es apropiado, la diferencia entre los valores observados y predichos debe ser pequeña. Este procedimiento se realiza en forma secuencial con cada uno de los puntos muestrales y así se obtiene un conjunto de n errores de predicción.

Supongase que se ha observado el valor del proceso $Z(s)$ sobre un conjunto de localizaciones s_1, \dots, s_n . Sea $\hat{Z}(s_i)$ el valor predicho y $\hat{\sigma}$ el error estándar de la predicción para la localización s_i obtenida a partir de las observaciones $s_1, \dots, s_{i-1}, \dots, s_n$. El valor predicho a partir de la validación cruzada es $\hat{Z}_{[i]}(s_i)$ y $\hat{\sigma}_{[i]}(s_i)$ la predicción para la desviación estándar en la localización s_i .

Los estadísticos usados para comparar los resultados de la predicción o interpolación son:

- Error medio de predicción (*EMP*).

$$EMP = \frac{\sum_{i=1}^n (\hat{Z}_{[i]}(s_i) - Z(s_i))}{n}, \quad (2-22)$$

- Raíz del error cuadrado medio de predicción (*RECM*).

$$RECM = \sqrt{\frac{\sum_{i=1}^n (\hat{Z}_{[i]}(s_i) - Z(s_i))^2}{n}}, \quad (2-23)$$

- Coeficiente de determinación (R^2).

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (\hat{Z}_{[i]}(s_i) - Z(s_i))}{\sum_{i=1}^n (Z(s_i) - \bar{Z})^2} \right) * 100, \quad (2-24)$$

donde $\bar{Z} = \sum_{j=1}^n Z(s_j)/n$.

2.4.4. Geoestadística multivariada a través de análisis de componentes principales.

Una forma de hacer geoestadística multivariada es usar análisis de componentes principales como reducción de variables en la fase inicial y posteriormente utilizar la teoría geoestadística vista en la sección (2.4.2). Las referencias consideradas para describir esta sección son tomadas de Aldás y Jiménez (2017), Pla (1986) y Srivastava (2002).

El análisis de componentes principales (*ACP*) es uno de los métodos más difundidos, que permite la caracterización de un conjunto de datos multivariados. Los objetivos de un *ACP* son: generar nuevas variables (combinaciones lineales) que puedan expresar la información contenida en el conjunto original de datos, reducir la dimensionalidad del problema de estudio y eliminar, cuando sea posible, algunas de las variables originales si éstas aportan poca información.

El *ACP* genera nuevas variables que se denominan *componentes principales (CP)*, que tienen algunas características estadísticas deseables:

- Los *CP* no están correlacionados y si, además, se supone multinormalidad son independientes.
- Cada *CP* sintetiza la máxima variabilidad residual contenida en los datos.

A continuación, se presenta el método para obtener el primer componente principal.

Obtención de la primera componente

Se considera una muestra de tamaño n , con p variables (X_1, X_2, \dots, X_p) y las observaciones están expresadas en desviaciones respecto a la media (centradas).

La primera componente, se expresa como la combinación lineal de las variables originales:

$$Z_{1i} = u_{11}X_{1i} + u_{12}X_{2i} + \dots + u_{1p}X_{pi} \quad (2-25)$$

Para el conjunto de las n observaciones muestrales la ecuación (2-25) se puede expresar matricialmente como:

$$\begin{bmatrix} Z_{11} \\ Z_{12} \\ \cdot \\ \cdot \\ Z_{1n} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{p1} \\ X_{12} & X_{22} & \dots & X_{p2} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ X_{1n} & X_{2n} & \dots & X_{pn} \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ \cdot \\ \cdot \\ u_{1p} \end{bmatrix}$$

o en notación matricial compacta:

$$\mathbf{z}_1 = \mathbf{X}\mathbf{u}_1 \quad (2-26)$$

La primera componente se obtiene de forma que su varianza sea máxima, sujeta a la restricción de que la suma de los pesos (u_{1j}) al cuadrado sea igual a 1. La varianza del primer componente, teniendo en cuenta que su media es 0, es:

$$Var(Z_1) = \frac{\sum_{i=1}^n Z_{1i}^2}{n} = \frac{1}{n} \mathbf{z}'_1 \mathbf{z}_1 = \frac{1}{n} \mathbf{u}'_1 \mathbf{X}' \mathbf{X} \mathbf{u}_1 = \mathbf{u}'_1 \left[\frac{1}{n} \mathbf{X}' \mathbf{X} \right] \mathbf{u}_1 \quad (2-27)$$

Sea \mathbf{V} la matriz de covarianza:

$$Var(Z_1) = \mathbf{u}'_1 \mathbf{V} \mathbf{u}_1 \quad (2-28)$$

La restricción de que la suma de los cuadrados de los pesos es igual a uno se expresa como:

$$\sum_{j=1}^p u_{1j}^2 = \mathbf{u}'_1 \mathbf{u}_1 = 1 \quad (2-29)$$

Incorporando la restricción (2-29) se forma el lagrangiano:

$$L = \mathbf{u}'_1 \mathbf{V} \mathbf{u}_1 - \lambda(\mathbf{u}_1 \mathbf{u}_1 - 1) \quad (2-30)$$

Para maximizar el valor del lagrangiano se deriva respecto a \mathbf{u}_1 y se iguala a 0:

$$\frac{\partial L}{\partial \mathbf{u}_1} = 2\mathbf{V}\mathbf{u}_1 - 2\lambda\mathbf{u}_1 = 0 \quad (2-31)$$

es decir:

$$(\mathbf{V} - \lambda \mathbf{I})\mathbf{u}_1 = 0 \quad (2-32)$$

Al resolver la ecuación $|\mathbf{V} - \lambda \mathbf{I}| = 0$, se obtienen p autovalores. Si se toma la raíz característica mayor (λ_1), se halla el autovector asociado a la misma \mathbf{u}_1 , aplicando la regla de normalización dada (2-29). El vector de ponderaciones que se aplica a las variables iniciales para obtener la primera componente principal es el autovector asociado al mayor autovalor de la matriz \mathbf{V}

Para determinar los componentes restantes, se puede expresar la ecuación (2-26) en forma genérica como:

$$\mathbf{z}_h = \mathbf{X}\mathbf{u}_h \quad (2-33)$$

Con $\mathbf{u}'_h \mathbf{u}_h = 1$, se imponen otras restricciones:

$$\mathbf{u}'_h \mathbf{u}_1 = \mathbf{u}'_h \mathbf{u}_2 = \dots = \mathbf{u}'_h \mathbf{u}_{h-1} = 0 \quad (2-34)$$

Las p componentes principales que se pueden calcular son una combinación lineal de las variables originales, en la que los coeficientes de ponderación son los correspondientes vectores característicos asociados a la matriz \mathbf{V} .

Varianzas de las componentes

La varianza de la componente h -ésima coincide con su autovalor, es decir:

$$\text{Var}(Z_h) = \mathbf{u}'_h \mathbf{V} \mathbf{u}_h = \lambda_h \quad (2-35)$$

La suma de la variabilidad de las variables originales es igual a la traza de la matriz \mathbf{V} . Se puede expresar teniendo en cuenta que:

$$\frac{\lambda_h}{\text{traza} \mathbf{V}} = \frac{\lambda_h}{\sum_{h=1}^p \lambda_h} \quad (2-36)$$

Correlación entre las componentes principales y las variables originales

Los vectores muestrales del componente Z_h y la variable X_j son:

$$\mathbf{x}_j = \begin{bmatrix} X_{j1} \\ X_{j2} \\ \cdot \\ \cdot \\ X_{jn} \end{bmatrix} \quad \mathbf{z}_h = \begin{bmatrix} Z_{h1} \\ Z_{h2} \\ \cdot \\ \cdot \\ Z_{hn} \end{bmatrix}$$

La covarianza muestral entre X_j y Z_h es:

$$\text{cov}(X_j, Z_h) = \frac{1}{n} \mathbf{x}'_j \mathbf{z}_h \quad (2-37)$$

El vector \mathbf{x}_j se puede expresar en función de matriz \mathbf{X} , utilizando el vector de orden p ($\boldsymbol{\delta}$):

$$\mathbf{x}'_j = \boldsymbol{\delta}' \mathbf{X} = [0 \dots, 1 \dots, 0] \begin{bmatrix} X_{11} & \dots & X_{li} & \dots & X_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ X_{j1} & \dots & X_{ji} & \dots & X_{jn} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ X_{p1} & \dots & X_{pi} & \dots & X_{pn} \end{bmatrix} \quad (2-38)$$

La covarianza (2-37) puede expresarse como:

$$\text{cov}(X_j, Z_h) = \frac{1}{n} \boldsymbol{\delta}' \mathbf{X}' \mathbf{X} \mathbf{u}_h = \boldsymbol{\delta}' \mathbf{V} \mathbf{u}_h = \boldsymbol{\delta}' \lambda_h \mathbf{u}_h = \lambda_h \boldsymbol{\delta}' \mathbf{u}_h = \lambda_h u_{hj} \quad (2-39)$$

La correlación entre la variable X_j y la componente Z_h es:

$$l_{jh} = \frac{\text{cov}(X_j, Z_h)}{\sqrt{\text{Var}(X_j)} \sqrt{\text{Var}(Z_h)}} = \frac{\lambda_h u_{hj}}{\sqrt{\text{Var}(X_j)} \sqrt{\lambda_h}} = \frac{u_{hj}}{\sqrt{\text{Var}(X_j)}} \sqrt{\lambda_h} \quad (2-40)$$

Una vez obtenidos los CP se puede usar un análisis geoestadístico con ellos, para interpretar de manera simultánea la variación espacial de las características consideradas. Para profundizar en el tema de análisis de componentes principales se sugiere el capítulo 11 de Srivastava (2002) y Anderson (2003).

3 Área de estudio e información disponible

En este capítulo en primer lugar, se presenta la descripción del área de estudio tomando como referencia lo mencionado por Insuasty et al. (2003), Osorio (2007) y Rodríguez et al. (2004). Posteriormente, se explica la procedencia de los datos utilizados en el análisis.

3.1. Hoya del Río Suárez (HRS)

La *HRS* está ubicada en la cuenca media del río del mismo nombre, en alturas comprendidas entre los 1200 y 1900 metros sobre el nivel del mar. La región tiene un brillo solar anual entre 1900 a 2000 horas equivalentes a una luminosidad de 5,7 horas/día, temperaturas medias entre los 18 y los 23 grados centígrados, con variaciones de temperatura entre el día y la noche que pueden estar entre los 10 y 12 grados, la temperatura máxima absoluta puede llegar hasta los 28,6 grados centígrados y la mínima absoluta hasta los 8,8 grados. La precipitación varía entre los 1200 a los 2700 milímetros al año, los cuales están repartidos en dos períodos lluviosos, donde el primero se da desde mediados de marzo a finales de junio y el segundo, desde finales de agosto a finales de diciembre, que representan 221 días con una buena distribución de lluvias. La humedad relativa del ambiente varía entre 65 y 85 %.

Esta zona incluye los municipios de Chitaraque, Moniquirá, Togüí, San José de Pare, Santana en el departamento de Boyacá y Barbosa, Chipatá, Güepsa, San Benito, Suaita y Vélez en el departamento de Santander (Figura 3-1). La topografía es de montaña ondulada y quebrada, con pendientes catalogadas de medianas a altas.

Históricamente en la *HRS* el cultivo de la caña panelera se inició en el año de 1939, donde los campesinos derivaban su sustento de las ventas de panela elaborada de manera rustica en viejos trapiches. En la actualidad es la actividad más importante que se desarrolla en la región junto con el cultivo y procesamiento de la guayaba. Las explotaciones son de tamaño mediano catalogadas como tecnificadas de forma empresarial, debido a la presencia

de diferentes agremiaciones o asociaciones que han contribuido al desarrollo de la región con diferentes aportes en tecnología. Estas características unidas a las condiciones geográficas y ambientales hacen que sea reconocida a nivel nacional como el mayor productor con un rendimiento medio de 120 toneladas de caña por hectárea (Rodríguez et al., 2004).



Figura 3-1: Derecha. El recuadro rojo enmarca la zona de estudio en Colombia. Izquierda. Municipios en color naranja son los sitios de los departamentos de Santander (Barbosa, Chipatá, Güepa, San Benito, Suaita, Vélez) y Boyacá (Chitaraque, Moniquirá, San José de Pare, Santana y Togüí) donde se midieron las variables.

3.2. Base de datos

Para realizar los análisis se parte de la información recopilada por Corpoica (Agrosavia) en el proyecto “Recomendaciones técnicas preliminares de manejo de suelos en ladera para el sistema de producción de caña panelera en la *HRS*”, el cual pertenece al macroproyecto “Generación de recomendaciones agronómicas y nuevas tecnologías en la producción de panela en Colombia” (CORPOICA, 2016). Los datos se tomaron a una profundidad de 20 centímetros, siguiendo la forma de una de rejilla rectangular, con una separación entre pun-

tos de muestreo de 700 metros, tanto a lo largo como a lo ancho; es decir, un punto de muestreo cada 49 hectáreas. Se midieron en total 932 puntos georreferenciados abarcando más de 40000 hectáreas (Figura 3-2). Los puntos de la rejilla se ubicaron en el terreno con un navegador Garmin Etrex 30 (Garmin, 2011).

Se tomaron 5 variables físicas, las cuales están directamente relacionadas con la erodabilidad de los suelos. El porcentaje de arenas (A), diámetro medio ponderado (DMP), capacidad de retención de agua disponible ($CRAD$), densidad aparente (DA) y real (DR), estas variables se definieron en la sección (2.2).

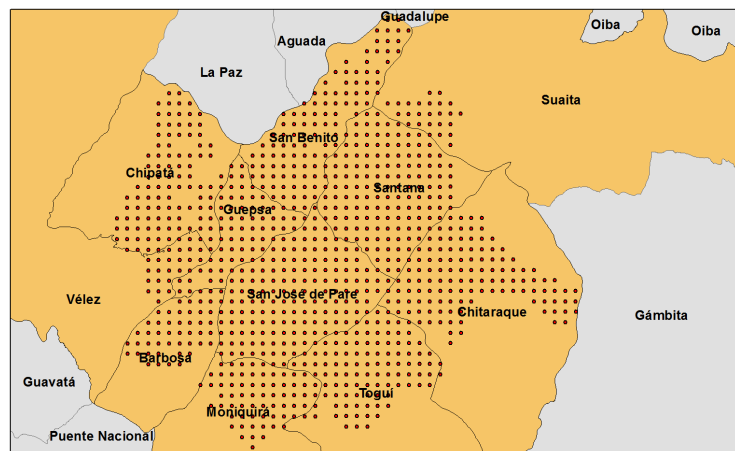


Figura 3-2: Esquema de muestreo. Los puntos rojos corresponden a los lugares de medición en cada uno de los municipios.

4 Metodología

En este capítulo se detallan los diferentes métodos utilizados para establecer zonas de erodabilidad, que se comparan para decidir el mejor, de acuerdo con consideraciones estadísticas (ajuste, sencillez) y comportamientos de las variables desde el punto de vista agronómico (limitantes del suelo para la producción agrícola).

En principio, se define la forma en que se manejaron las variables a espacializar, en segundo lugar, se menciona como se realizan las interpolaciones y predicciones planteadas. Finalmente se hace la evaluación de las alternativas resultantes. Los análisis y gráficos para las delimitaciones de áreas se realizaron con el software estadístico R (R Core Team, 2016).

4.1. Definición del método de manejo de las variables

Se plantearon tres formas de análisis de la base de datos (sección 3.2) referente a propiedades físicas del suelo: por variables individuales, conjuntas (selección de componentes principales del análisis multivariado) y el cálculo del índice de erodabilidad (K), cuyas definiciones se encuentran en las secciones (2.4.4 y 2.2.5).

4.1.1. Análisis individual de las variables

Se analiza de forma independiente cada una de las variables seleccionadas (A , DMP , DA , DR , $CRAD$). Al final se selecciona la variable que mejor describe la erodabilidad. Las unidades empleadas fueron para A (%), DMP (mm), DA (g/cm^3), DR (g/cm^3) y $CRAD$ (%).

