

# Desarrollo de una aplicación en Python para mapeo de variabilidad espacial en la agricultura digital

## Development of an Application in Python for Mapping of Spatial Variability in Digital Agriculture

Sebastián Saavedra-Rincón\* ORCID: 0000-0001-6911-2818

Óscar Chaparro-Anaya ORCID: 0000-0001-5834-0786

Universidad Nacional de Colombia – Sede Palmira. Palmira, Valle del Cauca, Colombia.

\*Autor de correspondencia: ssaavedrar@unal.edu.co

### Resumen

**Objetivo:** Desarrollar un sistema de información geográfica portable con el uso de módulos especializados para el análisis de datos en el ecosistema Python y comparar sus resultados con el motor geostatístico SAGA GIS para validar el sistema. **Métodos:** Se desarrolló la aplicación GeoStatLiteUN v2021.6.1 implementando los módulos especializados: SciPy v1.20.3, NumPy v1.20.0, statsmodels v0.12.2, PyKrige v1.6.0, Matplotlib.pyplot v3.4.2, seaborn v0.11.1 y Tkinter de Python (Python 3.9.5) en un *script* que, bajo la lógica de la geostatística, automatizó los procedimientos repetitivos para analizar la variabilidad espacial (dentro y fuera de los lotes) mediante interpolación Kriging ordinario, cuyo ejecutable posibilita a los usuarios procesar y analizar datos espaciales y construir mapas de variabilidad espacial de forma ágil. **Resultados:** GeoStatLiteUN al ser probado mostró que los resultados teóricos son válidos si se tiene en cuenta la similitud con los resultados de SAGA GIS (para los mismos datos) y que 99 de 100 (99% de confianza en el método) modelos espaciales construidos con GeoStatLiteUN a partir de las bases de datos

### Abstract

**Objective:** To develop a portable geographic information system with the use of specialized modules for data analysis in the Python ecosystem and to compare its results with the SAGA GIS geostatistical engine to validate the system. **Methods:** The GeoStatLiteUN v2021.6.1 application was developed implementing the specialized modules: SciPy v1.20.3, NumPy v1.20.0, statsmodels v0.12.2, PyKrige v1.6.0, Matplotlib.pyplot v3.4.2, seaborn v0.11.1 and Python Tkinter (Python 3.9.5) in a script that, under the logic of geostatistics, automated repetitive procedures to analyze spatial variability (inside and outside of cropfields) using ordinary Kriging interpolation, which executable allows users to process and analyze spatial data and construct spatial variability maps in an agile way. **Results:** GeoStatLiteUN when tested showed that the theoretical results are valid taking into account the similarity with the SAGA GIS results (for the same data) and that 99 out of 100 (99% confidence in the method) spatial models built with GeoStatLiteUN from the databases for testing, they will base their predictions on the sample means. **Conclusion:**

de prueba basarán sus predicciones en las medias de las muestras. *Conclusión:* La contribución innovadora de GeoStatLiteUN v2021.6.1 es la integración de protocolos automáticos para la depuración basada en la prueba de Tukey para atípicos leves y valida la estacionariedad de sentido amplio; además se comparte como un archivo portable que facilita el uso del sistema y el método geoestadístico Kriging ordinario a los profesionales del agro en el campo, sin requerir conocimientos previos en los algoritmos matemáticos a usar.

### Palabras clave

Agrícola, geoestadística, interpolación, Kriging, SIG.

The innovative contribution of GeoStatLiteUN v2021.6.1 is the integration of automatic debugging protocols based on the Tukey test for mild outliers and the validation of broad sense stationarity; In addition, it is shared as a portable file that will facilitate the use of the ordinary Kriging system and the geostatistical method for agricultural professionals in the field without requiring prior knowledge of the mathematical algorithms to be used.

### Keywords

Agricultural, geostatistics, interpolation, Kriging, GIS.

## Introducción

La cuarta revolución industrial está siendo involucrada en el ámbito de la producción agrícola con el desarrollo del Big Data, el Internet de las cosas, la robótica y la inteligencia artificial; el término agricultura digital surge cuando se utiliza el conocimiento científico, las tecnologías de la información y de la comunicación para obtener un desarrollo sostenible y competitivo en los sistemas de producción agrarios; en este sentido, la agricultura digital es una modernización del concepto de agricultura de precisión, siendo esta última una estrategia de gestión que recopila, procesa y analiza datos temporales y espaciales de los agroecosistemas para mejorar la eficiencia en el uso de los recursos, la productividad, la calidad, la rentabilidad y la sostenibilidad de la producción agraria.<sup>1</sup>

En la etapa de recolección, los elementos físicos más modernos involucrados en los sistemas de producción (tractores, sembradoras, cosechadora, abonadoras, etcétera) generan grandes cantidades de datos que se transfieren a través de redes inalámbricas y se almacenan en bases de datos. Si el procesamiento y análisis de estas bases de datos involucran un hardware potente y softwares que permitan usar la programación para manipular rápidamente la información, de tal forma que sea apta para aplicar técnicas estadísticas o geoestadísticas que respalden las decisiones de gestión de los sistemas de producción agrícola, es entonces cuando se puede hablar de agricultura digital (Kamilaris *et al.*, 2017; Bassoi *et al.*, 2019).

Un programa ampliamente usado (por ser de libre acceso) para procesar y analizar los datos es Quantum GIS (QGIS) —un sistema de información geográfica profesional que está construido sobre software libre y de código abierto—, el cual usa un sistema

---

<sup>1</sup> Adaptación de la definición de agricultura de precisión publicada por la International Society of Precision Agriculture (ISPA) en: <https://www.ispag.org/about/definition>

automatizado para el análisis geoestadístico agrupados en módulos llamado SAGA GIS; aunque se pueden encontrar otros programas con una amplia gama de aplicaciones para la agricultura como ArcGIS (software privativo) o GRASS GIS (código abierto).

La técnica usada comúnmente para la interpolación espacial de variables geofísicas se conoce como Kriging ordinario y se puede definir como el mejor interpolador lineal óptimo insesgado desde la estadística convencional, que permite hacer inferencias al calcular el error de estimación de las variables (modelo probabilístico) cuando el fenómeno que se estudia es estacionario y de media desconocida.

Gutiérrez-López *et al.* (2012) explican de forma detallada el método Kriging y concluyen que es el privilegiado cuando la precisión de los resultados es lo importante, así como para conocer el error de estimación.

El método Kriging ordinario se usa en los estudios sobre la variabilidad espacial de propiedades físicas de los suelos agrícolas o parámetros de las operaciones de mecanización agrícola sobre suelos gestionados; por citar unos ejemplos, Yéscas-Coronado *et al.* (2018), Loayza *et al.* (2020) y Salgado-Velázquez *et al.* (2020) aplican la técnica para estimar la variabilidad espacial del carbono orgánico e inorgánico del suelo en la Comarca Lagunera, México; del carbono orgánico en suelos de Ecuador y de propiedades fisicoquímicas en un suelo cultivado con caña de azúcar en el sureste de México, respectivamente.

La generación de grandes cantidades de datos (espaciales y temporales) sobre los parámetros que influyen en los sistemas productivos agrícolas (suelo, agua, insumos, maquinas, etcétera) trae consigo el reto de desarrollar herramientas informáticas que los procesen y analicen de forma rápida para la generación de estrategias y escenarios con óptimas opciones por unidad de área para una adecuada toma de decisiones, una forma de crear estas herramientas es a través del uso de lenguajes de programación para automatizar tareas repetitivas propias de la gestión de los datos (Cravero y Sepúlveda, 2021).

