

COMPARACIÓN DE LOS INTERPOLADORES IDW Y KRIGING EN LA VARIACIÓN ESPACIAL DE pH, CA, CICE Y P DEL SUELO¹

Mario Villatoro^{2/*}, Carlos Henríquez^{**}, Freddy Sancho*

Palabras clave: Interpolador, Kriging, IDW, semivariograma, análisis espacial

Keywords: Interpolator, Kriging, IDW, semivariogram, spatial analysis

Recibido: 13/12/07

Aceptado:27/03/08

RESUMEN

Con el objetivo de comparar los interpoladores Kriging y el IDW (Inverse Distance Weighting), por ser los más utilizados en los estudios de análisis de la variación espacial, en un área de 2 467 m² se procedió a georeferenciar 61 puntos a una distancia de 3,5 m entre sí. Se tomó muestras de suelo a una profundidad de 0-15 cm en cada uno de los puntos. El programa GS+ para Windows se usó en los análisis de variogramas, interpolación y validación cruzada. Con valores de pH, Ca, CICE y P del suelo, se procedió a las interpolaciones. Con el Kriging se calculó los semivariogramas y también se determinó que el modelo esférico fue el de mejor ajuste. Como medidas de precisión se calculó el promedio absoluto del error (PAE) y el promedio del cuadrado del error (PCE); y, como medida de efectividad, el estimado de efectividad de predicción (E). Aunque ambos interpoladores tuvieron un desempeño similar, el Kriging fue superior al predecir de una mejor manera la variación de pH, Ca, y CICE, mientras que el IDW lo fue con el P.

ABSTRACT

Comparison of IDW and Kriging interpolators in the spatial variation of soil, pH, Ca, CICE, and P. The objective of this study was to compare the Kriging and IDW (Inverse Distance Weighting), as the most frequently-used interpolators in spatial analysis studies. From a 2467 m² area, 61 soil samples in a quadric design were taken at 0-15 cm depth. The average distance between sample points was 3.5 m. From the chemical analysis, only the variables pH, Ca, CICE, and P were used for the interpolations. Every single soil sample point was georeferenced. The GS+ for Windows program was used for variogram analysis, interpolation and cross validation. The spherical model was preferred on almost all the variograms. For comparison purposes, a cross validation was performed. As accuracy measurements, the Mean Absolute Error (MAE) and the Mean Square Error (MSE) were calculated. Also as a goodness-of-fit measurement, the prediction (E) was obtained, in order to compare both interpolators. Both procedures were satisfactory, however the Kriging method was better for pH, Ca, and CICE predictions, while IDW was better for P prediction.

1 Proyecto VI-510-A3-143 financiado por la Vicerrectoría de Investigación de la Universidad de Costa Rica.

2 Autor para correspondencia. Correo electrónico: mario.villatoro@ucr.ac.cr

* Centro de Investigaciones Agronómicas. Universidad de Costa Rica. San José, Costa Rica.

** Sede del Atlántico y Centro de Investigaciones Agronómicas. Universidad de Costa Rica.

INTRODUCCIÓN

La interpolación de datos ofrece la ventaja de proyectar mapas o superficies continuas a partir de datos discretos; sin embargo, la utilización de una buena cantidad de puntos del área en estudio limita su utilización (Johnston *et al.* 2001). Dependiendo del tipo de datos analizados, su costo y dificultad de obtención determinan que tan valioso es finalmente el uso de la interpolación. Otro aspecto a mencionar, es que la precisión en el mapa generado, a partir de las características de un suelo en particular, depende en gran medida de la estructura espacial de los datos, donde entre más fuerte la correlación espacial, mejor la calidad del mapeo (Kravchenko 2003).

Dentro de los interpoladores usados existe un grupo llamado Kriging, nombre dado por su creador, el ingeniero en minas surafricano D.G. Krige. Hasta el día de hoy, todos los interpoladores geoestadísticos están en el grupo de los Kriging (con sus variantes), los cuales ofrecen no solo predicciones y superficies de respuesta requeridas, sino también mapas de probabilidades y cuantiles (Johnston *et al.* 2001). El método Kriging cuantifica la estructura espacial de los datos -mediante el uso de variogramas llamados algunas veces semivariogramas debido a su similitud en el cálculo- y los predice mediante la interpolación, usando estadística. Se asume que los datos más cercanos a un punto conocido tienen mayor peso o influencia sobre la interpolación, influencia que va disminuyendo conforme se aleja del punto de interés. La medición de la probabilidad, efectuada por los métodos Kriging, hace la diferencia con respecto a los métodos determinísticos para interpolaciones espaciales, de los cuales los más usados son el de ponderación de distancias inversas (IDW: inverse distance weighting) y “splines” o ajuste por curvas (Burrough y McDonnell 1998).