4.1.2. Análisis conjunto por componentes principales

Se estudia en conjunto las variables seleccionadas, para determinar si es conveniente resumirlas en uno o dos componentes principales, que se constituyen en nuevas variables que se puedan espacializar. Las características A , DMP , DA , DR , $CRAD$ se cruzan por pares para

determinar los coeficientes de la correlación lineal de Pearson, usando la función *cor.test* de la librería *stats* (R Core Team, 2016). Los componentes principales revisados en la sección (2.4.4) se generaron con la función *princomp* de la librería *stats* (R Core Team, 2016). La selección del número de componentes se determinó teniendo en cuenta que el porcentaje de varianza explicada fuera mayor al 60 %.

4.1.3. Análisis del índice de erodabilidad K

Se trabaja el modelo de regresión propuesto por Ramírez et al. (2009), para estimar el valor de K , en donde se tiene en cuenta el contenido de A , DMP y DA . La fórmula para calcular K se encuentra en la sección (2.2.5).

En cada caso se realizaron de manera previa estudios descriptivos. Se calcularon medidas de tendencia central (media y mediana) y variabilidad (varianza y coeficiente de variación). Para examinar la distribución de los datos e identificar valores extremos, se construyeron diagramas de cajas y bigotes (mediante la función *boxplot* de la librería *gstat* (Pebesma, 2004)).

4.2. Evaluación de los métodos de interpolación y predicción espacial

Con las formas de manejo de las variables establecidos en la sección anterior (4.1), se procede a predecir los valores de las variables en sitios no muestreados. Se emplearon los métodos de interpolación y predicción descritos en las secciones (2.4.1 y 2.4.2).

4.2.1. Establecimiento de la rejilla para interpolación y predicción

Se usó una rejilla rectangular de 150 metros de largo por 80 metros de ancho generando 3850 puntos. Cada punto representa un área de 1,2 hectáreas.

4.2.2. Interpolación determinística

Se utilizó el método del inverso de la distancia dado en la sección (2-3). Las predicciones se realizaron mediante la función *idw* de la librería *gstat* (Pebesma, 2004). Se aplicó con las

variables definidas en la sección (4.1).

4.2.3. Predicción estocástica

Las etapas del análisis geoestadístico descritas en la sección (2.4.2) utilizadas en el desarrollo de este estudio se detallan a continuación:

Evaluación de la estacionariedad y correlación espacial de las variables

En primer lugar, se observa si hay tendencia espacial de las variables en la zona evaluada. La estacionariedad de las variables con respecto a su localización geográfica en cada zonificación se realizó con la función *sm.regression* de la librería *sm* (Bowman y Azzalini, 2014). Con este método se crea la estimación de un modelo de regresión no paramétrico Kernel (2-4) y un gráfico de dispersión en dos dimensiones, con ejes Y (variable evaluada), X (latitud o longitud) y una banda de confianza de no efecto de la superficie de regresión estimada (Bowman y Azzalini, 2014) (sección 2.4.2). Para evaluar la anisotropía de los procesos espaciales estudiados (sección 2.4.2), se construyeron para cada variable semivariogramas direccionales en 4 orientaciones 0, 45, 90 y 135 grados, usando la función *variog4* de la librería *geoR* (Ribeiro y Diggle, 2016).

Luego se construyeron los semivariogramas para las variables. Se calculó el semivariograma experimental descrito en la sección (2.4.2) con cada variable, utilizando la función *variogram* de la librería *gstat* (Pebesma, 2004) y se ajustó un modelo teórico (2-10, 2-11), a través de la función *fit.variogram* de la librería *gstat* (Pebesma, 2004). También, se estimó el estadístico de ajuste (2-12), para seleccionar el mejor modelo.

Predicción por Kriging

Si el resultado del análisis indica que el proceso estudiado es estacionario se usa predicción por kriging ordinario o universal si hay tendencia. Para esto se emplea la función *krige* de la librería *gstat* (Pebesma, 2004) para los predictores de kriging ordinario y kriging universal presentados en las secciones (2-13 y 2-18), con sus respectivas varianzas revisadas en la sección (2-17).

4.3. Evaluación de las alternativas resultantes

Con las interpolaciones y predicciones de las variables de estudio definidas en la sección (3.2), se grafican los mapas para las diferentes zonificaciones y por último se comparan para seleccionar la mejor alternativa utilizando los criterios revisados en la sección (2.4.3).

4.3.1. Zonificación para las variables resultantes

Usando los valores estimados por los métodos de interpolación y predicción se elaboran los mapas de isolíneas usando la función *splot* de la librería *sp* (Pebesma y Bivand, 2005) para cada método de análisis planteado en la sección (4.1). La variable utilizada para zonificar se dividió en categorías para representar en el mapa zonas susceptibles al deterioro de la calidad del suelo si no se realizan prácticas adecuadas de conservación. Dichas categorías están representadas con diferentes colores y los valores correspondientes se muestran en la tabla (4-1).

Tabla 4-1: Valores para definir condiciones favorables y susceptibles de las variables porcentaje de arenas (A), diámetro medio ponderado (DMP), densidad aparente (DA), densidad real (DR) y capacidad de retención de agua disponible ($CRAD$).

Variables	Condiciones favorables del suelo	Condiciones limitantes del suelo
A (%)	< 50	> 50
DMP (mm)	> 0,5	< 0,5
DA (g/cm ³)	< 1,4	> 1,4
DR (g/cm ³)	< 2,6	> 2,6
$CRAD$ (%)	> 5	< 5

Información de la tabla fue adaptada de Torres et al. (2017); Gómez y Osorio (2016); Anshau et al. (2009); Buckman y Brady (1966); Montenegro y Malagón (1990).

Luego, se graficaron los valores estimados usando la función *splot* de la librería *sp* (Pebesma y Bivand, 2005).

4.3.2. Selección final de métodos de zonificación

Para determinar cuál de los esquemas describe mejor la erodabilidad de los suelos y que método de interpolación o predicción es el más adecuado, se utilizó la validación cruzada revisada en la sección (2.4.3). Para esto se empleó la función *krige.cv* de la librería *gstat* (Pebesma, 2004). Los estadísticos de comparación (2-22, 2-23 y 2-24), se calcularon con la función *criterio.cv* de la librería *geospt* (Melo et al., 2012).

5 Resultados y discusión

Los resultados descriptivos y el análisis geoestadístico se presentan en tres secciones. En primer lugar, con el estudio individual de las características físicas de suelo (sección 3.2). En segunda instancia a través del comportamiento conjunto de las variables en la zona de interés. Por último, usando el índice de erodabilidad (K) revisado en la sección (2.2.5).

5.1. Análisis individual de las variables

Las medidas de tendencia central y variabilidad para cada una de las características del suelo evaluadas (A , DMP , DA , DR y $CRAD$) se presentan en la tabla (5-1) y figura (5-1). Los valores del cuartil superior de A , DA y DR son semejantes a las categorías establecidas de las condiciones de susceptibilidad al deterioro de los territorios (Tabla 4-1). Esto sugiere que las observaciones localizadas arriba de este cuartil representan lugares con riesgo a degradación. Estos se establecen en diferentes sitios en la HRS . El contenido alto de A se encuentra principalmente en los municipios de Santana, San José de Pare, Togüí y Chitaraque en el departamento de Boyacá. En cuanto a DA se ubican en los municipios de Chipatá y San Benito (departamento de Santander). Mientras que la DR se manifiestan en los dos departamentos sin una agrupación en particular.

Los valores medios de las variables A , DMP , DA , DR y $CRAD$ son acordes a los reportados en estudios puntuales realizados en los municipios de Santana y Güepsa (Rueda et al., 2015; Ahumada, 2015). En cuanto a variabilidad A y $CRAD$ presentan los mayores registros ($CV > 35\%$), que pueden atribuirse a la diversidad de relieves y suelos que tiene la región (IGAC, 2003, 2005). Esta variación, refleja como el contenido de arenas afecta la capacidad de retención de agua (Montenegro y Malagón, 1990) y cuando alguna de estas propiedades es heterogénea, la otra tiende a presentar el mismo comportamiento. Se reportan CV bajos en otras investigaciones referentes a características físicas (Vásquez et al., 2010), pero las distancias de los muestreos en estos estudios fueron menores (< 700 metros) comparados con los de este trabajo.

La DA y el DMP tienen variación moderada (CV entre 16 - 35%), mientras que la DR fue más baja (CV menor a 15%). Las dos primeras propiedades mencionadas pueden ser alteradas por el manejo que se realiza en las fincas, puesto que la intervención de los suelos en el establecimiento de los cultivos hace que en la mayor parte del territorio se implementen prácticas de labranza, como son la arada y los continuos pases de rastrillo, que dejan el terreno bien mullido (condiciones favorables en el desarrollo de las plantas) para que la caña alcance su crecimiento óptimo (Manrique et al., 2000), y así mismo, uniformizan las características en cuanto a DA y DMP . Finalmente, la DR es una propiedad inherente al material que conforma al suelo siendo los valores descritos en la tabla (5-1) acordes a otros trabajos (Monroy et al., 2017).

Tabla 5-1: Medidas de localización y variabilidad de las variables porcentaje de arenas (A), diámetro medio ponderado (DMP), densidad aparente (DA), densidad real (DR) y capacidad de retención de agua disponible ($CRAD$). Calculados con datos obtenidos en la HRS en 2016.

MEDIDA	A (%)	DMP (mm)	DA (g/cm ³)	DR (g/cm ³)	% $CRAD$
Media	38,1	2,6	1,1	2,4	4,6
Mediana	37,6	2,5	1,1	2,5	4,1
Mínimo	1,6	0,5	0,5	1,7	0,1
Máximo	87,6	5,2	1,7	2,9	24,4
Cuartil inferior	27,6	1,9	0,9	2,4	3,1
Cuartil superior	46,8	3,2	1,2	2,5	5,4
Desviación estándar	13,8	0,8	0,2	0,1	2,7
Coefficiente de variación CV	36,3	31,0	18,0	5,1	59,6

Las observaciones atípicas (Figura 5-1) que pueden indicar riesgo a erosión son las de mayor valor de A , DA y DR . Los contenidos altos de A se ubicaron en los municipios de Chitaraque, Moniquirá, San José de Pare, Santana y Togüí en el departamento de Boyacá, mientras que la DA se observó en el municipio de Chipatá (departamento de Santander). Para la DR se distribuyeron en los departamentos de Boyacá (municipios de Santana, Togüí) y Santander

(municipios de Chipatá, Güepsa, San Benito y Vélez). Estos datos no se agrupan en áreas específicas lo que puede ser un indicio de pequeñas zonas ya degradadas.

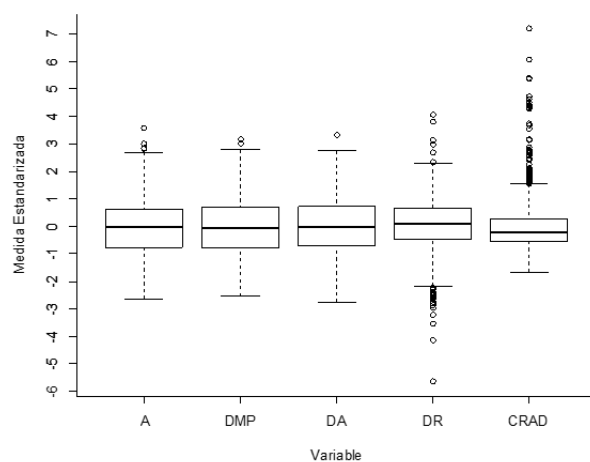


Figura 5-1: Diagramas de caja y bigotes (datos estandarizados) de las variables porcentaje de arenas (A), diámetro medio ponderado (DMP), densidad aparente (DA), densidad real (DR) y capacidad de retención de agua disponible ($CRAD$). Calculados con datos obtenidos en la *HRS* en 2016.

Los gráficos de regresión no paramétrica y las correspondientes bandas de confianza para la hipótesis de no efecto definidos en la sección (2.4.2) calculados con la información de las variables A , DMP , DR y $CRAD$ son dadas en las figuras (5-2 a 5-6). Estos indican que en general no hay tendencia espacial en la media. Un comportamiento similar se reporta en otros estudios (Ceddia et al., 2009; Jaramillo et al., 2011). Una excepción es DA , donde las curvas de regresión salen de las bandas de confianza (Figura 5-4). En este caso se presenta una tendencia respecto a la longitud. Un análisis detallado en la variación de las coordenadas revela que las magnitudes más grandes de DA corresponden a sitios ubicados en los municipios de Barbosa, Güepsa y Chipatá, que se deben a la intensidad de las prácticas de preparación de los terrenos que se implementan en esas zonas. La labranza produce la alteración de la condición estructural del terreno, mejorando temporalmente la aireación, aunque al realizarlo reiteradamente se agrupan nuevamente las partículas por la acción de lluvias, riego y tráfico de maquinaria agrícola incrementando así la DA por compactación (Montenegro y Malagón, 1990).

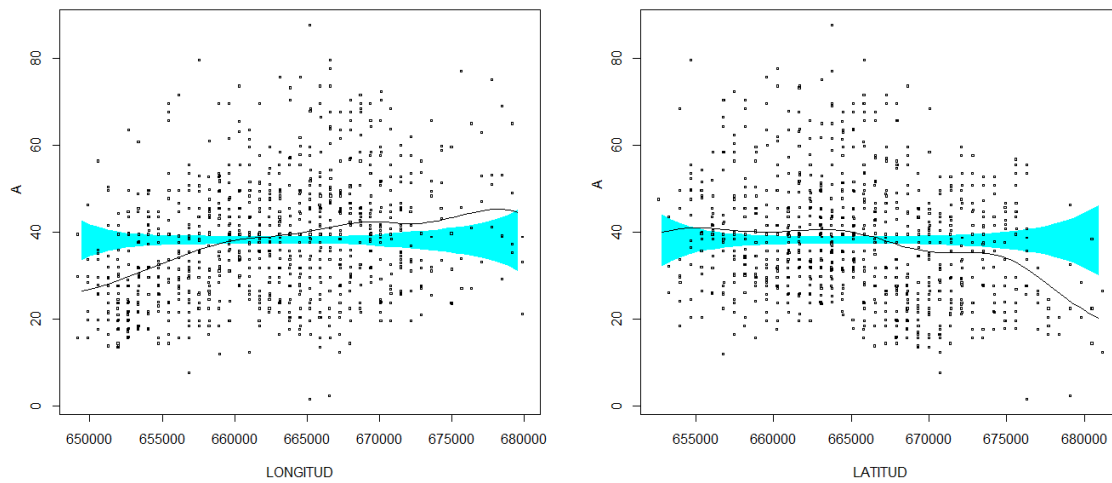


Figura 5-2: Gráfico de dispersión de valores de porcentaje de arenas (A) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016.

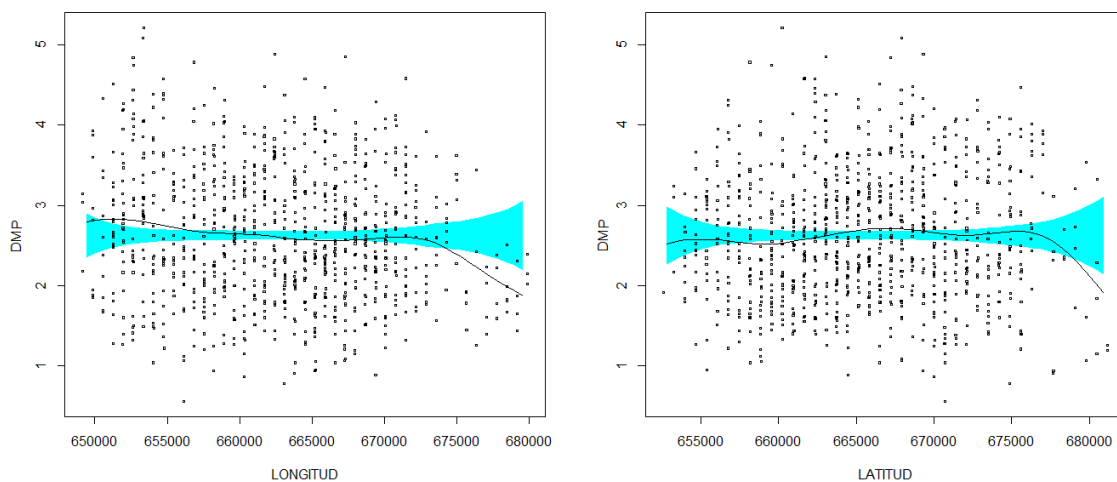


Figura 5-3: Gráfico de dispersión de valores de diámetro medio ponderado (DMP) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016.