Con el uso de lenguajes de programación, como lo es Python, se pueden generar *scripts* que con un solo comando se depuran bases de datos para someterlas a posteriores técnicas de la geoestadística, como el Kriging ordinario; sin duda, es una opción que eficientizaría el trabajo de los agricultores y los profesionales del agro (Kim *et al.*, 2016).

Lago-González *et al.* (2011) desarrollaron una aplicación —de las primeras en su tipo— para la creación de mapas de rendimiento en las consolas TechAgro integradas a cosechadoras, de uso amigable para cualquier usuario, incluso para quienes no tengan conocimientos previos de los algoritmos matemáticos.

Posteriormente, Herrera-Gómez y Reinoso-Gutiérrez (2016) desarrollaron una aplicación web en Java para el monitoreo y supervisión de variables espaciotemporales (sin ningún tipo de procesamiento geoestadístico), medidas por sensores en campo, que permite a los agricultores conocer con certeza la condiciones en la que se encuentran sus cultivos en tiempo real, las 24 horas del día.

Así mismo, Paccioretti *et al.* (2020) mostraron una experiencia exitosa en el desarrollo de una aplicación web para mapeo de variabilidad espacial en agricultura y ambiente, con una interfaz amigable que utiliza el lenguaje de programación R, el cual mejoró la

zonificación de métodos convencionales de clasificación y se delimitaron zonas con mayores diferencias en variables de suelo y topografía.

Los objetivos de este trabajo fueron desarrollar una aplicación portable con interfaz gráfica de usuario (IGU) que utiliza módulos especializados para el análisis de datos y presentación de los resultados en el ecosistema Python; bajo la lógica de la geoestadística se automatizaron los procedimientos repetitivos para analizar la variabilidad espacial (dentro y fuera de los lotes) y posteriormente generar mapas mediante interpolación Kriging ordinario de parámetros asociados a operaciones de mecanización agrícola con covariables espaciales, de forma ágil y sin depender de una conexión a internet, que facilitan el uso de estos métodos en campo a los profesionales del agro que no cuentan con los conocimientos previos de los algoritmos matemáticos, y posteriormente comparar los modelos de superficie continua producidos por la aplicación desarrollada con los mapas de variabilidad construidos por defecto, con un modelo semejante a partir del motor geoestadístico SAGA GIS que se usa en el software QGIS para validar el sistema.

## Materiales y métodos

### Desarrollo de la aplicación

Se desarrolló la aplicación GeoStatLiteUN con IGU implementando módulos especializados del lenguaje Python (Python 3.9.5), en un *script* que sigue la lógica de análisis de bases de datos espaciales con el método de interpolación Kriging ordinario (estudio descriptivo, depuración, supuesto de estacionariedad, análisis estructural y creación de superficie continua); para ello lee las bases de datos estandarizadas (archivos estructurados y separados por coma en extensión .csv con tres columnas de interés, que deben ser nombradas con identificadores DV, X y Y, siendo la variable a estudiar y sus covariables espaciales, respectivamente), posteriormente aplica pruebas de la estadística descriptiva, muestra histogramas y gráficos cuantil a cuantil (Q-Q) de probabilidad normal, aplica protocolos para depuración (según la prueba de Tukey para atípicos leves), genera una regresión lineal de la variable de interés con covariables espaciales y crea un modelo de semivariograma y un mapa de superficie continua con interpolación Kriging ordinario (Córdoba *et al.*, 2012; Jaramillo-Jaramillo, 2012; Fernández y Ribes-Dasi, 2014).

La lectura de la base de datos estandarizada se programó con el módulo *Pandas 1.2.4*; las pruebas de la estadística descriptiva, la depuración y la regresión lineal multivariada se programaron con el módulo *SciPy 1.20.3*, *NumPy 1.20.0* y *Statsmodels 0.12.2*; los protocolos para el semivariograma y la interpolación Kriging ordinaria se programaron con el módulo *PyKrige 1.6.0*; los histogramas y los gráficos Q-Q se programaron con los módulos *Matplotlib.pyplot 3.4.2* y *Seaborn 0.11.1* y la IGU se desarrolló con el paquete *Tkinter* de Python.

### Bases de datos de prueba

Se utilizaron para procesar en GeoStatLiteUN y SAGA GIS dos bases de datos correspondientes a suelos gestionados para la producción de caña de azúcar (*Saccharum*

*officinarum*). Las bases de datos respondieron a los resultados de una operación de fertilización mecánica a tasa variada en un lote de seis hectáreas (con 3 630 datos asociados) y una operación de roturación vertical con cincel de cuatro brazos (dos anteriores de 40 cm de profundidad y dos posteriores de 60 cm de profundidad) en un lote de 27 hectáreas (con 12 318 datos asociados). Las variables de interés fueron: dosis de nitrógeno entregado por hectárea ( $\text{kg}_{\text{nitrógeno}}/\text{ha}$ ) y profundidad alcanzada (cm); a cada una se les asignó atributos espaciales en coordenadas proyectadas del Datum UTM zona 18N, que fueron capturadas durante las operaciones a través de un sistema global de navegación por satélite con corrección de doble frecuencia.

## Análisis estadístico

### Procesamiento de datos

En el programa GeoStatLiteUN se hizo un análisis exploratorio de las dos bases de datos que indicó la necesidad o no de depurar, las cuales se sometieron a pruebas de la estadística descriptiva en los que se determinó la media, la desviación estándar, los valores máximo y mínimo, el coeficiente de variación, la simetría, el estudio de la normalidad y la estacionariedad, así como las pruebas esenciales para determinar la idoneidad de la información, previo a la aplicación del método de interpolación Kriging ordinario.

La estadística descriptiva y prueba de normalidad en el sistema GeoStatLiteUN se hicieron con los métodos: *describe()*, *skew()* y *shapiro()* (con significancia de 5%) de la librería SciPy; la regresión lineal multivariada para la prueba de estacionariedad de sentido amplio se programó utilizando el método *From\_formula()* de la librería Statsmodels y el método de interpolación Kriging ordinario se implementó con el método *OrdinaryKriging()* de la librería PyKrige.

La depuración de las bases de datos (cuando su distribución no fue simétrica) consistió en la eliminación de valores atípicos considerados leves según la prueba de Tukey; para esto GeoStatLiteUN identificó el cuantil uno (Q1) y el cuantil tres (Q3), el rango intercuantil (IQR) y eliminó todos los datos que estuvieran a una y media (1.5) veces el IQR por debajo del Q1 y por encima del Q3, estableciendo así nuevos mínimos y máximos para el atributo de interés de las bases de datos analizadas; este proceso se programó utilizando los métodos *quantile()*, *min()* y *max()* de Python.

### Generación de superficies continuas

GeoStatLiteUN es un programa que genera superficies continuas univariadas (una sola variable dependiente con covariables espaciales) mediante el método Kriging ordinario bajo el supuesto de isotropía, el cual depende de un modelo de semivariograma ajustado. El programa es capaz de ajustar cuatro modelos de semivariogramas teóricos: Linear, Exponential, Spherical y Gaussian; se usa para estimar las predicciones como una ponderación que asigna mayor peso a los valores que están cerca entre sí y menor peso a las que están más distantes entre sí; es decir, los parámetros del semivariograma determinan los pesos de los datos que serán usados para la estimación de las predicciones.

Para la selección del modelo de mejor ajuste a la variabilidad espacial de los datos analizados, GeoStatLiteUN proporciona información sobre tres geoestadísticos conocidos como:  $Q_1$ ,  $Q_2$  y  $cR$ ; donde  $Q_1$  representa el error medio del modelo,  $Q_2$  representa la varianza del error y  $cR$  representa la media geométrica de los residuos al cuadrado (que tiene que ver con la variabilidad espacial que no es capaz de explicar el modelo). Un buen modelo de Kriging ordinario debe arrojar un  $Q_1$  cercano a cero, un  $Q_2$  cercano a uno y el menor valor posible de  $cR$  (Stephan *et al.*, 2016).

### *Evaluación de GeoStatLiteUN*

El objetivo del software System for Automated Geoscientific Analyses (SAGA GIS) es brindar una plataforma eficaz y fácil de aprender para la implementación de métodos geocientíficos y que sean accesibles mediante su interfaz gráfica de usuario. En este estudio se compararon los resultados de GeoStatLiteUN y SAGA GIS a través de los modelos de superficies continuas, creados con los datos de prueba en ambos programas (analizados en GeoStatLiteUN, ya que SAGA GIS no ofrece esa facilidad), a partir de los semivariogramas construidos con un número de iteraciones igual a treinta (intervalos de clase).

La comparación se hizo a través de las visualizaciones de los resultados producidos, las características de los semivariogramas (Nugget, Rango y Full Sill), y los estadísticos de los modelos espaciales (mínimo, máximo, media y desviación estándar). Adicionalmente, como un parámetro de evaluación, se construyeron intervalos de confianza con 99% de confiabilidad y 1% de significancia para las medias de los modelos construidos con GeoStatLiteUN y SAGA GIS (Diez *et al.*, 2019); y se consideró razonable decir que los modelos construidos con GeoStatLiteUN son aceptables si su intervalo contuvo la media de la muestra.

## Resultados

### Aplicación

Se desarrolló el programa *GeoStatLiteUN 2021.6.1* con IGU (figura 1), una aplicación que implementa métodos estadísticos y geoestadísticos utilizando módulos especializados de Python y permite al usuario, de forma rápida, procesar y analizar bases de datos con un atributo y sus covariables espaciales, así como de obtener resultados de forma automática. GeoStatLiteUN se encuentra disponible para su descarga libre en [https://drive.google.com/file/d/1wIhKBHm4CdL0cXzDtbTvfaxRkPGwb6\\_H/view?usp=drivesdk](https://drive.google.com/file/d/1wIhKBHm4CdL0cXzDtbTvfaxRkPGwb6_H/view?usp=drivesdk).

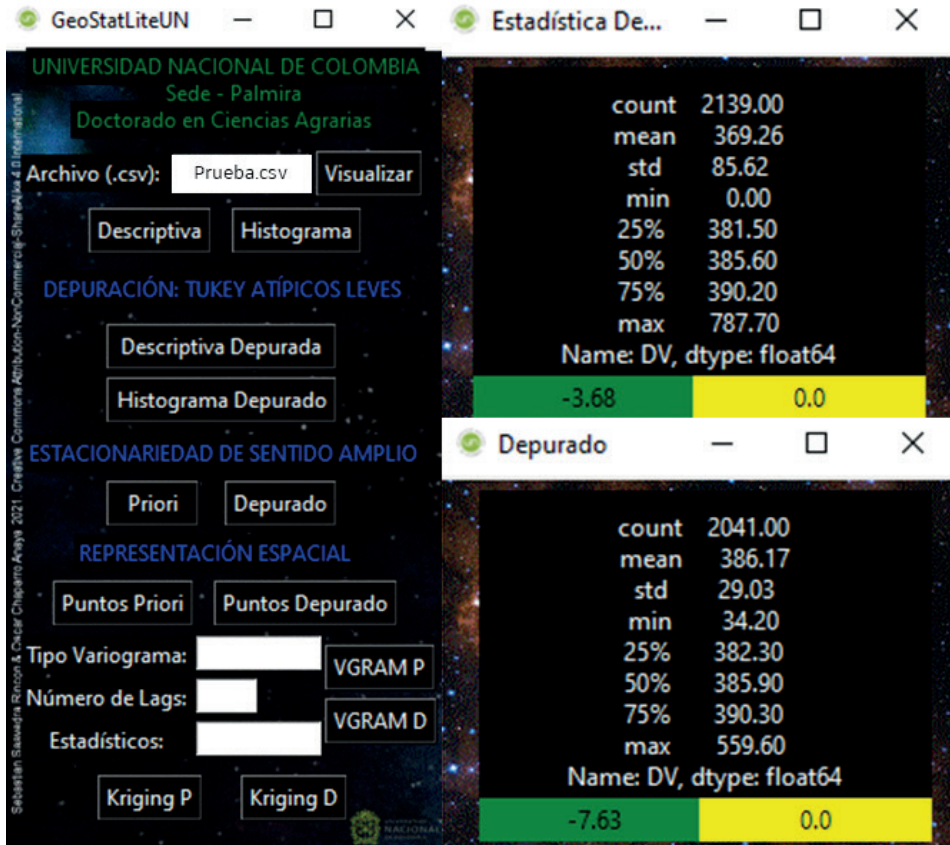
### Análisis estadístico

#### *Procesamiento de datos: Operación de fertilización mecánica a tasa variada*

En el cuadro 1 se muestran los estadísticos de la base de datos depurada de forma automática en GeoStatLiteUN, correspondiente a la operación de fertilización mecánica a tasa variada. Para su procesamiento y análisis, la base de datos contó con un total de 1 183

datos del atributo de interés con sus respectivas covariables espaciales, cuyos valores fueron: mínimo de 124.80 kg<sub>nitrogeno</sub>/ha, máximo de 282.90 kg<sub>nitrogeno</sub>/ha, media de 214.21 kg<sub>nitrogeno</sub>/ha y desviación estándar de 20.12 kg<sub>nitrogeno</sub>/ha.

Figura 1  
GeoStatLiteUN v2021.6.1 IGU. Ejemplo



Cuadro 1

Estadística descriptiva. Variable dosis en la operación de fertilización mecánica

Dosis (kg <sub>nitrogeno</sub> /ha)			
Cuenta	1 183.00	Cuantil 1	202.35
Media	214.21	Cuantil 2	213.60
desviación estándar	20.12	Cuantil 3	226.30
Mínimo	124.80	Máximo	282.90

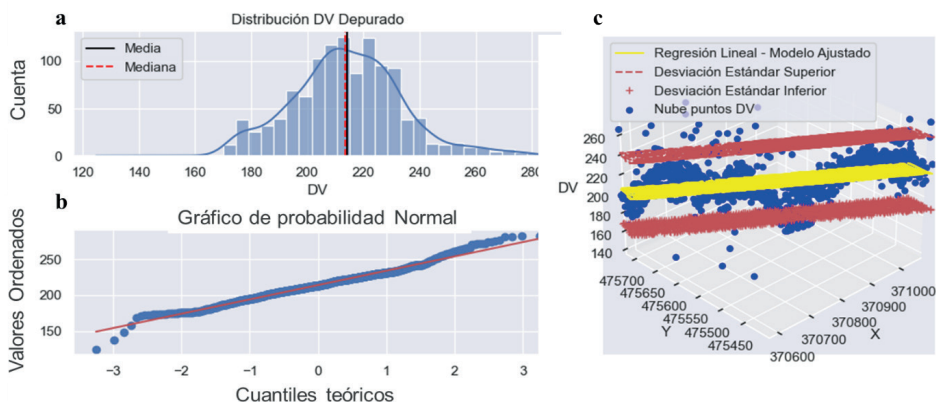


El histograma y la gráfica Q-Q para la variable (figura 2a y figura 2b) mostró que la distribución no sigue estrictamente la forma de una distribución normal y, debido al número de datos que se analizó, fue posible aplicar el teorema del límite central y decir que la media de la distribución de la muestra tiende a ser normal y simétrica (Diez *et al.*, 2019).