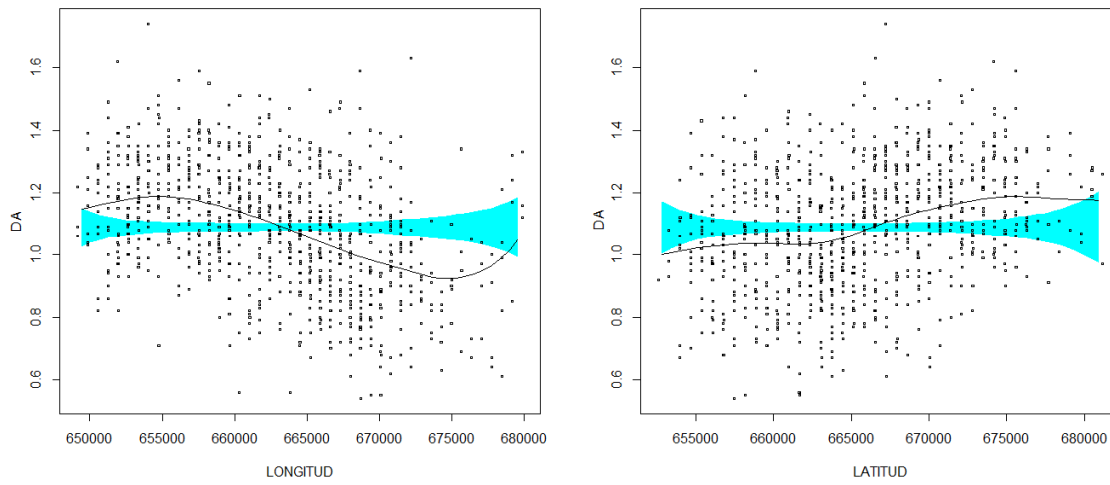


Figura 5-4: Gráfico de dispersión de valores de densidad aparente (DA) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016.

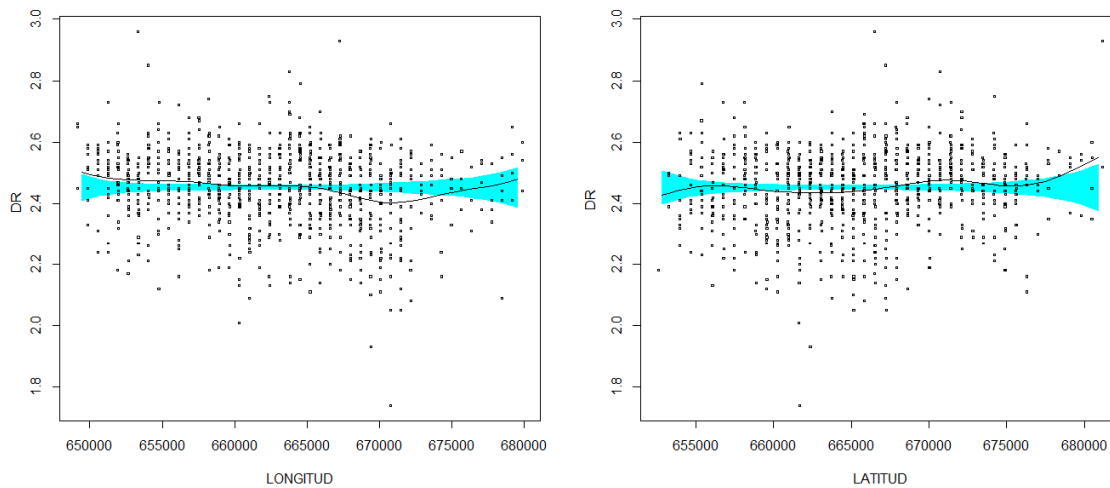


Figura 5-5: Gráfico de dispersión de valores de densidad real (DR) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016.

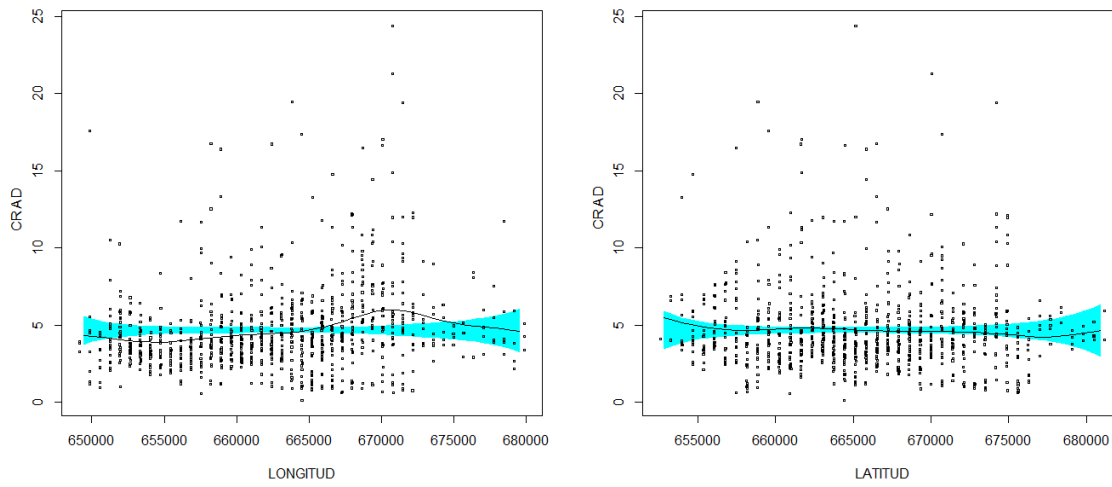


Figura 5-6: Gráfico de dispersión de valores de porcentaje de capacidad de retención de agua disponible ($CRAD$) respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016.

En lo referente al estudio de anisotropía, las figuras (5-7) a (5-9) muestran que para distancias cercanas al origen (menores a 10 kilómetros) los semivariogramas son similares. Indican por consiguiente que las propiedades evaluadas no presentan variación considerable en su estructura de correlación. Este patrón también se registró en Ceddia et al. (2009). En general los gráficos anteriormente mencionados sugieren que hay estacionariedad en los procesos estudiados. Es común que bajo estacionariedad los semivariogramas direccionales tengan diferencias en distancias grandes, puesto que hay un número pequeño de parejas de sitios para hacer la estimación (Warrick, 2002).

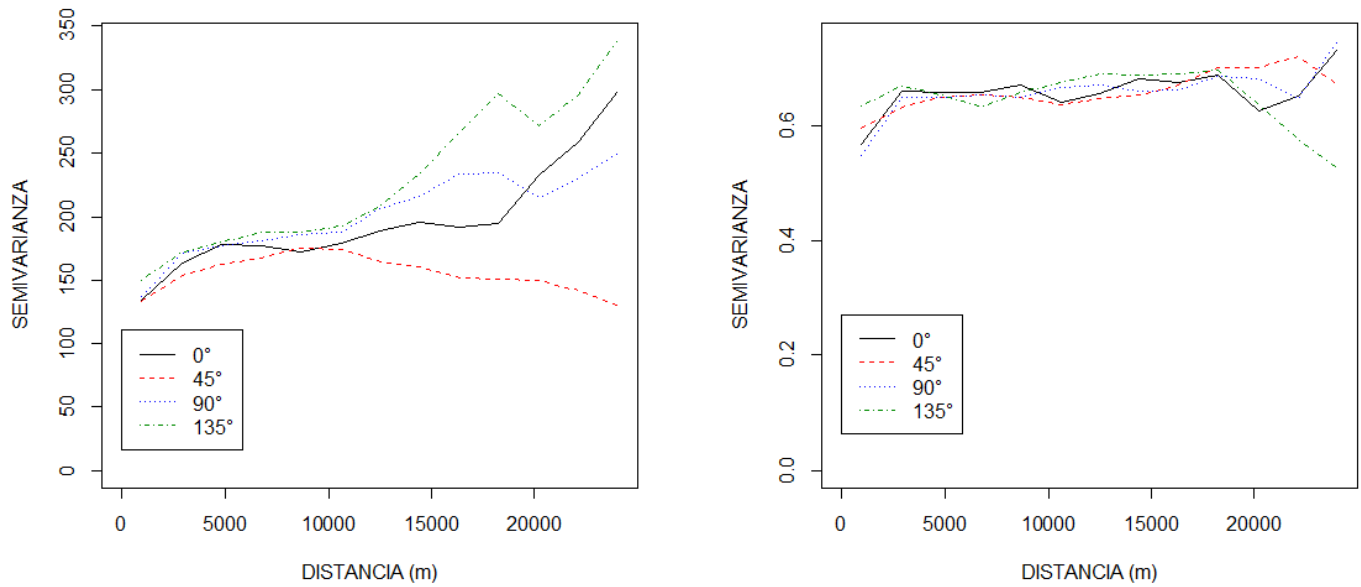


Figura 5-7: Semivariogramas direccionales (0, 45, 90 y 135 grados), para los procesos de porcentaje de arenas (izquierda) y diámetro medio ponderado (derecha).

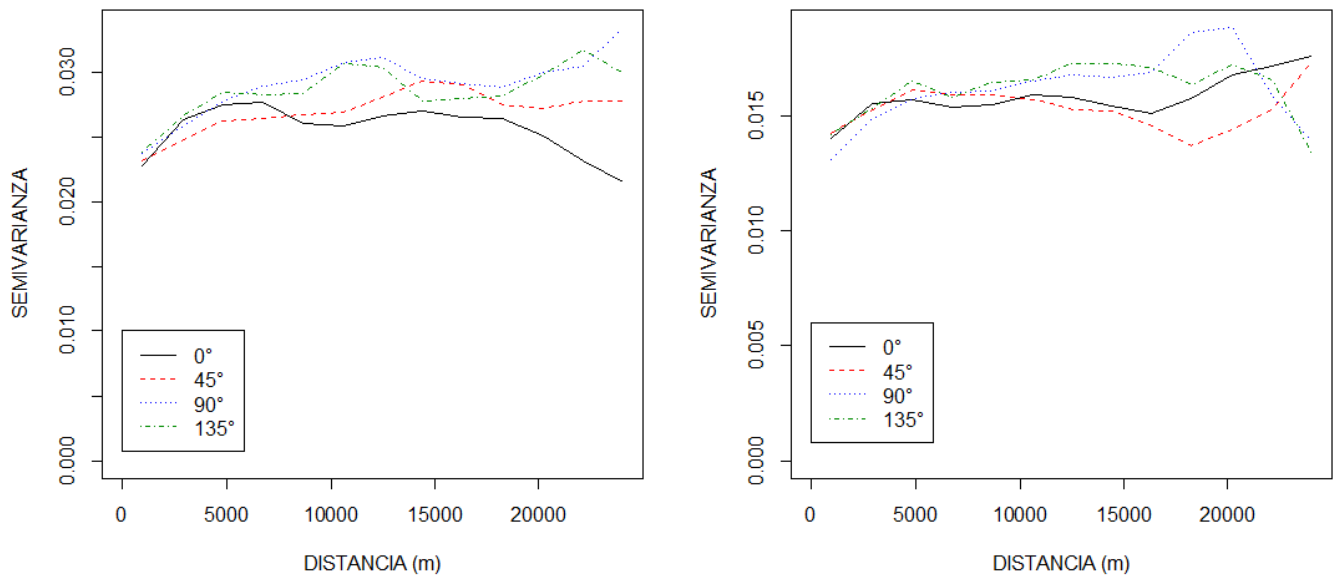


Figura 5-8: Semivariogramas direccionales (0, 45, 90 y 135 grados), para el proceso de densidad aparente (izquierda) y densidad real (derecha).

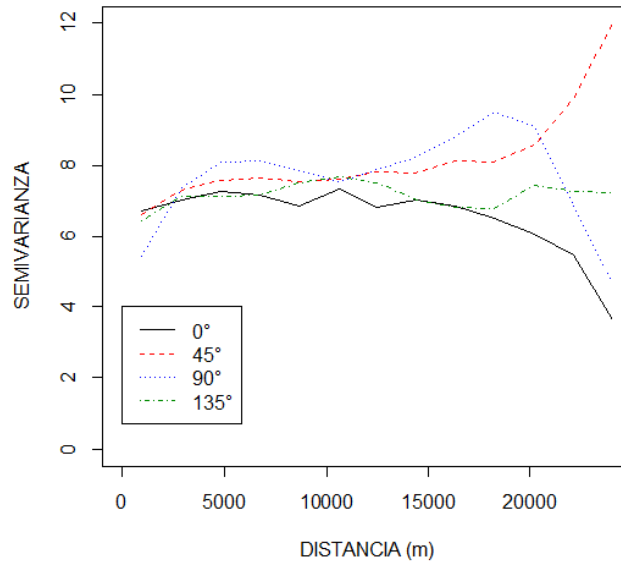


Figura 5-9: Semivariogramas direccionales (0, 45, 90 y 135 grados), para la capacidad de retención de agua disponible .

Los modelos ajustados a los semivariogramas experimentales (Figuras 5-10 a 5-12) tienen en todos los casos valores de $\hat{\tau}^2$ altos (por encima del 50% del valor de la meseta), lo que sugiere que hay errores de medición (variabilidad instrumental) o que la distancia de muestreo (700 m) no permite estimar bien este parámetro (no hay observaciones del variograma experimental a distancias menores de 700 m, lo que dificulta su estimación cerca al origen). Los rangos ($\hat{\phi}$) en general menores a 5000 m, que indican baja o media dependencia espacial, teniendo en cuenta que hay puntos muestreados separados por distancias superiores a 20 kilómetros. Los modelos teóricos considerados en la sección (2.4.2) que mejor se ajustaron son el esférico (*DMP* y *CRAD*) y el exponencial (*A*, *DA* y *DR*). Un resultado similar es encontrado por Vásquez et al. (2010), Bravo et al. (2011) y Maestre y de Andreu (2018). En estos estudios se reportaron valores mayores para $\hat{\phi}$, lo que posiblemente podría estar asociado a que consideraron un enmallado más denso (distancias inferiores a 30 m). Teniendo en cuenta los resultados del estudio de tendencia (Figuras 5-2 a 5-6) se decidió emplear predicción kriging ordinario en todos los casos, salvo en la *DA*, en el que se usó kriging universal.

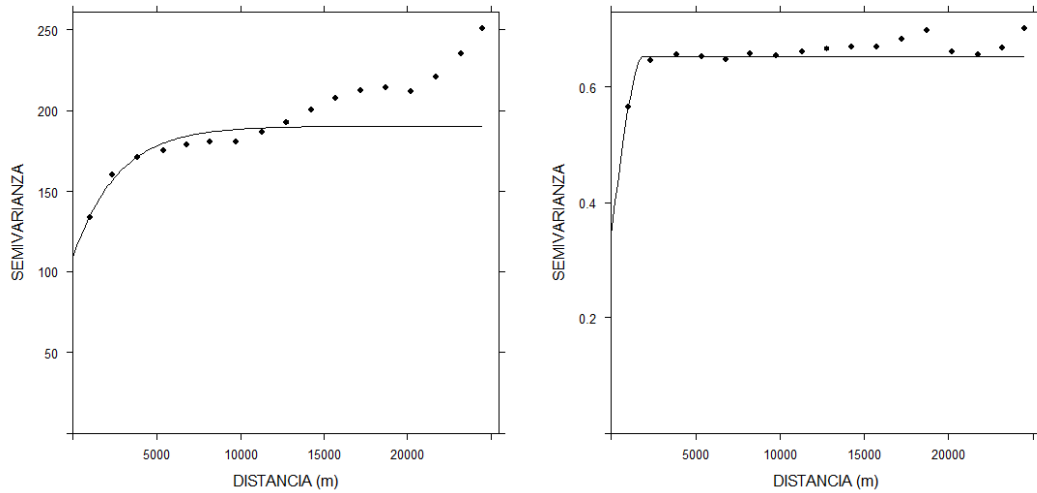


Figura 5-10: Semivariogramas experimentales (puntos negros) calculados con datos muestrales y ajuste de un modelo exponencial ($\hat{\tau}^2 = 108,88$, $\hat{\sigma}^2 = 81,64$ y $\hat{\phi} = 2639,03$) y uno esférico ($\hat{\tau}^2 = 0,34$, $\hat{\sigma}^2 = 0,31$ y $\hat{\phi} = 1922,91$) para las variables porcentaje de arenas (A) (izquierda) y diámetro medio ponderado (DMP) (derecha), respectivamente.