Por otro lado, el modelo de regresión lineal multivariado (figura 2c) mostró que los datos del atributo de interés no siguieron una tendencia lineal frente a sus covariables espaciales y se agruparon alrededor de la media del modelo de regresión lineal, de forma uniforme con distancia casi igual a una desviación estándar de la muestra. Lo que indica que el fenómeno fue un proceso estocástico estacionario de sentido amplio (Jaramillo-Jaramillo, 2012). Como consecuencia, la base de datos depurada, correspondiente a la operación de fertilización mecánica a tasa variada, fue apta para la creación de un semivariograma y un modelo de superficie continua con el método Kriging ordinario.

Figura 2

Histograma (a), Q-Q plot (b) y regresión lineal multivariada (c).  
Variable dosis en la operación de fertilización mecánica



### Procesamiento de datos: Operación de roturación vertical con cinceles rígidos

Los datos registrados durante la operación de roturación vertical con cinceles rígidos no requirieron ser sometidos al protocolo automático de depuración ofrecido por la herramienta GeoStatLiteUN, debido a que los valores de la media y el cuantil 2 son iguales, con un valor de 50 cm (distribución simétrica). Para el análisis exploratorio, se trabajó con 4 105 datos del atributo de interés con sus respectivas covariables espaciales para un valor mínimo de 45 cm, un valor máximo de 55 cm, una media de 50 cm y una desviación estándar de 2.65 cm (cuadro 2).



Cuadro 2

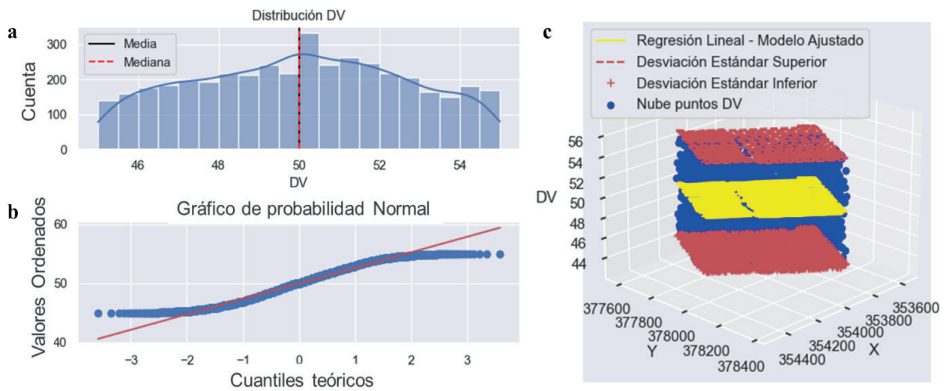
Estadística descriptiva. Variable profundidad en la operación de roturación vertical

Profundidad (cm)			
Cuenta	4 105.00	Cuantil 1	47.9
Media	50.02	Cuantil 2	50
desviación estándar	2.65	Cuantil 3	52.10
Mínimo	45.00	Máximo	55.00

El histograma y la gráfica Q-Q para la variable (figura 3a y figura 3b) mostraron que la distribución no sigue estrictamente la forma de una distribución normal y, debido al número de datos que se analizó, fue posible aplicar el teorema del límite central, además se encontró que la media de la distribución de la muestra tiende a ser normal y simétrica (Diez *et al.*, 2019).

Figura 3

Histograma (a), Q-Q plot (b) y regresión lineal multivariada (c).  
Variable profundidad en la operación de roturación vertical



Por otro lado, el modelo de regresión lineal multivariado (figura 3c) muestra que los datos del atributo de interés no siguieron una tendencia lineal frente a sus covariables espaciales y se agruparon alrededor de la media del modelo de regresión lineal de forma uniforme con distancia igual a una desviación estándar de la muestra, lo que indica que el fenómeno es un proceso estocástico estacionario de sentido amplio (Jaramillo-Jaramillo, 2012). Como consecuencia, la base de datos, correspondiente a la operación de roturación vertical con cinceles rígidos, fue apta para la creación de un semivariograma y un modelo de superficie continua con el método Kriging ordinario.

*Generación de superficie continua: Fertilización mecánica a tasa variada*

En GeoStatLiteUN se crearon cuatro semivariogramas empíricos y teóricos del atributo de interés y sus covariables espaciales. El modelo que mejor se ajustó a la base de datos depurada fue el esférico, con el menor valor de cR igual a 202.93, los geoestadísticos para este modelo fueron: 0.11 y 0.56 respectivamente para  $Q_1$  y  $Q_2$  (cuadro 3).

Cuadro 3

## Geoestadísticos. Variable dosis en la operación de fertilización mecánica

Geoestadísticos	Modelo Semivariograma			
	Lineal	Exponencial	Esférico	Gaussiano
$Q_1$	0.13	0.09	0.11	0.21
$Q_2$	0.58	3.18	0.56	0.76
cR	212.41	188.28	202.93	281.60

Las características estructurales del semivariograma de mejor ajuste a la base de datos depurada (figura 4a) se pueden ver en el cuadro 4, al igual que los parámetros estadísticos del modelo de superficie continua que representó espacialmente el fenómeno estudiado. Es importante resaltar que el modelo de superficie continua disminuyó la variabilidad del atributo de interés y mostró un valor mínimo mayor y uno máximo menor al que se encontró en la estadística descriptiva del cuadro 1.

Con la información estructural de la base de datos, se construyó el modelo Kriging ordinario de superficie continua que modeló espacialmente el comportamiento de la operación de fertilización mecánica en términos de dosis de nitrógeno aplicada (figura 4b).

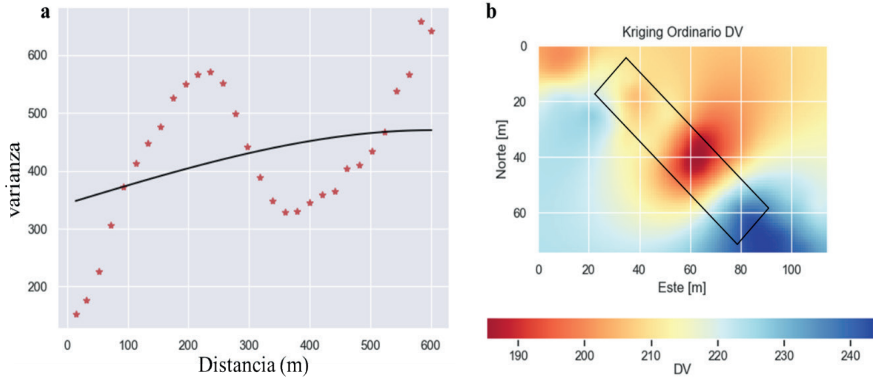
Cuadro 4

## Características estructurales del semivariograma ajustados y estadístico del modelo de superficie continua Kriging ordinario. Variable dosis en la operación de fertilización mecánica

Semivariograma ajustado (características)	
Nugget ( $\text{kg}_{\text{nitrógeno}}/\text{ha}$ ) <sup>2</sup>	343.50
Rango (m)	601.10
Full Sill ( $\text{kg}_{\text{nitrógeno}}/\text{ha}$ ) <sup>2</sup>	470.40
Modelo Kriging ordinario ( $\text{kg}_{\text{nitrógeno}}/\text{ha}$ )	
Mínimo	185.21
Máximo	244.16
Media	214.77
Desviación estándar	10.93

Figura 4

Semivariograma (a) y superficie continua Kriging (b).  
Variable dosis en la operación de fertilización mecánica



*Generación de superficie continua: Roturación vertical con cincelos rígidos*

En GeoStatLiteUN se crearon cuatro semivariogramas empíricos y teóricos del atributo de interés y sus covariables espaciales. El modelo que mejor se ajustó a la base de datos fue el exponencial, con el menor valor de cR de 6.49 y los geoestadísticos para este modelo fueron 0.02 y 1.03 respectivamente para  $Q_1$  y  $Q_2$  (cuadro 5).

Cuadro 5

Geoestadísticos. Variable profundidad en la operación de roturación vertical

Geoestadísticos	Modelo Semivariograma			
	Lineal	Exponencial	Esférico	Gaussiano
$Q_1$	0.02	0.02	0.02	0.02
$Q_2$	1.02	1.03	1.03	1.03
cR	7.07	6.49	6.58	6.60

Las características estructurales del semivariograma que mejor se ajustó a la base de datos depurada (figura 5a) se presentan en el cuadro 6, al igual que los parámetros estadísticos del modelo de superficie continua que representó espacialmente el fenómeno estudiado. Es importante resaltar que el modelo de superficie continua disminuyó la variabilidad del atributo de interés con un valor mínimo mayor y uno máximo menor al que se encontró en la estadística descriptiva del cuadro 2.

Con la información estructural de la base de datos se construyó el modelo Kriging ordinario de superficie continua que modeló espacialmente el comportamiento de la operación de fertilización mecánica en términos de profundidad alcanzada en el suelo (figura 5b).

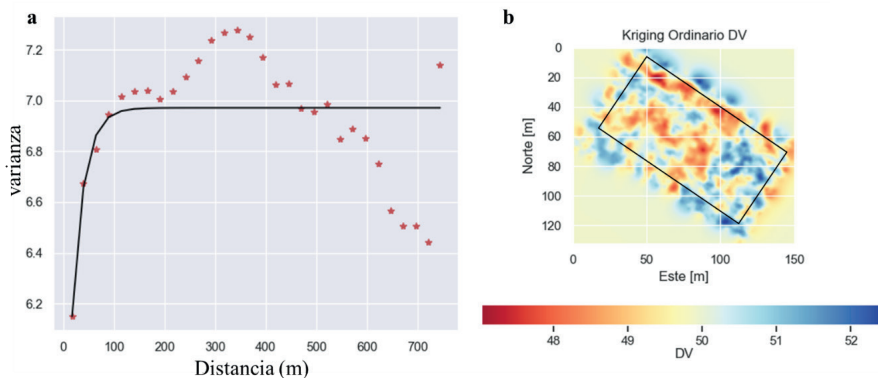
### Cuadro 6

Características estructurales del semivariograma ajustados y estadístico del modelo de superficie continua Kriging ordinario. Profundidad en la operación de roturación vertical

Semivariograma ajustado (características)	
Nugget (cm <sup>2</sup> )	5.27
Rango (m)	69.97
Full Sill (cm <sup>2</sup> )	6.97
Modelo Kriging ordinario (cm)	
mínimo	47.03
máximo	52.48
media	49.93
desviación estándar	0.68

Figura 5

Semivariograma (a) y superficie continua Kriging (b). Variable profundidad en la operación de roturación vertical



## Generación de superficies continuas en SAGA GIS – Operaciones: Fertilización mecánica a tasa variada y roturación vertical con cinceles rígidos

### *Fertilización mecánica a tasa variada*

Las características estructurales del semivariograma que mejor se ajustó a la base de datos depurada (figura 6a) se muestran en el cuadro 7, al igual que los parámetros estadísticos del modelo de superficie continua que representó espacialmente el fenómeno estudiado. Es importante resaltar que el modelo de superficie continua creado con SAGA GIS, al igual que en GeoStatLiteUN, disminuye la variabilidad del atributo de interés y mostró

un valor mínimo mayor y un valor máximo menor al que se determinó en la estadística descriptiva del cuadro 1; sin embargo, la desviación estándar del modelo SAGA GIS de  $18.35 \text{ kg}_{\text{nitrogeno}}/\text{ha}$  fue más cercano a la desviación de  $20.12 \text{ kg}_{\text{nitrogeno}}/\text{ha}$  que se presentó en el cuadro 1.

### Cuadro 7

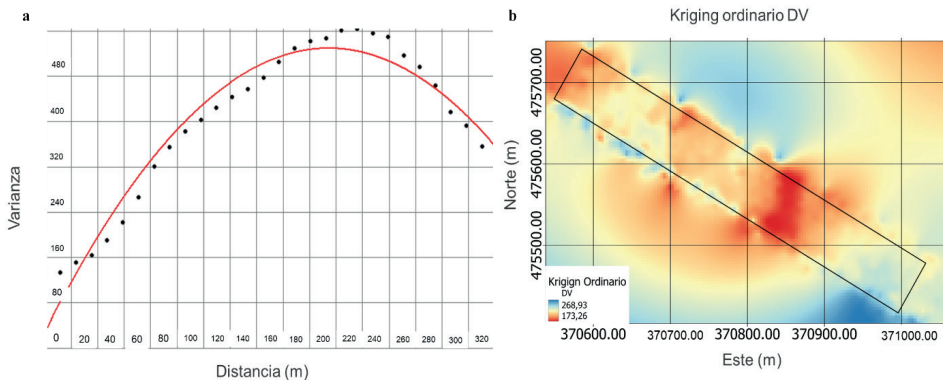
Características estructurales del semivariograma esférico ajustado por defecto en SAGA GIS y estadísticos del modelo de superficie continua Kriging ordinario. Variable dosis en la operación de fertilización mecánica

Semivariograma ajustado (características)	
Nugget ( $\text{kg}_{\text{nitrogeno}}/\text{ha}$ ) <sup>2</sup>	28.86
Rango (m)	338.50
Full Sill ( $\text{kg}_{\text{nitrogeno}}/\text{ha}$ ) <sup>2</sup>	360.00
Modelo Kriging ordinario	
(kg <sub>nitrogeno</sub> /ha)	
Mínimo	173.26
Máximo	268.93
Media	217.91
Desviación estándar	18.35

Con la información estructural de la base de datos se construyó el modelo Kriging ordinario de superficie continua en SAGA GIS que modeló espacialmente el comportamiento de la operación de fertilización mecánica en términos de dosis de nitrógeno aplicada (figura 6b).