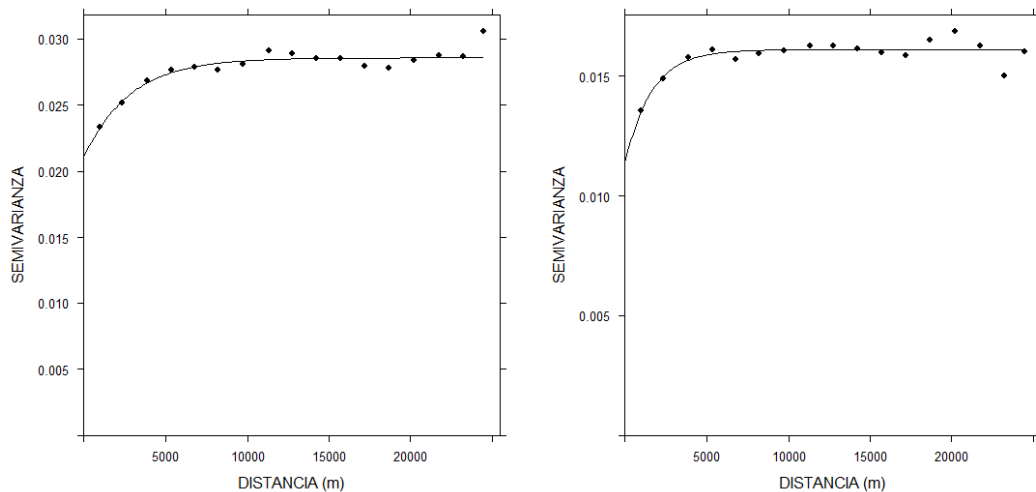


Figura 5-11: Semivariogramas experimentales (puntos negros) calculados con datos muestrales y ajuste de modelos exponenciales ($\hat{\tau}^2 = 0,021$, $\hat{\sigma}^2 = 0,007$ y $\hat{\phi} = 2818,13$) y ($\hat{\tau}^2 = 0,011$, $\hat{\sigma}^2 = 0,005$ y $\hat{\phi} = 1637,79$) para las variables de densidad aparente (DA) (izquierda) y real (DR) (derecha), respectivamente.

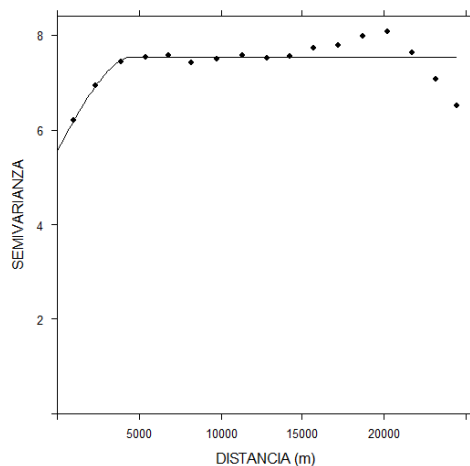


Figura 5-12: Semivariograma experimental (puntos negros) calculados con datos muestrales y ajuste de modelo esférico ($\hat{\tau}^2 = 5,55$, $\hat{\sigma}^2 = 1,98$ y $\hat{\phi} = 4638,14$) para la variable capacidad de retención de agua disponible (*CRAD*) (línea continua).

Usando los métodos de interpolación y predicción espacial descritos en las secciones (2.4.1 y 2.4.2) se elaboraron mapas de isolíneas (Figuras 5-13 a 5-17). Cabe resaltar que el uso de los métodos geoestadísticos permite obtener en este caso mayor definición de las áreas y además posibilita la evaluación de la incertidumbre en las predicciones (Figuras 5-18 a 5-20).

De acuerdo con los mapas de predicción (Figuras 5-13 a 5-17), se puede establecer que los menores valores de A ($< 30\%$), se presentan en los municipios de Chipatá y el norte de San Benito (Figura 5-13). Por el contrario, se observa una mayor concentración de A en el norte de Togüí, sur de Santana, oriente de San José de Pare y occidente de Chitaraque (Figura 5-13). El alto contenido de A disminuye la atracción de las partículas del suelo que son fácilmente dispersadas por el impacto de las lluvias destruyendo los agregados, favoreciendo así el transporte de estos por las aguas de escorrentía (Rivera et al., 2010). Además, los contenidos de arena superiores al 50 % (Tabla 4-1) pueden representar limitaciones en la retención de humedad, incrementos en los riesgos de lixiviación de agroquímicos y menor almacenamiento de agua disponible para los cultivos (Torres et al., 2017). Esto indica que las zonas con las condiciones anteriormente mencionadas tienen más riesgo de susceptibilidad a la erosión.

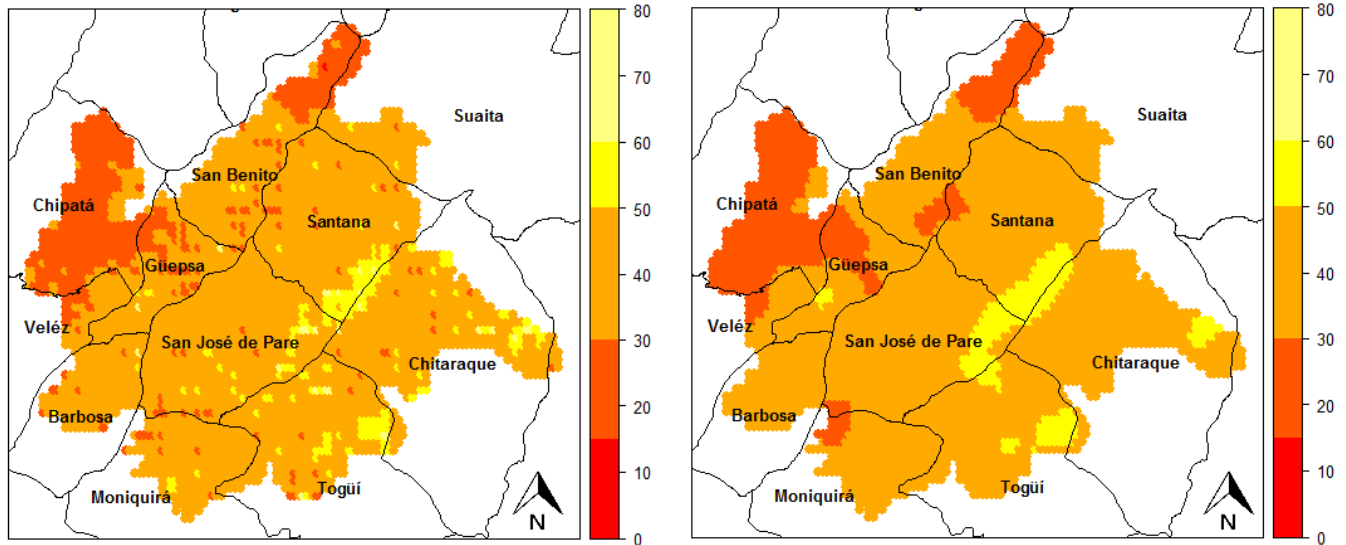


Figura 5-13: Distribución espacial para el porcentaje de arenas (A). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).

Con respecto al *DMP*, los mapas de predicción (Figura 5-14) revelan que las zonas aledañas a Togui, San José de Pare, Santana y Chitaraque (departamento de Boyacá), son moderadamente estables (Tabla 2-2), mientras que hacía el departamento de Santander el nivel de estabilidad del suelo es mayor. En general en la zona de estudio no se presenta una condición limitante (Tabla 4-1) para el desarrollo de la caña panelera en cuanto a firmeza estructural del suelo.

Las condiciones ideales de *DA* para el desarrollo de la caña están entre 1,1 y 1,2 gr/cm^3 (Anschau et al., 2009). Se observa (Figura 5-15) que en los municipios de Chipatá, Güepa y sur de San Benito, se encuentran valores por encima del rango ideal. Esta propiedad tiene una relación directa con la erodabilidad (*K*), puesto que entre más grande sea el valor de *DA*, mayor será el *K* de los suelos (Rivera et al., 2010) y por tanto, constituye un indicativo de que las zonas anteriormente mencionadas puedan tener riesgo considerable de erosión con respecto al resto del territorio de la *HRS*.

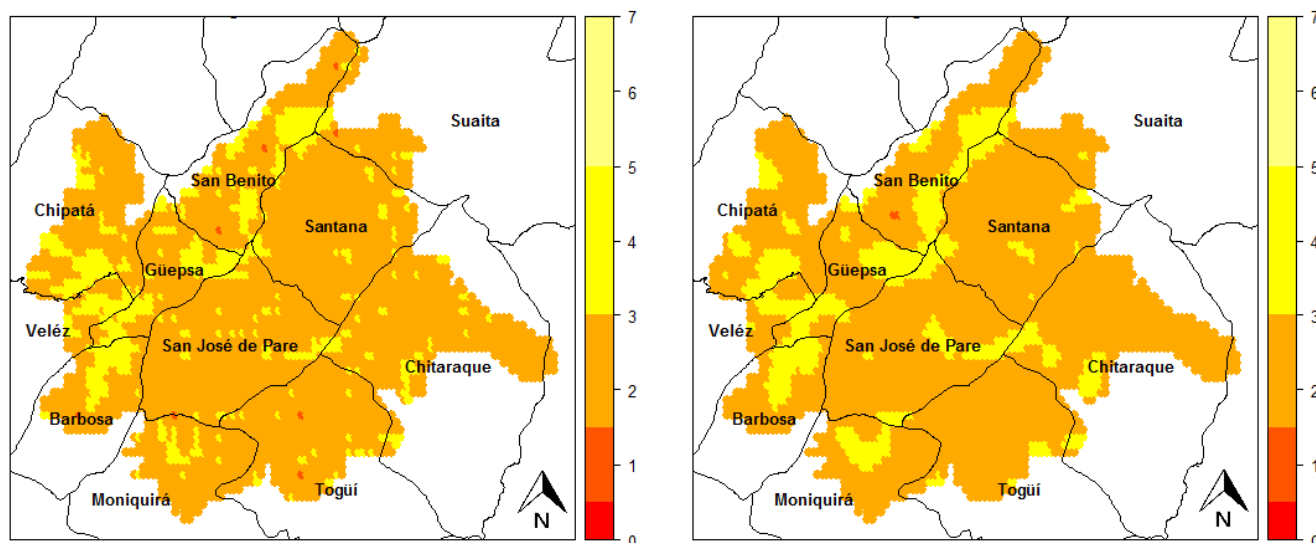


Figura 5-14: Distribución espacial para el diámetro medio ponderado (DMP). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).

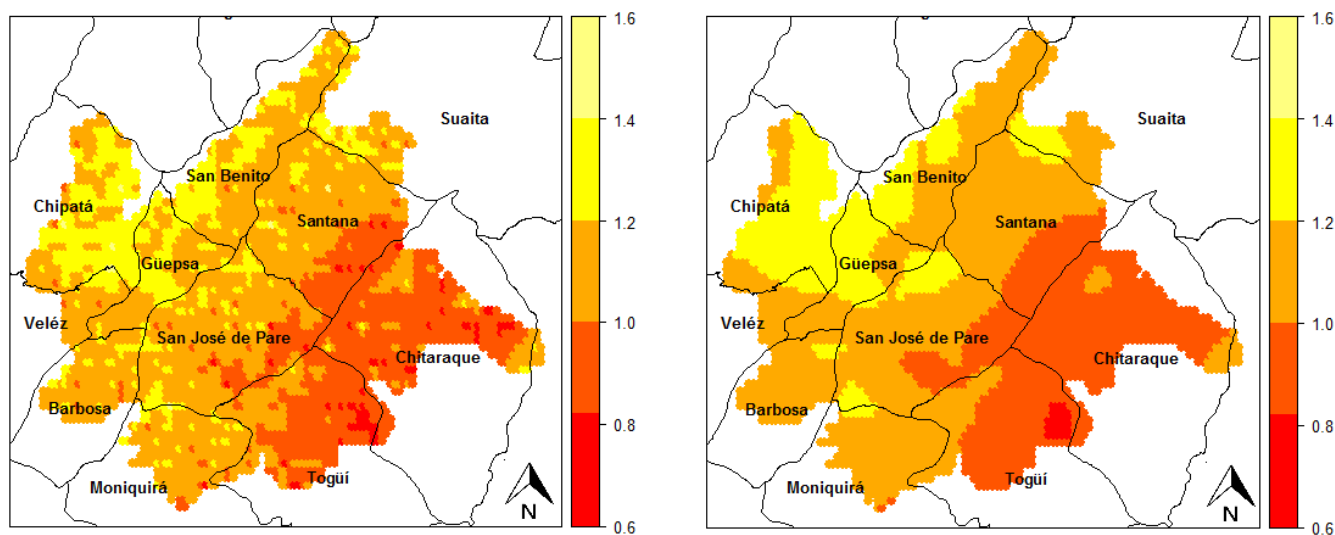


Figura 5-15: Distribución espacial para la densidad aparente (DA). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging universal (derecha).

Para la *DR* (Figura 5-16) se pueden identificar valores mayores a $2,3 \text{ gr/cm}^3$ similar a lo reportado por IGAC (2005). Los territorios con menor *DR* se encuentran en los municipios de Barbosa, Chitaraque, San José de Pare, Suaita y Togüí. En estos lugares los rendimientos de caña panelera son superiores a 11 toneladas por hectárea (cifras de Agronet, 2015), posiblemente porque las zonas con menos *DR* tienen mayor contenido de materia orgánica (*MO*), una propiedad fundamental en el cultivo (Lal y Shukla, 2004) como se revisó en la sección (2.1). Por el contrario, no se identifican áreas con niveles de *DR* que presenten alto riesgo de susceptibilidad a degradación en suelos.

En relación a *CRAD*, se encontraron valores inferiores al 15 % en la zona de estudio (Figura 5-17). Esto sugiere que hay baja o muy baja capacidad de almacenamiento de agua disponible para las plantas (Tabla 2-3). Los resultados coinciden con los reportados por IGAC (2005). El cultivo de caña puede verse afectado en el crecimiento vegetativo y además, a una menor *CRAD* el suelo es más susceptible al proceso erosivo (Montenegro y Malagón, 1990).

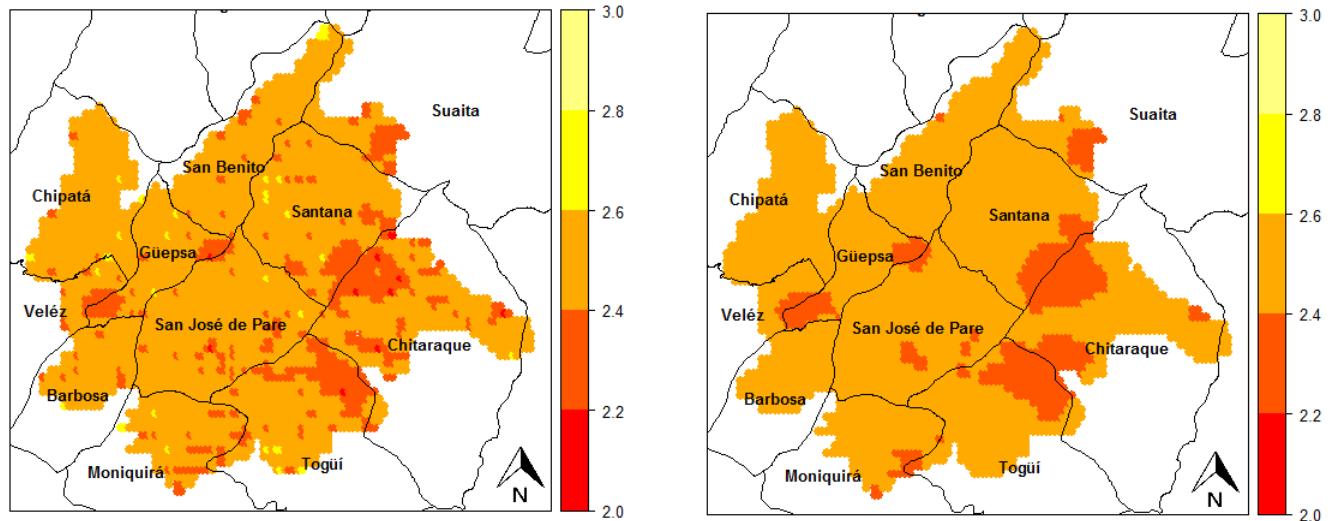


Figura 5-16: Distribución espacial para la densidad real (*DR*). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).

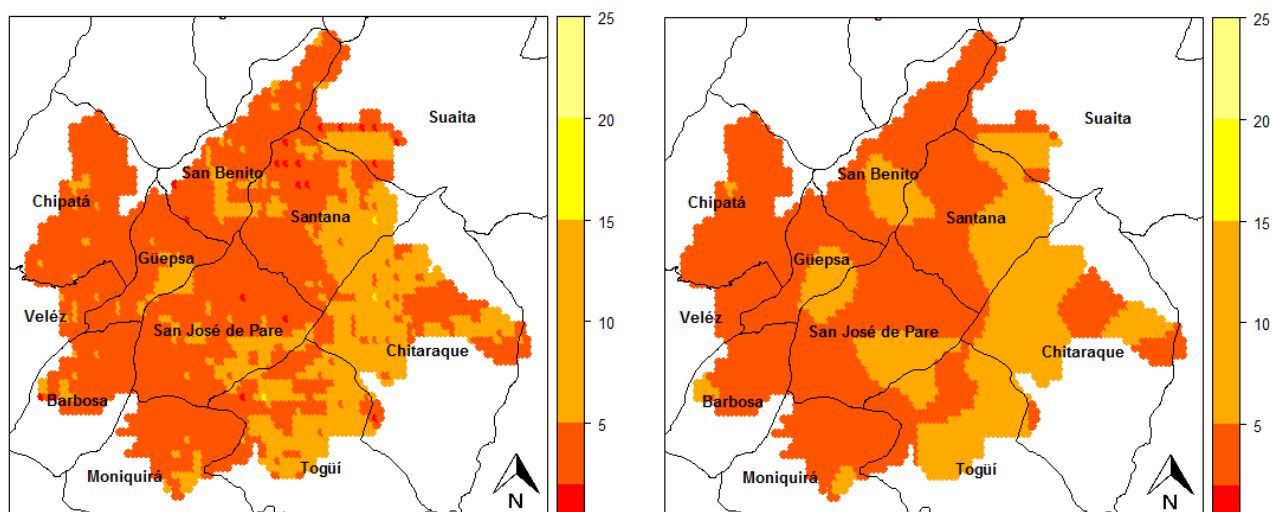


Figura 5-17: Distribución espacial de la capacidad de retención de agua disponible (*CRAD*). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).

En general, al observar las diferentes zonificaciones realizadas (Figuras 5-13 a 5-17), se puede establecer que los municipios de Chipatá, Güepsa, sur de San Benito y norte de San José de Pare, son territorios de la *HRS* susceptibles a erosionarse, puesto que hay valores altos de *DA*, *DR* y bajos contenidos de *CRAD*. Un estudio realizado por CORPOBOYACÁ (2014), determinó que el municipio de Chipatá presenta un inadecuado manejo de los residuos producidos por los trapiches paneleros y vertientes de agua altamente contaminadas que afectan la calidad de los suelos. En Güepsa y San Benito se evidencia el deterioro de la capa arable por el aumento de las construcciones de viviendas. También se registran quemas de terrenos con fines agrícolas e invasión de las reservas forestales en el municipio de Güepsa. Las áreas referenciadas anteriormente se pueden visualizar en mayor medida con las delimitaciones usando la *DA* y el *DMP* (Figuras 5-14 y 5-15), siendo estas las que pueden representar la erodabilidad en la *HRS* con la metodología de análisis individual de las variables.

Las predicciones en todos los casos tienen mejores resultados de validación cruzada comparados con las interpolaciones realizadas (Tabla 5-2). Además, los mapas de las varianzas (generados con los métodos de predicción) sugieren que la incertidumbre es homogénea en gran parte del área de estudio (Figuras 5-18 a 5-20), mostrando solo en el contorno de la *HRS* mayor variabilidad, debido al menor número de sitios de muestreo realizado en estas áreas.

Tabla 5-2: Estadísticos de validación cruzada (Error medio de predicción EMP , raíz del error cuadrado medio de predicción $RECM$ y coeficiente de determinación R^2) de los análisis de predicción (Kriging ordinario KO y universal KU) e interpolación (inverso de la distancia DIP), realizados con las variables porcentaje de arenas (A), diámetro medio ponderado (DMP), densidad aparente (DA), densidad real (DR) y capacidad de retención de agua disponible ($CRAD$).

VARIABLE	EMP	$RECM$	R^2
ARENAS (DIP)	-0,14	11,94	25,5
ARENAS (KO)	-0,01	11,75	29,7
DMP (DIP)	-0,01	0,77	0,9
DMP (KO)	0,01	0,77	1,1
DA (DIP)	-0,01	0,16	34,8
DA (KU)	-0,01	0,16	36,1
DR (DIP)	0,01	0,12	10,3
DR (KO)	0,01	0,12	11,3
$CRAD$ (DIP)	-0,01	2,56	12,6
$CRAD$ (KO)	-0,01	2,55	13,4

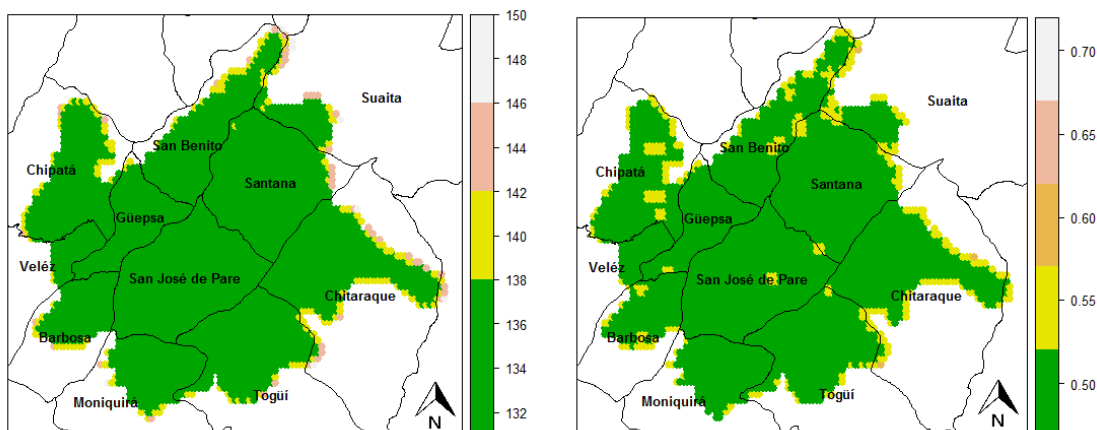


Figura 5-18: Distribución espacial de varianzas de la predicción del kriging del porcentaje de arenas (A) (izquierda) y diámetro medio ponderado (DMP)(derecha).

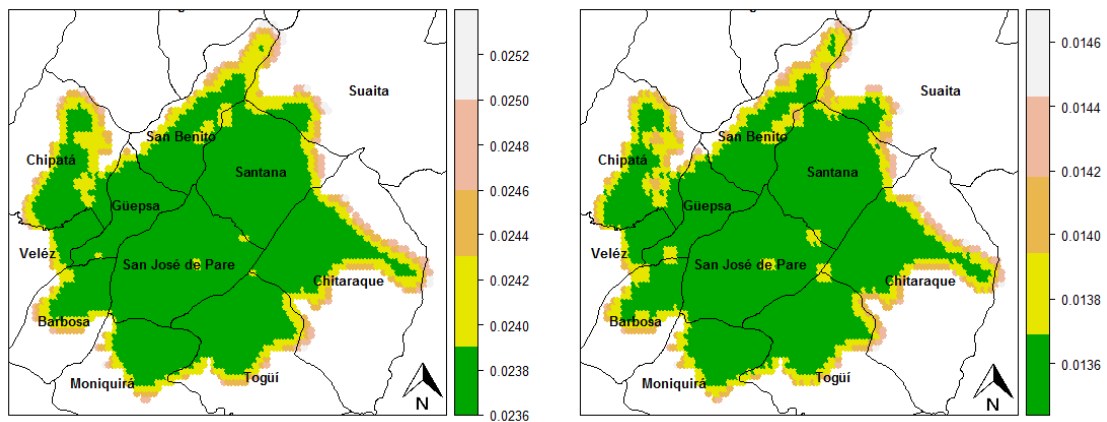


Figura 5-19: Distribución espacial de varianzas de la predicción del kriging de la densidad aparente (DA) (izquierda) y real (DR) (derecha).

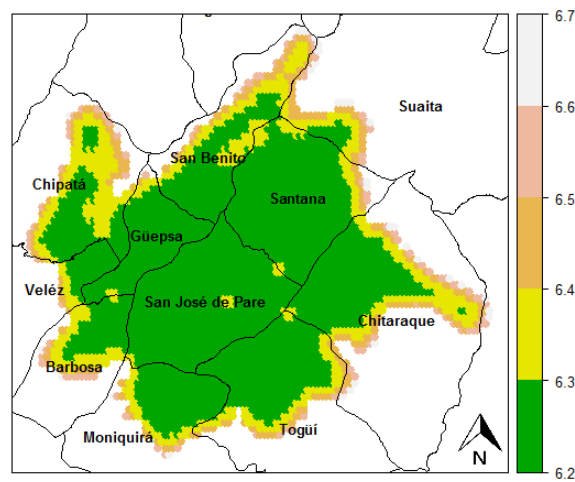


Figura 5-20: Distribución espacial de varianzas de la predicción del kriging de la capacidad de retención de agua diaponible ($CRAD$).

5.2. Análisis conjunto por componentes principales

Los coeficientes de correlación son en general significativos (valor $p < 0,05$) indicando que hay relación entre las variables. Dos excepciones a este patrón son las debidas a la relación entre *DMP* y *DA* y entre *DMP* y *CRAD* (Tabla 5-3). Lo anterior indica que es factible hacer un estudio conjunto del comportamiento espacial de las variables estudiadas en aras de identificar las zonas susceptibles a erosión. Cabe aclarar que la significancia de los coeficientes de correlación debe interpretarse con precaución, puesto que las variables son autocorrelacionadas.

Entre *A* y *CRAD* y entre *DA* y *DR* hay correlación positiva (directa). En todos los otros casos la asociación (cuando es significativa) es inversa (Tabla 5-3). La relación de *A* y *CRAD* no establece una clara dependencia, pero el estudio de Vargas et al. (2015) muestra que *A* tiene un vínculo opuesto con el contenido de agua por la dinámica de la humedad con los componentes del suelo, como la distribución de *A*, porosidad y *MO* (Burbano y Silva, 2010). La correlación de *DA* y *DR* es considerable ($> 0,5$), indicando que sitios con valores altos de *DA* tienen *DR* elevados. Los resultados de Salamanca y Sadeghian (2006) son acordes a los de este trabajo presentando coeficientes de correlación mayores a 0,5 en las características antes mencionadas.

La relación inversa entre *DA* y *CRAD* es esperable, pues según Blanco (2009) a menor valor de *DA* (alta porosidad) mayor capacidad de retención de agua en el suelo. Un resultado semejante al anterior es reportado en Varón et al. (2014). Algo similar ocurre con la relación entre *DA* y *A*, ya que a mayor contenido de arena, hay más espacio poroso y por consiguiente menor valor de *DA*. Además, al aumentar el contenido de *A* la estabilidad de los terrenos (*DMP*) se disminuye, como lo muestra la correlación negativa entre estas características (Tabla 5-3). Resultados parecidos se reportan en Novillo et al. (2018) y Vargas et al. (2015).

Tabla 5-3: Coeficientes de correlación de Pearson (valor-p), calculados con la información de las variables físicas del suelo; arenas (A), diámetro medio ponderado (DMP), densidad aparente (DA), densidad real (DR) y capacidad de retención de agua disponible ($CRAD$), medidas en la HRS en 2016.

Coeficiente correlación	A	DMP	DA	DR	$CRAD$
A	1				
DMP	-0,119 ***	1			
DA	-0,322 ***	-0,051	1		
DR	-0,312 ***	-0,107 **	0,522 ***	1	
$CRAD$	0,136 ***	0,035	-0,417 ***	-0,282 ***	1
Significancia	*** < 0,001	** < 0,01	* < 0,05		

A partir de la matriz de correlación (Tabla 5-3) y de acuerdo a lo descrito en la sección (2.4.4) se redujo el conjunto de variables mediante un ACP . Los resultados de valores y vectores propios son dados en las tablas (5-4) y (5-5). El primer componente principal ($CP1$) explica el 40,5% de la variabilidad y está asociado a A , DMP y $CRAD$ en sentido directo y a DA y DR en sentido inverso. El componente principal dos ($CP2$) representa el 21,7% de varianza y hace referencia a DMP en sentido directo y A en sentido inverso. El primer componente hace referencia a la humedad del suelo ($CRAD$), mientras que el segundo tiene relación con la estabilidad del suelo (DMP). Los otros componentes principales aportan poco individualmente, por lo cual no son considerados en la explicación.

Tabla 5-4: Valores propios y porcentajes de varianza explicados por los componentes principales generados con la información de cuatro variables físicas del suelo, medidas en la HRS en 2016.

COMPONENTE	VALOR PROPIO	% VARIANZA	% ACUMULADO
1	2,027	40,56	40,56
2	1,084	21,69	62,25

Tabla 5-5: Vectores propios de los dos primeros componentes principales.

VARIABLE	COMPONENTE 1	COMPONENTE 2
<i>A</i>	0,401	-0,468
<i>DMP</i>	0,054	0,865
<i>DA</i>	-0,585	-0,042
<i>DR</i>	-0,546	-0,099
<i>CRAD</i>	0,442	0,140

Con los ejes factoriales generados se llevó a cabo un análisis exploratorio y se obtuvieron mapas de distribución espacial. En la figura (5-21) se identifica la presencia de varias observaciones atípicas, la mayoría de ellas correspondientes a valores muy altos del *CP1*. Estos corresponden a puntos de muestreo en los municipios de Chitaraque, San José de Pare, Santana y Togüi, que presentan riesgo bajo de erosión (es decir valores bajos de *DA*, *DR* y *CRAD*). Los dos valores atípicos negativos en el *CP2* están asociados también a sitios de baja susceptibilidad a la erosión. Estos fueron obtenidos en Chipatá.

El dato atípico negativo en el *CP1* corresponde al municipio de Chipatá. El dato atípico positivo en el *CP2* es un punto del muestreo ubicado en el municipio de San José de Pare. En ambos hay condiciones que favorecen la erosión (menor estabilidad del suelo y mayor *A*).

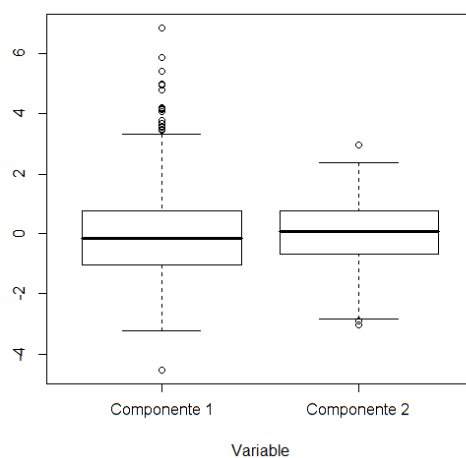


Figura 5-21: Diagrama de caja y bigotes de los dos primeros componentes principales, generados a partir de información medida en la *HRS* en 2016.

Los modelos de regresión no paramétricos de los dos primeros componentes principales respecto a las coordenadas de ubicación espacial son presentados en las figuras (5-22) y (5-23). Las curvas ajustadas en ambos casos (Figuras 5-22 y 5-23) están ligeramente por fuera de las bandas de confianza de no efecto, lo cual muestra que no hay una clara tendencia espacial y que por lo tanto puede asumirse el supuesto de estacionariedad para llevar a cabo el análisis geoestadístico de estos dos factores.

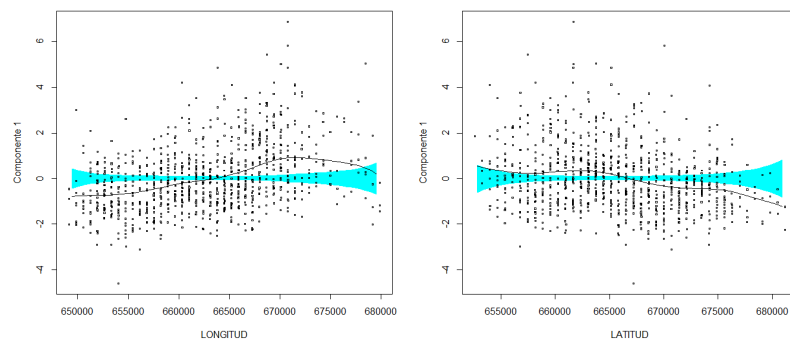


Figura 5-22: Gráfico de dispersión de valores “scores” del componente principal uno, respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en *HRS* en 2016.