Figura 6

Semivariograma (a) y superficie continua Kriging (b) SAGA GIS. Variable dosis en la operación de fertilización mecánica



### Roturación vertical con cinceles rígidos

Las características estructurales del semivariograma que mejor se ajustó a la base de datos depurada (figura 7a) se pueden ver en el cuadro 8, al igual que los parámetros estadísticos del modelo de superficie continua que representó espacialmente el fenómeno estudiado. Es importante resaltar que el modelo de superficie continua creado con SAGA GIS, al igual que en GeoStatLiteUN, disminuyó la variabilidad del atributo de interés y mostró un valor mínimo mayor y uno máximo menor al que se determinó en la estadística descriptiva del cuadro 2; sin embargo, la desviación estándar del modelo SAGA GIS de 1.02 cm fue más cercano a la desviación de 2.65 cm presentado en el cuadro 2.

Con la información estructural de la base de datos se construyó el modelo Kriging ordinario de superficie continua en SAGA GIS que modeló espacialmente el comportamiento de la operación de roturación vertical en términos de profundidad alcanzada (figura 7b).

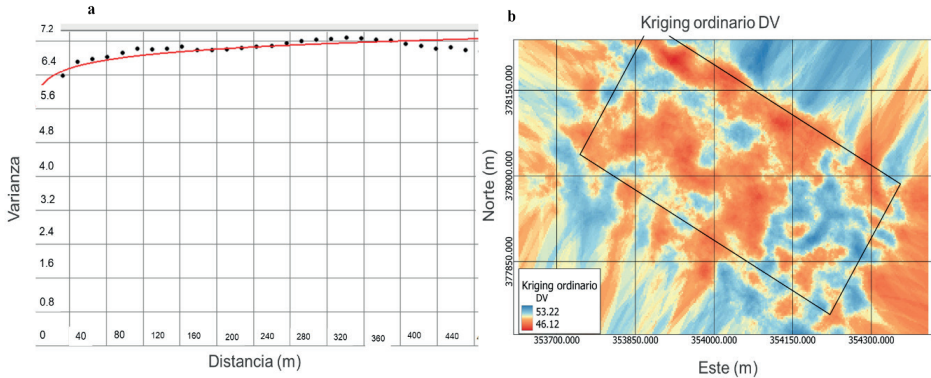
#### Cuadro 8

Características estructurales del semivariograma exponencial ajustado por defecto en SAGA GIS y estadísticos del modelo de superficie continua Kriging ordinario. Profundidad en la operación de roturación vertical

Semivariograma ajustado (características)	
Nugget (cm <sup>2</sup> )	5.60
Rango (m)	488.26
Full Sill (cm <sup>2</sup> )	7.10
Modelo Kriging ordinario (cm)	
Mínimo	46.12
Máximo	53.22
Media	49.96
Desviación estándar	1.02

Figura 7

Semivariograma (a) y superficie continua Kriging (b) SAGA GIS.  
Variable profundidad en la operación de roturación vertical



## Evaluación de GeoStatLiteUN

### *Modelos espaciales: Operación de fertilización mecánica a tasa variada*

En términos estructurales, los valores más bajos para las características del semivariograma se obtuvieron en el programa SAGA GIS (cuadro 9). Esto se explica porque los modelos utilizados en ambos programas son similares más no iguales, lo que resulta en un comportamiento diferente de las características de los modelos ajustados.

El modelo GeoStatLiteUN fue conservador (menos variable con relación a la media de la muestra) frente al modelo SAGA GIS (más variable con relación a la media de la muestra), lo anterior se puede afirmar con base en el valor del rango (su magnitud expresa la variabilidad espacial del atributo: en cuanto más grande sea, la variabilidad espacial es menor) para GeoStatLiteUN, que fue mayor frente al valor de SAGA GIS (Villatoro *et al.*, 2007; Acevedo *et al.*, 2008; Saavedra-Rincon, 2018).

Lo anterior influyó en los valores mínimos y máximos para los modelos de superficie continua, donde fueron, para SAGA GIS, más cercanos a sus pares mostrados en el cuadro 1, que los valores para GeoStatLiteUN, lo que se reflejó en una desviación estándar para SAGA GIS ( $18.35 \text{ kg}_{\text{nitrogeno}}/\text{ha}$ ) similar a la desviación estándar exploratoria del cuadro 1 ( $20.12 \text{ kg}_{\text{nitrogeno}}/\text{ha}$ ) en comparación con la desviación estándar para GeoStatLiteUN ( $10.93 \text{ kg}_{\text{nitrogeno}}/\text{ha}$ ).



### Cuadro 9

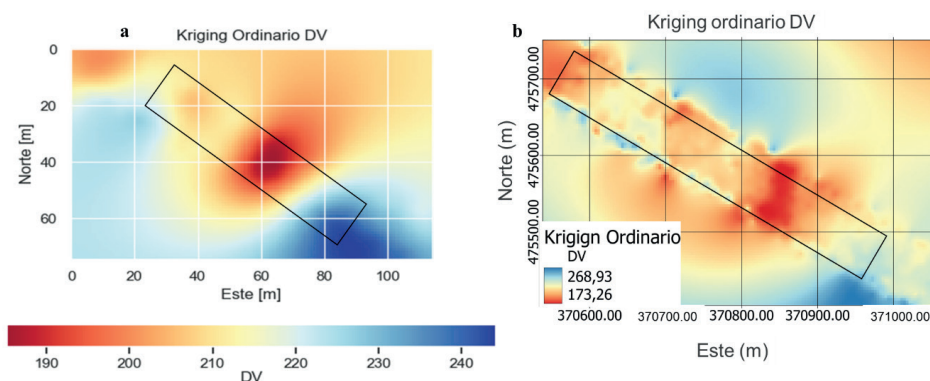
#### Características estructurales de los semivariogramas esféricos y estadísticos de los modelos de superficie continua Kriging ordinario. GeoStatLiteUN y SAGA GIS

Semivariograma ajustado	GeoStatLiteUN	SAGA GIS
Nugget ( $\text{kg}_{\text{nitrogeno}}/\text{ha})^2$	343.50	28.86
Rango (m)	601.10	338.50
Full Sill ( $\text{kg}_{\text{nitrogeno}}/\text{ha})^2$	470.40	360.00
Modelo Kriging ordinario ( $\text{kg}_{\text{nitrogeno}}/\text{ha}$ )	GeoStatLiteUN	SAGA GIS
mínimo	185.21	173.26
máximo	244.16	268.93
media	214.77	217.91
desviación estándar	10.93	18.35

La comparación visual de los modelos GeoStatLiteUN (figura 8a) y SAGA GIS (figura 8b) evidenció concordancia en su distribución; es decir que, aunque el rango de valores que el atributo tomó, con relación a las covariables espaciales en ambos modelos, fue de diferentes mínimos, medias y máximos; el comportamiento espacial del fenómeno analizado fue similar (superficies continuas con modelos de colores distribuidos de forma similar).

Figura 8

Modelos Kriging ordinario: GeoStatLiteUN (a) y SAGA GIS (b). Variable dosis en la operación de fertilización mecánica



Se puede afirmar, según los intervalos de confianza para los modelos GeoStatLiteUN y SAGA GIS (cuadro 10), que la media de la muestra estará en 99% de los intervalos de confianza construidos con los estadísticos de los modelos

desarrollados con base en la muestra (Diez *et al.*, 2019), y que fue razonable decir que el modelo construido con GeoStatLiteUN fue aceptable (el intervalo de confianza contiene la media de la muestra), aunque es un modelo conservador (menos dependiente de los valores extremos) frente al construido con SAGA GIS.