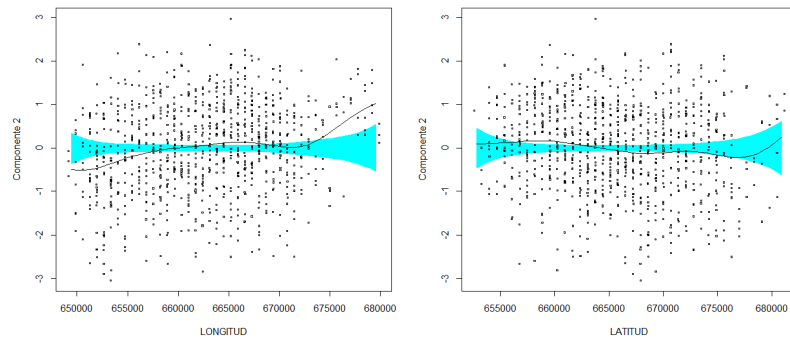


Figura 5-23: Gráfico de dispersión de valores “scores” del componente principal dos, respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016.

Como en el caso del análisis individual, el estudio de los variogramas direccionales (Figura 5-24) y el ajuste de los modelos al semivariograma omnidireccional (Figura 5-25), indica que por un lado es razonable considerar que las variables son isotrópicas y por el otro muestra que hay poca dependencia espacial (valores altos de $\hat{\tau}^2$ y pequeños de $\hat{\phi}$ en ambos casos).

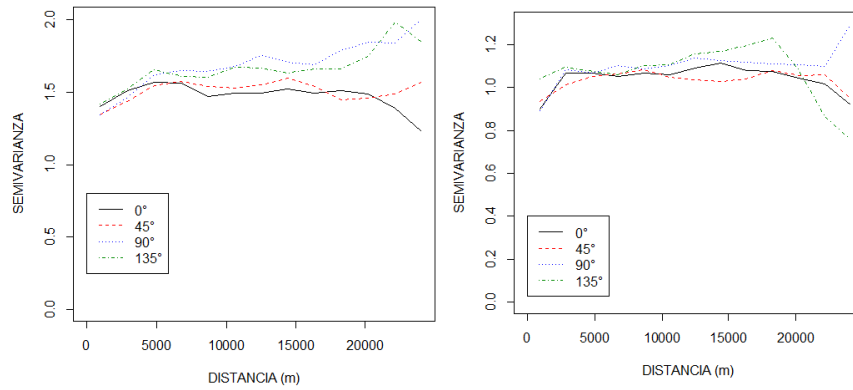


Figura 5-24: Semivariogramas direccionales (0, 45, 90 y 135 grados), para los componentes principales uno (izquierda) y dos (derecha).

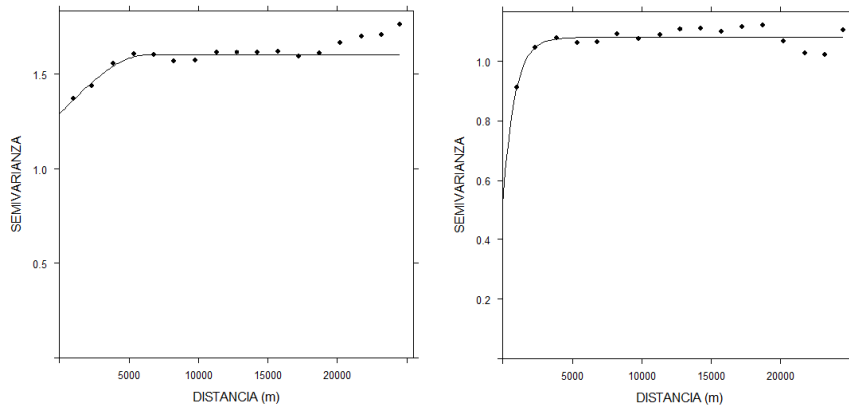


Figura 5-25: Semivariogramas experimentales (puntos negros) calculados con datos muestrales y ajuste de modelos esférico ($\hat{\tau}^2 = 1,30$, $\hat{\sigma}^2 = 0,31$ y $\hat{\phi} = 6308,48$) y exponencial ($\hat{\tau}^2 = 0,51$, $\hat{\sigma}^2 = 0,56$ y $\hat{\phi} = 837,02$) para las variables componente principal uno (izquierda) y dos (derecha), respectivamente.

Los mapas de distribución espacial con la información de los dos primeros componentes principales se muestran en las figuras (5-26) y (5-27). De acuerdo con los resultados del *CP1* parecen encontrarse dos regiones. Una con alta susceptibilidad en el norte de la *HRS* (municipios de Chipatá, Barbosa, Vélez, Güepesa, San Benito y parte de San José de Pare y Santana) que tienen valores negativos (es decir, hay valores altos de *DA*, *DR* y bajos *CRAD*), lo que desde el punto de vista práctico indica que estas son propensas a la erosión, ya que al aumentar la *DA* se incrementa la compactación, la porosidad disminuye y se afectan las condiciones de retención de humedad que limitan el crecimiento de las raíces de las plantas (Salamanca y Sadeghian, 2006). La otra región con valores positivos (bajos *DA*, *DR* y altos *CRAD*) sugieren menor susceptibilidad en los municipios de Chitaraque, Togüi y parte de San José de Pare y Santana. Estas zonas se caracterizan por ser moderadamente estables a muy estables (IGAC, 2005).

El mapa de *CP2* (Figura 5-27) ayuda a delimitar mejor las regiones susceptibles. Los valores más negativos (que tienen altos contenidos de *A* y menor *DMP*) se ubican al sur de San Benito y en la parte oriental de los municipios de San José de Pare y Chitaraque. Estas zonas son propensas a degradarse por que al tener alto *A* y baja estabilidad se favorece el desprendimiento de los agregados de los terrenos (Rivera et al., 2010). Por el contrario, en el sur de Chipatá y oriente de San Benito tienen valores positivos, es decir valores por debajo del promedio en *A* y alto *DMP*, lo que indica que allí hay menor riesgo de erosión.

Los resultados descritos anteriormente a través de la interpretación de los gráficos (5-26 y 5-27) son completamente coherentes con los reportados en el análisis individual de las características físicas en la sección (5.1). Las predicciones usando la información de los componentes principales tienen mejores valores de validación cruzada comparados con las interpolaciones realizadas (Tabla 5-6). Además, los mapas de las varianzas (generados con los métodos de predicción) para los dos componentes, sugieren que la incertidumbre es homogénea en gran parte del área de estudio (Figura 5-28), mostrando solo en el contorno de la *HRS* mayor variabilidad, debido al menor número de sitios de muestreo realizado en estas áreas.

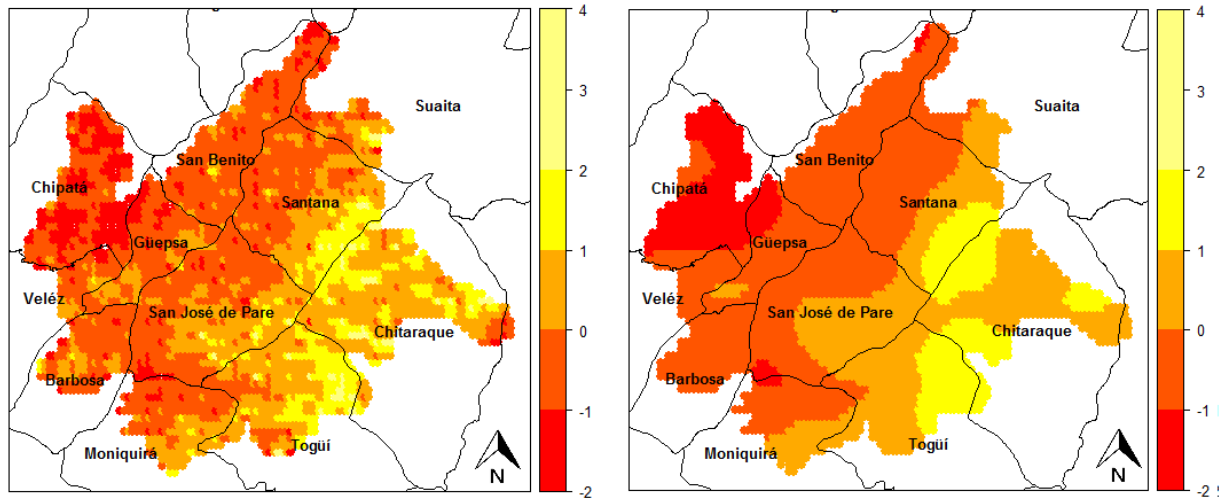


Figura 5-26: Distribución espacial del primer componente principal. Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).

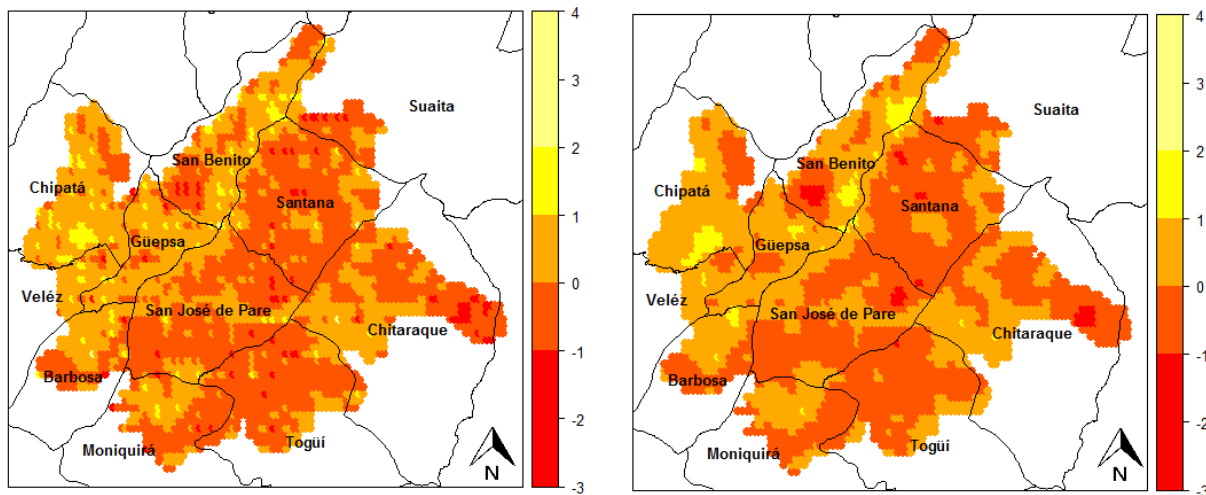


Figura 5-27: Distribución espacial del segundo componente principal. Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging universal (derecha).

Tabla 5-6: Estadísticos de validación cruzada (Error medio de predicción EMP , raíz del error cuadrado medio de predicción $RECM$ y coeficiente de determinación R^2) de los análisis de predicción (Kriging ordinario KO y universal KU) e interpolación (inverso de la distancia DIP), realizados con los componentes principales uno y dos.

VARIABLE	EMP	$RECM$	R^2
COMPONENTE 1 (DIP)	-0,005	1,216	27,1
COMPONENTE 1 (KU)	-0,001	1,203	28,5
COMPONENTE 2 (DIP)	0,003	0,976	12,9
COMPONENTE 2 (KO)	-0,001	0,966	14,2

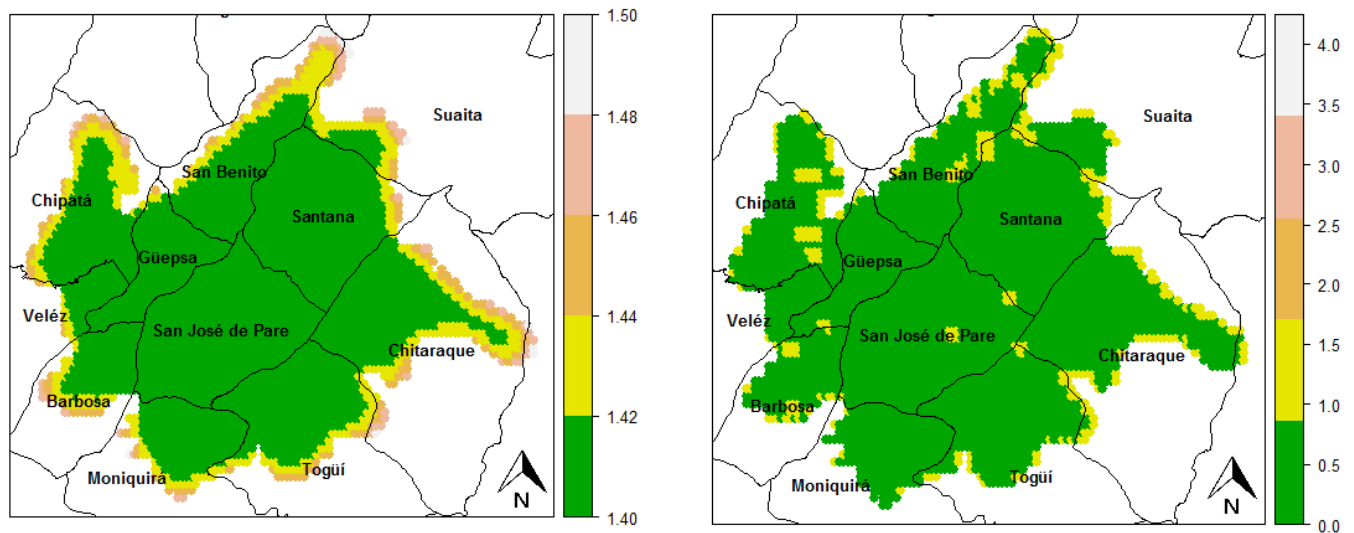


Figura 5-28: Distribución espacial de varianzas de la predicción del kriging del primer (izquierda) y segundo componente principal (derecha).

5.3. Análisis del índice de erodabilidad (K)

Los valores del índice K son en todos los casos menores a 0,01 y no hay observaciones atípicas (Tabla 5-7 y Figura 5-29). Esto indica que la región se clasifica con erodabilidad natural (Tabla 2-4). Lo anterior sugiere que en el área de interés hay alta resistencia a la erosión. No obstante algunos autores han reportado lo contrario (IGAC, 2003, 2005; IDEAM, 2015). En este sentido, cabe resaltar que es posible que el modelo utilizado en la estimación de K (Ramírez et al., 2009) no se ajuste a las condiciones de la HRS (aunque este mismo haya sido evaluado en relieves similares). Otros métodos para estimar K son más precisos (Piscitelli et al., 2010) y pueden dar valores más altos de K pero requieren información adicional no disponible en este trabajo.

Los gráficos de regresión no paramétrica y las correspondientes bandas de confianza para la hipótesis de no efecto definidos en la sección (2.4.2), calculados con la información de K se presentan en la figura (5-30). Estos indican que no hay tendencia espacial, puesto que las curvas de regresión se mantienen en gran medida dentro de las bandas de confianza. De acuerdo con esto, el índice K en esta región puede considerarse como un proceso estacionario.

Tabla 5-7: Medidas de localización y variabilidad de la variable índice de erodabilidad (K). Calculados con datos obtenidos en la HRS en 2016.

MEDIDA	K
Media	0,0015
Mediana	0,0014
Mínimo	-0,0055
Máximo	0,0082
Cuartil inferior	-0,0002
Cuartil superior	0,0033
Desviación estándar	0,0026

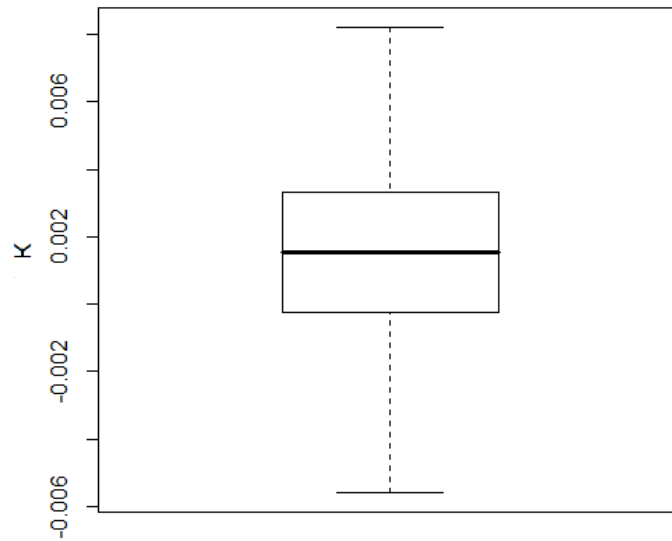


Figura 5-29: Diagrama de caja y bigotes del índice de erodabilidad (K), generados a partir de información medida en la *HRS* en 2016.

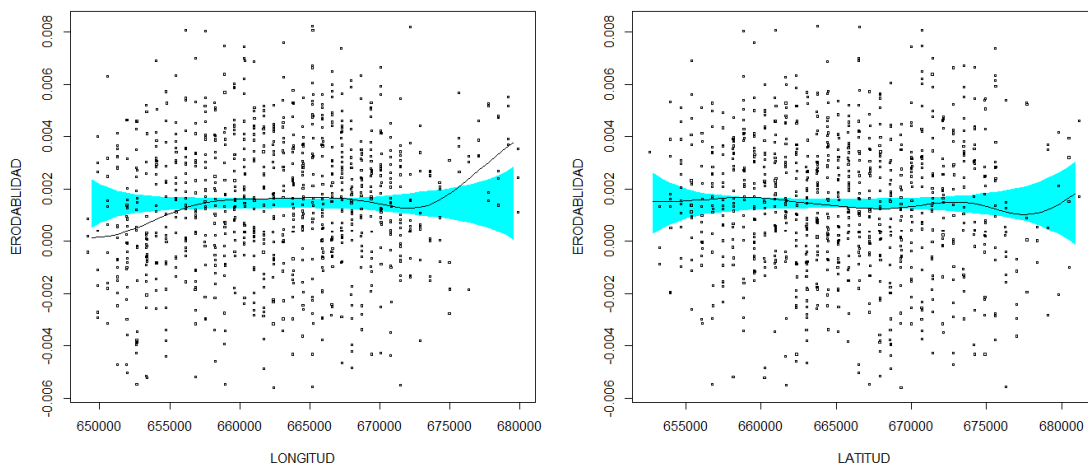


Figura 5-30: Gráfico de dispersión de los valores de erodabilidad (puntos negros), respecto a coordenadas planas (puntos negros). Curva de regresión ajustada por métodos Kernel (línea continua negra). Banda de confianza de no efecto (banda azul). Datos tomados en la *HRS* en 2016.

Los semivariogramas direccionales tienen un patrón similar (Figura 5-31) lo que sugiere isotropía. El semivariograma omnidireccional y en particular el modelo ajustado a este (Figura 5-31), indica que hay poca dependencia espacial ($\hat{\phi} = 609,23$). Este comportamiento difiere del reportado por Addis y Klik (2015) quien encontró una alta dependencia espacial para esta variable en una zona de Etiopía (en ese caso se usó un modelo Gaussiano y K fue calculado con la metodología de Wischmeier et al. (1971)).

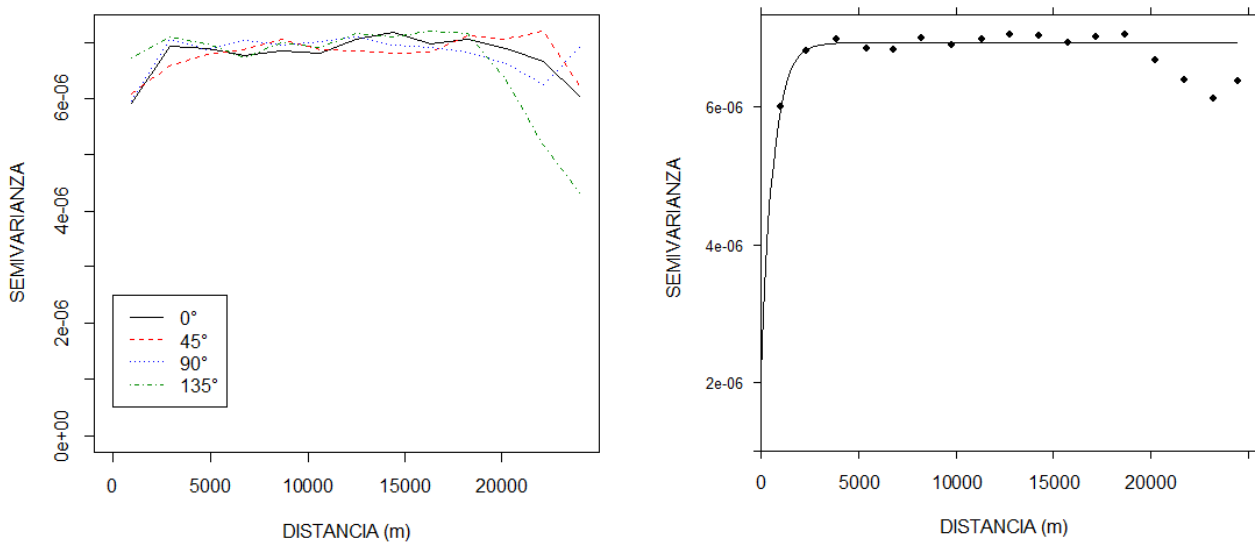


Figura 5-31: Semivariogramas direccionales (0, 45, 90 y 135 grados) (izquierda), y experimental (puntos negros) calculados con datos muestrales con ajuste de modelo teórico exponencial ($\hat{\tau}^2 = 2,1e - 06$, $\hat{\sigma}^2 = 4,8e - 06$ y $\hat{\phi} = 609,23$) (derecha), para el índice de erodabilidad (K).

Los mapas de distribución espacial del índice K se muestran en la figura (5-32). Estos sugieren que en general los valores de K son bajos y que hay muchas micro zonas, es decir que dentro del mismo municipio pueden darse diversas condiciones de riesgo a erodabilidad. Este patrón difiere al reportado en otros estudios, donde se establecen grandes zonas claramente definidas (Ayoubi y Alizadeh, 2005; Addis y Klik, 2015).

Al igual que los análisis individuales y el conjunto, las predicciones de K tienen mejores resultados de validación cruzada comparados con las interpolaciones realizadas (Tabla 5-8). Además, el mapa de las varianzas de predicción (Figura 5-33) como en los casos anteriores,

sugiere que la incertidumbre es homogénea en gran parte del área de estudio.

La zonificación realizada con el índice K no determina regiones de susceptibilidad a erosión. Caso contrario se observa en el análisis individual de variables que genera zonas amplias, mientras que el estudio conjunto delimita áreas de menor extensión en los cuales se facilita dirigir planes orientados a la prevención de la degradación en la HRS . Con el fin de seleccionar y diseñar los métodos de manejo más eficaces para el control del deterioro de la capa arable, es fundamental comprender los factores que fijan estas condiciones (Robins y Burgoa, 2016). Por tal motivo, el siguiente paso después de establecer los territorios de riesgo, es reconocer en campo que situaciones determinan en mayor medida la afectación al suelo.

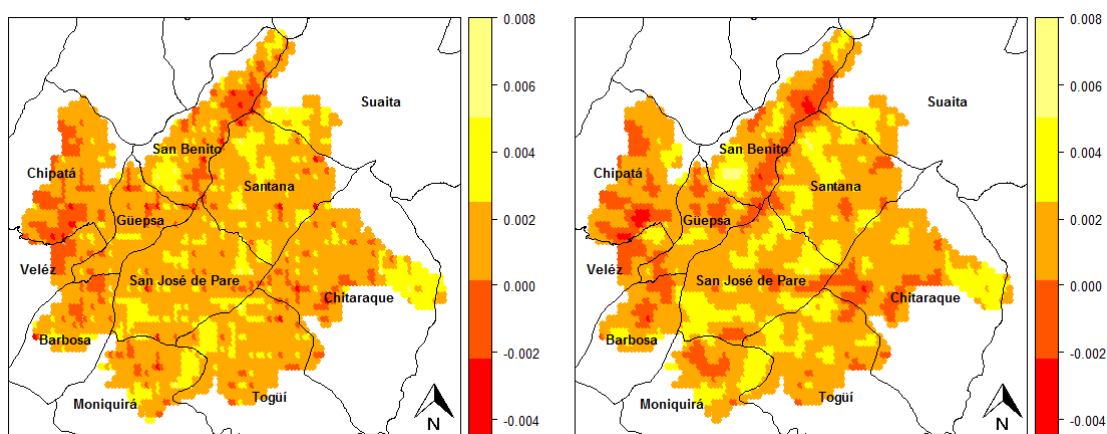


Figura 5-32: Distribución espacial para el índice de erodabilidad (K). Mapas construidos con los métodos del inverso de la distancia (izquierda) y kriging ordinario (derecha).

Tabla 5-8: Estadísticos de validación cruzada (Error medio de predicción EMP , raíz del error cuadrado medio de predicción $RECM$ y coeficiente de determinación R^2) de los análisis de predicción (Kriging ordinario KO) e interpolación (inverso de la distancia DIP), realizados con el índice de erodabilidad (K).

VARIABLE	EMP	$RECM$	R^2
ERODABILIDAD DIP	-0,0001	0,0025	0,9150
ERODABILIDAD KO	-0,0001	0,0024	0,9771

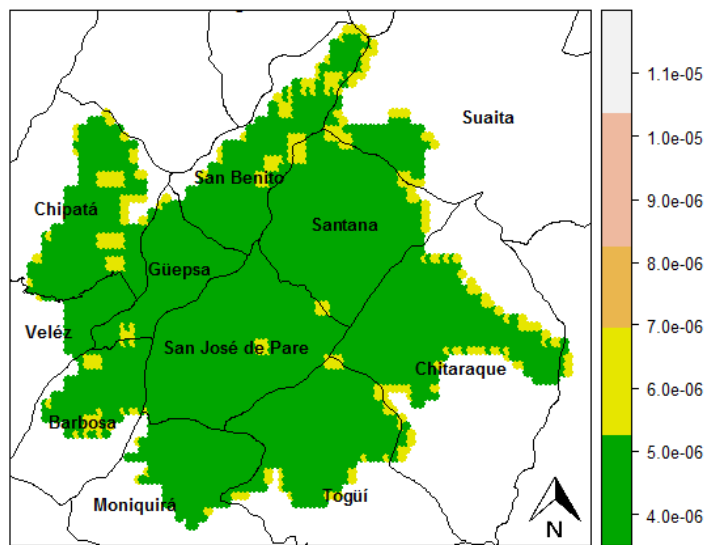


Figura 5-33: Distribución espacial de varianzas de la predicción de kriging ordinario del índice de erodabilidad (K).

6 Conclusiones

En la Hoya del Río Suárez se identificaron las áreas al sur de San Benito (en el departamento de Santander) y la parte oriental de los municipios de San José de Pare y Chitaraque (ambos en el departamento de Boyacá) como los más susceptibles a erosión. Esto se debe a la menor estabilidad estructural y altos contenidos de arena que presentan los suelos de las regiones mencionadas. Estas condiciones favorecen el desprendimiento de los agregados del terreno y su eventual degradación.

Las regiones colindantes entre los municipios de Chitaraque con Santana y Togüí (departamento de Boyacá) son áreas de menor susceptibilidad a la erosión. Estos sitios presentan valores bajos de densidad aparente y real con mayor retención de humedad. Además, los suelos de los territorios referidos anteriormente se caracterizan por ser estables.

De acuerdo con los resultados de las metodologías de predicción consideradas en el análisis (en especial el uso de técnicas geoestadísticas) se evidenció que contribuyen a la explicación de zonas susceptibles a erosión en la región y son herramientas útiles en la predicción y posterior interpretación del comportamiento de la física del suelo.

El análisis espacial de los componentes principales generados, a partir de las variables del suelo, permitió obtener una mejor caracterización de la susceptibilidad a la erosión, que los establecidos por medio del estudio individual de cada una de ellas y a través del índice de erodabilidad. Lo anterior sugiere que el uso combinado de técnicas multivariadas y métodos de análisis espacial puede facilitar la descripción del patrón de variación de las propiedades físicas del suelo.

7 Recomendaciones

Para futuros estudios sería conveniente considerar las propiedades químicas del suelo como el porcentaje de materia orgánica, pH, entre otras, ya que son indicadores de fertilidad y estabilidad de las zonas de cultivo.

Se sugiere que al realizar el análisis espacial de las variables físicas, los puntos de muestreo se ubiquen a una menor distancia teniendo en cuenta la alta heterogeneidad de los terrenos.

En la actualidad el país no cuenta con programas dedicados a la prevención y manejo de los suelos en la *HRS*. Se hace necesario por lo tanto continuar con estudios que visualicen la amenaza de la erosión en los terrenos, y el uso de herramientas estadísticas es fundamental en la predicción y modelación de estos fenómenos.

Bibliografía

- Addis, H. y Klik, A. (2015). Predicting the spatial distribution of soil erodibility factor using usle nomograph in an agricultural watershed, ethiopia. *International Soil and Water Conservation Research*, 3(4):282–290.
- Ahumada, J. (2015). Efectos de la Tecnología Utilizada en la Producción Panelera en las Transformaciones Ambientales de la Hoya del Río Suárez. Tesis de Maestría, Universidad Nacional de Colombia - Sede Bogotá. Colombia.
- Alarcón, S. y Reyes, A. (2013). Evaluación de la Erodabilidad de los Suelos Typic Dystrudepts, Typic Hapludands y Andic Dystrudepts Mediante Simulador de Lluvia en la Microcuenca La Centella (Dagua - Valle del Cauca). *EIDENAR*, 12(1):59–67.
- Aldás, J. y Jiménez, E. (2017). *Análisis Multivariante Aplicado con R*. Ediciones Paraninfo, S. A. Madrid, España.
- Anderson, T. (2003). *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, Ney Jersey. United States.
- Anschau, R., Flores, N., Carballo, S. y Hilbert, J. (2009). Evaluación del potencial de producción de biocombustibles en Argentina, con criterios de sustentabilidad social, ecológica y económica, y gestión ordenada del territorio. el caso de la caña de azúcar y el bioetanol. Informe Técnico, Instituto Nacional de Tecnología Agropecuaria, Buenos Aires, Argentina.
- Ayoubi, S. y Alizadeh, M. (2005). Assessment spatial variability of soil erodibility by using of geostatistic and gis (case study mehr watershed of sabzevar). *Sociedade & Natureza*, 1(1):78–86.
- Bivand, R., Pebesma, E. y Gomez, V. (2013). *Applied Spatial Data Analysis with R*. Springer. New York. United States.
- Blanco, R. (2009). La relación entre la densidad aparente y la resistencia mecánica como indicadores de la compactación del suelo. *Agrociencia. Texcoco, México.*, 43(3):231–239.

- Bosque, J. (1992). *Sistemas de Información Geográfica*. Ediciones Rialp S. A. Madrid, España.
- Bowman, A. y Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis. The Kernel Approach With S-Plus Illustrations*, volume 18. Oxford University Press Inc., New York. United States.
- Bowman, A. W. y Azzalini, A. (2014). *R package sm: nonparametric smoothing methods (version 2.2-5.4)*. University of Glasgow, UK and Università di Padova, Italia.
- Bravo, C., Cabrera, J., M, C., Cánchica, H., Ramírez, E. y González, I. (2011). Variabilidad espacial de propiedades físicas y químicas del suelo y su relación con el rendimiento de frijol (*vigna unguiculata l.*) en una parcela experimental de los llanos centrales venezolanos. *Estudios en la Zona no Saturada del Suelo. Salamanca, España*, 10(ZNS'11):79–84.
- Briggs, L. y Shantz, H. (1911). A wax seal method for determining the lower limit of available soil moisture. *Botanical gazette. University of Chicago, United States*, 51(3):210–219.
- Buckman, H. y Brady, N. (1966). *The Nature and Properties of Soils*. The Macmillian Company. New York, United States.
- Buol, S., Sánchez, P., Cate, R. y Granger, M. (1974). Clasificación de suelos en base a su fertilidad. *Manejo de Suelos en la América Tropical. Cali, Colombia.*, 1(1):79–92.
- Burbano, H. y Silva, F. (2010). *Ciencia del Suelo: Principios Básicos*. Sociedad Colombiana de la Ciencia del Suelo. Bogotá, D.C. Colombia.
- Burgess, T. y Webster, R. (1980). Optimal interpolation and isarithmic mapping of soil properties. *Journal of Soil Science*, 31(2):315–331.
- Ceddia, M., Vieira, S., Villela, A., Mota, L., Anjos, L. y Carvalho, D. (2009). Topography and spatial variability of soil physical properties. *Scientia Agricola*, 66(3):338–352.
- Cook, H. (1937). The nature and controlling variables of the water erosion process. *Soil Science Society of America Journal*, 1(C):487–494.
- CORPOBOYACÁ (2014). Actualización del plan de ordenación y manejo de la cuenca hidrográfica del río medio y bajo suárez. Informe Técnico, Corporación Autónoma Regional de Boyacá (CORPOBOYACÁ).