Cuadro 10

Intervalos de confianza para los modelos GeoStatLiteUN y SAGA GIS

Límites - Intervalos de Confianza	GeoStatLiteUN	SAGA GIS
Límite inferior (kg <sub>nitrogeno</sub> /ha)	213.95	216.86
Límite superior (kg <sub>nitrogeno</sub> /ha)	215.59	219.28

*Modelos espaciales: Operación de roturación con cinceles rígidos*

En términos estructurales, los valores más bajos para las características del semivariograma se obtuvieron en el programa GeoStatLiteUN (cuadro 11), esto se explica porque los modelos utilizados en ambos programas son similares más no iguales, lo que resulta en un comportamiento diferente de las características de los modelos ajustados.

Cuadro 11

Características estructurales de los semivariogramas esféricos y estadísticos de los modelos de superficie continua Kriging ordinario. GeoStatLiteUN y SAGA GIS

Semivariograma ajustado	GeoStatLiteUN	SAGA GIS
Nugget (cm) <sup>2</sup>	5.27	5.6
Rango (m)	69.97	488.26
Full Sill (cm) <sup>2</sup>	6.97	7.10
Modelo Kriging ordinario (cm)	GeoStatLiteUN	SAGA GIS
Mínimo	47.03	46.12
Máximo	52.48	53.22
Media	49.93	49.96
Desviación estándar	0.68	1.02

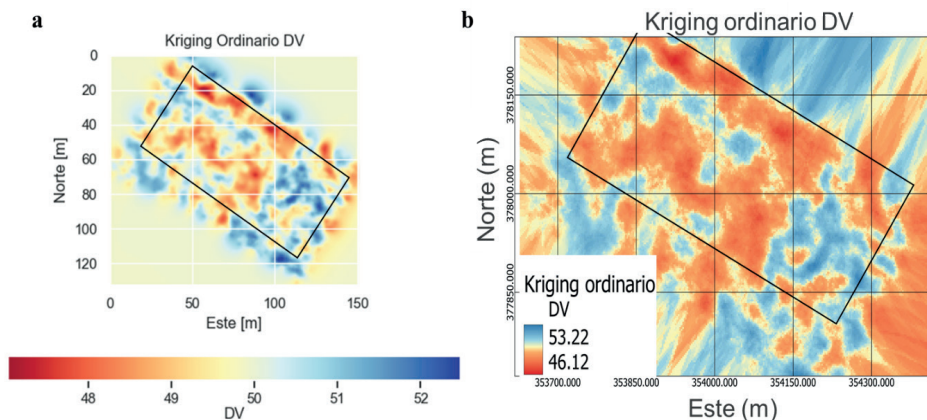
La comparación visual de los modelos GeoStatLiteUN (figura 9a) y SAGA GIS (figura 9b) evidenció concordancia en su distribución, es decir que el comportamiento espacial del fenómeno analizado fue similar.

El modelo GeoStatLiteUN fue más variable frente al modelo SAGA GIS, debido a que el valor del rango para GeoStatLiteUN fue menor frente al valor de SAGA GIS (Villatoro *et al.*, 2007; Saavedra-Rincon, 2018). Sin embargo, lo anterior no influyó en los valores mínimos y máximos para los modelos de superficie continua en donde, para

SAGA GIS (46.12 cm y 53.22 cm, respectivamente) y GeoStatLiteUN (47.03 cm y 52.48 cm, respectivamente) sus valores fueron similares y cercanos a sus pares mostrados en el cuadro 2 (45 cm y 55 cm, respectivamente), lo que se reflejó en una desviación estándar para SAGA GIS (1.02 cm) similar a la desviación estándar para GeoStatLiteUN (0.68 cm), esto se explica por la poca variabilidad del atributo que se evidenció en el análisis exploratorio del cuadro 2, lo que supeditó los modelos espaciales que en sí muestran ser conservadores frente a dicha variabilidad.

Figura 9

Modelos Kriging ordinario: GeoStatLiteUN (a) y SAGA GIS (b).  
Variable profundidad en la operación de roturación vertical



Según los intervalos de confianza para los modelos GeoStatLiteUN y SAGA GIS (cuadro 12), la media de la muestra estará en 99% de los intervalos de confianza construidos con los estadísticos de los modelos espaciales desarrollados con base en la muestra (Diez *et al.*, 2019), con lo anterior es razonable decir que el modelo construido con GeoStatLiteUN fue aceptable porque el intervalo de confianza contiene la media de la muestra y fue, además, muy similar al construido con SAGA GIS.

Cuadro 12

Intervalos de confianza para los modelos GeoStatLiteUN y SAGA GIS

Límites - Intervalos de confianza	GeoStatLiteUN	SAGA GIS
Límite inferior (cm)	49.90	49.92
Límite superior (cm)	49.96	50.00

## Discusión

El uso del sistema GeoStatLiteUN permitió entender cómo se comportaron espacialmente las operaciones de fertilización mecánica a tasa variada y roturación vertical con cinceles rígidos en dos suelos gestionados para el cultivo de caña de azúcar (*Saccharum officinarum*), cumpliendo con los requerimientos fundamentales para la creación de un mapa de variabilidad espacial a través del método Kriging ordinario, tal como lo desarrolló Armstrong (1998) en su libro *Basic Linear Geostatistics*; cumpliéndose los requerimientos de conocer el volumen de datos recolectado, la distribución de la muestra, sus estadísticos descriptivos (media, desviación estándar, etcétera), conocer si el fenómeno es estacionario, calcular varios semivariogramas con sus geoestadísticos y seleccionar el de mejor ajuste.

El análisis de las bases de datos correspondientes mostró que las medias de la distribución de los atributos de interés en las muestras fueron normales y simétricas, y que el comportamiento de los atributos y sus covariables espaciales fueron estacionarios de sentido amplio.

La comparación de los mapas de variabilidad espacial del sistema GeoStatLiteUN con SAGA GIS para las dos operaciones de mecanización agrícola estudiadas reflejó similitud en sus resultados; tanto en la distribución del atributo de interés como en los valores predichos, con base en ponderaciones de la media de la muestra, aunque los modelos de GeoStatLiteUN fueron más conservadores; lo que permitió afirmar con 99% de confianza, que los modelos hechos con GeoStatLiteUN basarán sus predicciones en intervalos que contendrán la media de la muestra con límites más robustos (cercaos al Q1 y Q3).

## Conclusiones

La contribución innovadora del sistema GeoStatLiteUN es la integración de protocolos automáticos para la depuración basada en la prueba de Tukey para atípicos leves, la validación de la estacionariedad de sentido amplio, la construcción y selección de superficies continuas (mapas de variabilidad espacial); además, que se comparte como un archivo portable que posterior a su descarga no requiere conexión a internet, lo que facilitará el uso del sistema y el método geoestadístico Kriging ordinario a los profesionales del agro en el campo, sin requerir conocimientos previos en los algoritmos matemáticos a usar.

El lenguaje de programación Python y sus módulos especializados permitieron la creación de un sistema de información geográfico de libre acceso enfocado en la construcción de mapas de variabilidad espacial con una interfaz gráfica de usuario (GeoStatLiteUN), que permitirá a los usuarios procesar y analizar datos espaciales en el ámbito de la agricultura digital.

El sistema GeoStatLiteUN logró resultados válidos al observar su similitud con el programa de amplia difusión SAGA GIS; sin embargo, es necesario validar en campo los mapas de variabilidad espacial para garantizar su eficiencia y apoyarse en un programa más robusto para la confección de planos convencionales.