- CORPOICA (2016). *Recomendaciones Técnicas Preliminares de Manejo de Suelos en Ladera para el Sistema de Producción de Caña Panelera en la Hoya del Río Suárez*. Corporación Colombiana de Investigación Agropecuaria. Mosquera, Colombia.
- Cressie, N. (1993). *Statistics for Spatial Data*. John Wiley & Sons, Inc. New York, United States.
- Delicado, P. (2008). *Curso de Modelos no Paramétricos*. Departamento de Estadística Investigación Operativa. Universidad Politécnica de Cataluña. España.
- Duque, G. y Escobar, C. (2002). *Mecánica de los Suelos*. Universidad Nacional de Colombia. Sede Manizales. Colombia.
- Eubank, R. (1999). *Nonparametric Regression and Spline Smoothing*. Marcel Dekker, Inc. New York, United States.
- Ferreras, L., Toresani, S., Faggioli, V. y Galarza, C. (2015). Sensibilidad de indicadores biológicos edáficos en un argiudol de la región pampeana argentina. *Spanish Journal of Soil Science*, 5(3):220–235.
- Gallardo, A. (2006). Geoestadística. *Ecosistemas.*, 15(3):48–58.
- García, H., Albarracín, L., Toscano, A., Santana, N. y Insuasty, O. (2007). *Guía Tecnológica para el Manejo Integral del Sistema Productivo de Caña Panelera*. Produmedios. Bogotá, D. C. Colombia.
- García, A. (2013). Manejo de la degradación del suelo. *Sociedad Colombiana de la Ciencia del Suelo, Bogotá, Colombia*.
- Garmin (2011). *Manual del usuario de eTrex*. Garmin Ltda. Kansas, United States.
- Ghaemi, M., Astaracai, A., Emami, H., Nassiri Mahalati, M. y Sanaeinejad, S. (2014). Determining soil indicators for soil sustainability assessment using principal component analysis of astan quds-east of mashhad-iran. *Journal of Soil Science and Plant Nutrition*, 14(4):1005–1020.
- Giménez, R. (2004). Física del suelo. *Cátedra de Edafología. Facultad de Agronomía. Universidad Nacional de Tucumán. Tucumán, Argentina*.
- Gómez, J. y Osorio, M. (2016). *Manual de Prácticas de Campo y del Laboratorio de Suelos*. Servicio Nacional de Aprendizaje. Bogotá, Colombia.

- Hengl, T. (2007). *A Practical Guide to Geostatistical Mapping of Environmental Variables*. Institute for Environment and Sustainability. Ispra, Lombardía. Italia.
- Henríquez, C., Killorn, R., Bertsch, F. y Sancho, F. (2005). La geostatística en el estudio de la variación espacial de la fertilidad del suelo mediante el uso del interpolador kriging. *Agronomía Costarricense*, 29(2):73–81.
- Hillel, D. (1998). *Environmental Soil Physics*. Academic Press. San Diego, United States.
- Hincapié, E. y Ramirez, F. (2013). Riesgo a la erosión en suelos de ladera de la zona cafetera. Informe Técnico, Centro Nacional de Investigaciones de Café (CENICAFÉ). Manizales. Colombia.
- IDEAM (2015). *Síntesis del Estudio Nacional de la Degradación de Suelos por Erosión en Colombia*. Instituto de Hidrología, Meteorología y Estudios Ambientales - IDEAM. Bogotá D. C., Colombia.
- IGAC (2003). *Estudio General de Suelos y Zonificación de Tierras. Santander*. Instituto Geográfico Agustín Codazzi. Colombia.
- IGAC (2005). *Estudio General de Suelos y Zonificación de Tierras del Departamento de Boyacá*. Instituto Geográfico Agustín Codazzi. Colombia.
- Insuasty, O., Manrique, R. y Palacio, O. (2003). *Catálogo de Variedades de Caña para la Producción de Panela en la Hoya del Río Suárez*. Litografía La Bastilla Ltda. Bucaramanga, Colombia.
- IPBES (2018). *Plataforma Intergubernamental Científico-normativa sobre Diversidad Biológica y Servicios de los Ecosistemas*. Recuperado de: <https://www.ipbes.net/>.
- James, G., Witten, D., Hastie, T. y Tibshirani, R. (2013). *An Introduction to Statistical Learning*, volume 112. Springer, New York. United States.
- Jaramillo, D., Anaya, M., Restrepo, C., González, H. y Álvarez, F. (2011). Variables físicas que explican la variabilidad de suelo aluvial y su comportamiento espacial. *Pesquisa Agropecuaria Brasileira*, 46(12):1707–1715.
- Johnson, L. (1987). Soil loss tolerance: Fact or myth? *Journal of Soil and Water Conservation*, 42(3):155–160.
- Kohnke, H. (1968). *Soil Physics*. McGraw-Hill Publishing. Bombay, New Delhi.

- Kooch, Y., Jalilvand, H., Bahmanyar, M. y Pormajidian, M. (2008). The use of principal component analysis in studying physical, chemical and biological soil properties in southern caspian forests (north of iran). *Pakistan Journal of Biological Sciences*, 11(3):366–372.
- Krige, D. (1966). *Two-dimensional weighted moving average trend surfaces for ore evaluation*. South African Institute of Mining and Metallurgy Johannesburg.
- Lal, R. y Shukla, M. (2004). *Principles of Soil Physics*. Marcel Dekker, Inc. New York. United States.
- León, J. (2001). *Estudio y Control de la Erosión Hídrica*. Universidad Nacional de Colombia. Medellín, Colombia.
- López, A. (2005). Manual de Edafología. *Departamento de Cristalografía, Mineralogía y Química. Agrícola de la Universidad de Sevilla. España*.
- Luna, C. (2006). Aumento de la productividad de caña de azúcar por unidad de área cultivada. *Asociación de Técnicos Cultivadores de Caña. TECNICAÑA*, 10(8):4–19.
- Maestre, R. B. y de Andreu, A. F. (2018). Variabilidad espacial de las propiedades físicas de dos suelos cultivados con palma aceitera en el estado monagas, venezuela— spatial variability of physical properties of two soils cultivated with oil palm in monagas state, venezuela. *UDO Agrícola*, 9(4):912–924.
- Manrique, R., Insuasty, O., Mora, C., Rodríguez, G., Blanco, R., Mejía, L., Pinto, J. y Sandoval, G. (2000). *Manual de Caña de Azúcar para la Producción de Panela*. Litografía La Bastilla Ltda. Bucaramanga, Colombia.
- Matheron, G. (1963). Principles of geostatistics. *Economic geology*, 58(8):1246–1266.
- Melo, C., Santacruz, A. y Melo, O. (2012). *geospt: An R package for spatial statistics*. R package version 1.0-0.
- Monroy, F., Álvarez, J. y Alvarado, O. (2017). Distribución espacial de algunas propiedades físicas del suelo en un transecto de la granja tunguavita, paipa. *Revista UDCA Actualidad & Divulgación Científica*, 20(1):91–100.
- Montenegro, H. y Malagón, D. (1990). *Propiedades Físicas de los Suelos*. Instituto Geográfico Agustín Codazzi. Bogotá, D. C. Colombia.

- Montgomery, D., Peck, E. y Vining, G. (2006). *Introducción al Análisis de Regresión Lineal*. Compañía Editorial Continental, San Juan Tlihuaca. México.
- Moreno, M. (2011). Variabilidad espacial de las propiedades físicas, químicas y biológicas de dos suelos agrícolas de la provincia de quebec, Canadá. *Revista de la Facultad de Agronomía*, 28(2):223–241.
- Morgan, R. (1986). *Soil Erosion and Conservation*. John Wiley & Sons, Inc. New York, United States.
- Mousavi, S., Sarmadian, F., Dehghani, S., Sadikhani, M. y Taati, A. (2017). Evaluating inverse distance weighting and kriging methods in estimation of some physical and chemical properties of soil in Qazvin plain. *Eurasian Journal of Soil Science*, 6(4):327.
- Narro, E. (2004). *Física De Suelos*. Editorial Trillas. México, D. F. México.
- Novillo, I., Carrillo, M., Cargua, J., Nabel, V., Albán, K. y Morales, F. (2018). Propiedades físicas del suelo en diferentes sistemas agrícolas en la provincia de los ríos, Ecuador. *Temas agrarios*, 23(2):177–187.
- Oliver, M. y Webster, R. (2015). *Basic Steps in Geostatistics: The Variogram and Kriging*. Springer, New York. United States.
- Osorio, G. (2007). *Manual Técnico Buenas Prácticas Agrícolas (BPA) y Buenas Prácticas de Manufactura (BPM) en la Producción de Caña y Panela*. CTP Print Ltda. Medellín, Colombia.
- Pebesma, E. (2004). Multivariable geostatistics in R: the gstat package. *Computers & Geosciences*, 30:683–691.
- Pebesma, E. y Bivand, R. (2005). Classes and methods for spatial data in R. *R News*, 5(2):9–13.
- Peláez, J. (2003). Métodos experimentales para el seguimiento y estudio de la erosión hídrica. *Instituto de Estudios Políticos y Relaciones Internacionales de la Universidad Nacional de Colombia. Medellín, Colombia.*, 6(1):39–56.
- Piscitelli, M., Varni, M., Sfeir, A. y Ares, G. (2010). Valores de erodabilidad (factor k-usle) para suelos representativos de la cuenca alta del arroyo azul (partido de azul-buenos aires). *Revista de la Facultad de Agronomía*, 109.

- Pla, L. (1986). *Análisis Multivariado: Método de Componentes Principales*. Secretaría General. Programa Regional de Desarrollo Científico y Tecnológico. OEA, Washington, DC (United States).
- Prado, W. y da Vega, M. (1992). Relación entre erosión y pérdida de fertilidad del suelo. *Memoria Taller Erosión de Suelos en América Latina. Chile. FAO.*, 1(1):55–79.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramírez, R. (1997). *Propiedades Físicas, Químicas y Biológicas de los Suelos*. Fenalce - Sena. Primera edición, Bogotá, Colombia.
- Ramírez, F., Hincapie, E. y Sadeghian, S. (2009). Erodabilidad de los suelos de la zona central cafetera del departamento de Caldas. *Centro Nacional de Investigaciones de Café.*, 60(1):58–71.
- Ribeiro, J. y Diggle, P. (2016). *geoR: Analysis of geostatistical data*. R package version 1.7-5.2.
- Rivera, J., Lal, R., Amézquita, E., Mesa, O. y Chaves, B. (2010). Predicción de la erodabilidad en algunos suelos de ladera de la zona cafetera colombiana. *Centro Nacional de Investigaciones de Café.*, 61(4):344–357.
- Robins, P. y Burgoa, B. (2016). *Guía de Prácticas para el Manejo de Erosión y Escorrentía Agrícola en Laderas*. Resource Conservation District of Monterey County (RCDMC). Salinas, California. 52p.
- Rodríguez, G., García, H., Roa, Z. y Santacoloma, P. (2004). *Producción de Panela como Estrategia de Diversificación en la Generación de Ingresos en Áreas Rurales de América Latina*. Food and Agriculture Organization of the United Nations, Roma. Italia.
- Rodríguez, G. y Gottret, M. (1998). *Aprendiendo del Pasado para Proyectarnos Hacia el Futuro: Adopción e Impacto de la Tecnología de Panela en la Hoya del Río Suárez y Cundinamarca (Colombia)*. Reporte Técnico, Centro Internacional de Agricultura Tropical (CIAT), Proyecto de Desarrollo de Agroempresas Rurales. Palmira, Colombia.
- Rucks, L., García, F., Kaplán, A., Ponce de León, J. y Hill, M. (2004). *Propiedades Físicas del Suelo*. Universidad de la República. Facultad de Agronomía. Montevideo, Uruguay.

- Rueda, F., Peñaranda, L., Velásquez, W. y Díaz, S. (2015). Aplicación de una metodología de análisis de datos obtenidos por percepción remota orientados a la estimación de la productividad de caña para panela al cuantificar el ndvi (índice de vegetación de diferencia normalizada). *Revista Corpoica: Ciencia y Tecnología Agropecuaria*, 16(1):25–40.
- Salamanca, A. y Sadeghian, S. (2006). La densidad aparente y su relación con otras propiedades en suelos de la zona cafetera colombiana. *Centro Nacional de Investigaciones de Café.*, 56(4):381–397.
- Schabenberger, O. y Gotway, C. (2005). *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC. Boca Raton, Florida. United States.
- Scholes, R., Montanarella, L., Brainich, A., Barger, N., ten Brink, B., Cantele, M., Erasmus, B., Fisher, J., Gardner, T., Holland, T. et al. (2018). Summary for policymakers of the thematic assessment report on land degradation and restoration. *Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services IPBES. Bonn, Alemania.*, pages 1–31.
- Silva, S. y Correa, F. (2009). Análisis de la contaminación del suelo: revisión de la normativa y posibilidades de regulación económica. *Semestre Económico*, 12(23):13–34.
- Soto, A. y López, L. (2005). *Análisis y Síntesis en Cartografía: Algunos procedimientos*. Universidad Nacional de Colombia. Departamento de Geografía. Bogotá, Colombia.
- Srivastava, M. (2002). *Methods of Multivariate Statistics*. John Wiley & Sons, Ney Jersey. United States.
- Torres, D., Álvarez, J., Contreras, J., Henríquez, M., Hernández, W., Lorbes, J. y Mogollón, J. (2017). Identificación de potencialidades y limitaciones de suelos agrícolas del Estado Lara, Venezuela. *Bioagro*, 29(3):207–218.
- Valero, A. (1999). *Capacidad de Retención de Agua Disponible (CRAD) para las plantas en suelos del Alto y Medio Aragón*. Escuela Politécnica Superior de Huesca. España.
- Vargas, R., Serrato, F. y Trujillo, A. (2015). Variabilidad espacial de las propiedades físicas de un suelo fluventic ustropepts en la cuenca baja del río las Ceibas-Huila. *Ingeniería y Región*, 1(13):113–123.
- Varón, V., Camacho, J. y Ordoñez, N. (2014). Estudio de la humedad aprovechable del suelo y su relación con algunos atributos físicos en un inceptisol. *Memorias XI Congreso Latinoamericano y del Caribe de Ingeniería Agrícola*.

- Vásquez, J., Baena, D. y Menjivar, J. (2010). Variabilidad espacial de propiedades físicas y químicas en suelos de la granja experimental de la universidad del magdalena (santa marta, colombia). *Acta Agronómica*, 59(4).
- Vauclin, M., Vieira, S., Bernard, R. y Hatfield, J. (1982). Spatial variability of surface temperature along two transects of a bare soil. *Water Resources Research*, 18(6):1677–1686.
- Veihmeyer, F. y Hendrickson, A. (1931). The moisture equivalent as a measure of the field capacity of soils. *Soil Science. University of California, United States.*, 32(3):181–194.
- Villatoro, M., Henríquez, C. y Sancho, F. (2008). Comparación de los interpoladores IDW y kriging en la variación espacial de pH, Ca, CICE y P del suelo. *Agronomía Costarricense*, 32(1).
- Waller, L. y Gotway, C. (2004). *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons, Inc., Publication. New Jersey. United States.
- Wand, M. y Jones, M. (1994). *Kernel Smoothing*. Chapman and Hall/CRC. Boca Ratón, Florida. United States.
- Warrick, A. (2002). *Soil Physics Companion*. CRC press LLC. Boca Ratón, Florida. United States.
- Wasserman, L. (2007). *All of Nonparametric Statistics*. Springer, New York. United States.
- Weisberg, S. (2014). *Applied Linear Regression*. John Wiley & Sons, Inc., New Jersey. United States.
- Wischmeier, W., Johnson, C. y Cross, B. (1971). Soil erodibility nomograph for farmland and construction sites. *Journal of soil and water conservation*, 71:189–191.
- Wischmeier, W. y Smith, D. (1978). *Predicting Rainfall Erosion Losses*. Department of Agriculture, Handbook N 537. Washington, D. C., United States.
- Zamora, T. (2010). Métodos de Regresión No Paramétrica en Muestreo en Poblaciones Finitas. Tesis de Maestría, Universidad de Granada. España.