## Literatura citada

- Acevedo, C.; Álvarez-Sánchez, M.E.; Hernández-Acosta, E.; Maldonado-Torres, R.; Pérez-Grajales, M. y Castro-Brindis, R. (2008). Variabilidad espacial de propiedades químicas del suelo y su uso en el diseño de experimentos. *Terra Latinoamericana*, 26(4): 317-324. [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0187-57792008000400004&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-57792008000400004&lng=es&tlng=es) (Consultado 19 agosto 2021).
- Armstrong, M. (1998). *Basic Linear Geostatistics*. Estados Unidos: Editorial Springer-Verlag Berlin Heidelberg. 154 pp. <https://www.springer.com/gp/book/9783540618454> (Consultada 24 junio 2021).
- Basso, L.H.; Bernardi, A.C.; Cruvinel, P.E.; Inamasu, R.Y.; Speranza, E.A. y Vaz, C.M. (2019). Agricultura de precisión e agricultura digital. *TECCOGS*, 0(20): 17-36. <https://doi.org/10.23925/1984-3585.2019i20p17-36> (Consultada 15 junio 2021).
- Córdoba, M.; Bruno, C.; Balzarini, M. y Costa, J. (2012). Análisis de componentes principales con datos georreferenciados. Una aplicación en agricultura de precisión. *Revista de la Facultad de Ciencias Agrarias*, 44(1): 27-39. <https://www.redalyc.org/articulo.oa?id=382837650003> (Consultado 19 agosto 2021).
- Cravero, A. y Sepúlveda, S. (2021). Use and Adaptations of Machine Learning in Big Data-Applications in Real Cases in Agriculture. *Electronics*, 10(5): 552. <https://doi.org/10.3390/electronics10050552> (Consultada 19 agosto 2021).
- Diez, D.; Barr, C. y Cetinkaya-Rundel, M. (2019). One-Sample Means with the T-Distribution. In: D. Diez, M. Cetinkaya-Rundel, y C. Barr (Edits.), *OpenIntro Statistics Fourth Edition* (pp. 251-261). Estados Unidos: OpenIntro Ed. <https://www.openintro.org/os> (Consultado 15 junio 2021).
- Fernández, D. y Ribes-Dasi, M. (2014). Uso de la geoestadística y los sistemas de información geográfica en agricultura. En: E. Mantovani y Magdalena Carlos (Edits.), *Manual de agricultura de precisión* (pp. 84-91). Uruguay: PROCISUR Ed. [http://www.procisur.org.uy/adjuntos/Manual\\_AP\\_01web.pdf](http://www.procisur.org.uy/adjuntos/Manual_AP_01web.pdf) (Consultada 14 junio 2021).
- Gutiérrez-López, A.; Fuentes, C.; Lebel, T.; Ramirez, A. y Santillán, O. (2012). El variograma y el correlograma, dos estimadores de la variabilidad de mediciones hidrológicas. *Rev. Fac. Ing. Univ. Antioquia*, 0(59): 193-202. <https://revistas.udea.edu.co/index.php/ingenieria/article/view/13824> (Consultada 18 junio 2021).
- Herrera-Gómez, A. y Reinoso-Gutiérrez, J. (2016). Diseño e implementación de una aplicación SIG para la supervisión y monitoreo de sensores en el campo de la agricultura de precisión. Tesis de grado, Universidad de los Llanos. Colombia. <https://repositorio.unillanos.edu.co/handle/001/1172> (Consultada 24 julio 2021).
- Jaramillo-Jaramillo, D. (2012). Variabilidad espacial del suelo: Bases para su estudio. *Revista de la Facultad de Ciencias*, 1(1): 73-87. <https://revistas.unal.edu.co/index.php/rfc/article/view/49004> (Consultado 19 junio 2021).
- Kamilaris, A.; Kartakoullis, A. y Prenafeta-Boldú, F. (2017). A Review on the Practice of Big Data Analysis in Agriculture. *Comput. Electron. Agric.*, 143: 23-37. <https://doi.org/10.1016/j.compag.2017.09.037> (Consultada 19 agosto 2021).
- Kim, S. Hee; Kang, M.S. y Jung, Y.G. (2016). Big Data Analysis Using Python in Agriculture Forestry and Fisheries. *International Journal of Advanced Smart Convergence*, 5(1): 47-50. <https://doi.org/10.7236/IJASC.2016.5.1.47> (Consultada 19 agosto 2021).
- Lago-González, C.; Sepúlveda-Peña, J.; Barroso-Abreu, R.; Fernández-Peña, F.; Maciá-Pérez, F. y Lorenzo, J. (2011). Sistema para la generación automática de mapas de rendimiento. Aplicación en la agricultura de precisión. *Idesia (Arica)*, 29(1): 59-69. <https://dx.doi.org/10.4067/S0718-34292011000100009> (Consultada 24 julio 2021).
- Loayza, V.; Sevilla, V.; Olivera, C.; Guevara, M.; Olmedo, G.; Vargas, R.; Oyonarte, C. y Jiménez, W. (2020). Mapeo digital de carbono orgánico en suelos de Ecuador. *Ecosistemas* 29(1): 1852. <https://doi.org/10.7818/ECOS.1852> (Consultada 19 agosto 2021).

- Paccioletti, P.; Balzarini, M. y Córdoba, M. (2020). Desarrollo de un software para mapeo de variabilidad espacial en agricultura y ambiente. *AgriScientia*, 37(1): 75-84. <https://doi.org/10.31047/1668.298x.v37.n1.27863> (Consultada 20 junio 2021).
- Saavedra-Rincón, S. (2018). Fertilización nitrogenada a tasa variada en seis campos de caña de azúcar: Hacia un manejo ambiental sostenible bajo el concepto de agricultura de precisión. Tesis de maestría, Universidad Nacional de Colombia. Palmira, Valle del Cauca, Colombia. <https://repositorio.unal.edu.co/handle/unal/69043> (Consultada 10 junio 2021).
- Salgado-Velázquez, S.; Salgado-García, S.; Rincón-Ramírez, J.A. et al. (2020). Spatial Variability of Soil Physicochemical Properties in Agricultural Fields Cultivated with Sugarcane (*Saccharum officinarum* L.) in Southeastern Mexico. *Sugar Tech*, 22:65-75. <https://doi.org/10.1007/s12355-019-00742-9> (Consultada 20 agosto 2021).
- Stephan, A.; Holzäpfel, F. y Schroll, J. (2016). Numerical Optimization of Plate-Line Design for Enhanced Wake-Vortex Decay. *Journal of Aircraft*, 54(3): 995-1010. <https://doi.org/10.2514/1.C033973> (Consultada 10 junio 2021).
- 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. *Agron. Costarricense*, 32(1): 95-105. <https://revistas.ucr.ac.cr/index.php/agrocost/article/view/6773> (Consultada 10 junio 2021).
- Yescas-Coronado, P.; Álvarez-Reyna, V.; García-Carrillo, M.; González-Cervantes, G.; Hernández-Hernández, V. y Segura-Castruita, M. (2018). Variabilidad espacial del carbono orgánico e inorgánico del suelo en la Comarca Lagunera, México. *Bol. Soc. Geol. Mex.*, 70(3): 591-610. <https://doi.org/10.18268/BSGM2018v70n3a2> (Consultada 14 junio 2021).

Recepción: 21 de junio 2021

Arbitraje: 13 de julio 2021

Dictamen: 19 de agosto 2021

Aceptado: 20 de agosto 2021