El método IDW es similar al Kriging ordinario, ya que da más peso a los valores cercanos a un punto, pero posee una menor complejidad del cálculo. El IDW utiliza un algoritmo simple basado en distancias (Johnston *et al.* 2001). Diversos

autores han comparado el interpolador Kriging con el IDW en condiciones no tropicales (Gotway *et al.* 1996, Kravchenko y Bullock 1999, Mueller *et al.* 2001, Schloeder *et al.* 2001, Kravchenko 2003, Mueller *et al.* 2004). Ambos modelos, Kriging ordinario e IDW, asumen que las predicciones son una combinación lineal de los datos, como lo muestra la ecuación (1) (Gotway *et al.* 1996, Schloeder *et al.* 2001).

$$\hat{z}(s_o) = \sum_{i=1}^n \lambda_i z(s_i) \quad i=1, \dots, n \quad (1)$$

donde $\hat{z}(s_o)$ es el valor estimado en el punto interpolado s_o ; n es el número de observaciones vecinas usadas para la estimación y λ_i es el peso dado al valor observado $z(s_i)$ en las cercanías del valor s_o (Lozano *et al.* 2004). Este último parámetro hace la diferencia entre Kriging y el IDW.

El método ordinario de Kriging obtiene los pesos (o influencia) de los valores, resolviendo la ecuación Kriging mostrada en la ecuación (2) (Schloeder *et al.* 2001).

$$\sum_{i=1}^n \lambda_i \gamma[d(s_i, s_j)] + m = \gamma[d(s_o, s_i)], \quad i=1, \dots, n; \quad \sum_{i=1}^n \lambda_i = 1 \quad (2)$$

donde n es el número de observaciones, m es el multiplicador Lagrange usado para la minimización de las restricciones, λ es el peso dado a cada una de las observaciones y la suma de todos los λ es igual a uno. Los subíndices i y j denotan los puntos muestreados, el subíndice o es el punto en estimación, s simboliza la medición efectuada (variable medida) y $d(s, s_o)$ es la distancia entre s_i y s_o a partir del semivariograma:

$$\gamma[d(s_i, s_o)] = \text{var}[z(s_i) - z(s_o)] \quad (3)$$

Esta semivarianza calculada es una medida para determinar la similitud entre observaciones, en donde a mayor similitud, menor semivarianza (Lozano *et al.* 2004).

Los pesos (λ) o las relevancias de los valores, son determinados con el fin de asegurar que el error promedio para el modelo sea cero y además la varianza del error es minimizada (Schloeder *et al.* 2001), lo cual ofrece una predicción no sesgada. Pese a ello y como menciona Benmostefa (2006), este método requiere de supuestos estadísticos muy fuertes, como que la hipótesis intrínseca de estacionalidad sea aceptada, lo cual raramente se observa en la naturaleza.

Los parámetros utilizados en el análisis del semivariograma son mostrados en la figura 1, en donde **Co** es la variancia de discontinuidad espacial también llamada el efecto pepita o ruido espacialmente no correlacionado (Burrough y McDonnell 1998). **C** es la variancia estructural o espacialmente dependiente, en donde entre mayor participación tenga en la suma de **C+Co**, las estimaciones son mejores (Muñoz *et al.* 2006). El ámbito o rango es el valor de la correlación espacial o punto (en distancia), a partir del cual los datos no tienen influencia sobre el punto en comparación (Demmers 1999). **C+Co** llamado meseta o cima representa donde las varianzas de las diferencias son máximas y de obtenerse el variograma este sería inadecuado (Demmers 1999). **lag(h)** representa la distancia de los puntos circundantes a cada uno de los puntos en comparación. El lag puede ser definido con una distancia dada (a criterio de quién analiza) antes de calcular y graficar el variograma.

Kravchenko y Bullock (1999), encontraron que para la mayoría de análisis de datos, teniendo un número de puntos vecinos óptimo, una selección cuidadosa del modelo para el variograma y una transformación logarítmica es necesaria para normalizar los datos, en este caso el Kriging hace mejores estimaciones que el IDW.

Otro aspecto involucrado en el análisis del Kriging es la tendencia de isotropía o anisotropía, esta última indica si una variable tiene dependencia espacial hacia una o varias direcciones. Si la anisotropía fuese más fuerte, esta puede servir para determinar el área más homogénea según la variable medida, lo cual puede ser útil

al determinar parcelas experimentales (Lozano *et al.* 2004).

Por otro lado, el IDW calcula el peso de los valores de acuerdo a la relación inversa de la distancia (Schloeler *et al.* 2001) con a la ecuación (4).

$$\lambda_i = \frac{[d(s_i, s_o)]^{-p}}{\sum_{i=1}^n [d(s_i, s_o)]^{-p}} \quad i=1, \dots, n \quad (4)$$

Donde p es el parámetro del exponente que controla que tan rápido los pesos de los puntos tienden a cero (al aumentar su valor) conforme aumenta la distancia del sitio de interpolación. Entre mayor sea p , mayor peso es dado a los puntos más cercanos y por consiguiente se obtiene superficies más continuas o suaves y las predicciones tienden hacia el promedio de la muestra (Schloeder *et al.* 2001). Los valores p usualmente están entre 1 y 3, donde 2 es el más común (Gotway *et al.* 1996); de acuerdo a resultados de este mismo autor, la exactitud del IDW tiende a aumentar conforme p aumenta (1, 2, y 4) en el uso de datos con coeficientes de variación menores al 25%. Este coeficiente de variación, sin embargo, no pareció afectar la exactitud de las predicciones con el Kriging las cuales fueron mejores que las del IDW.

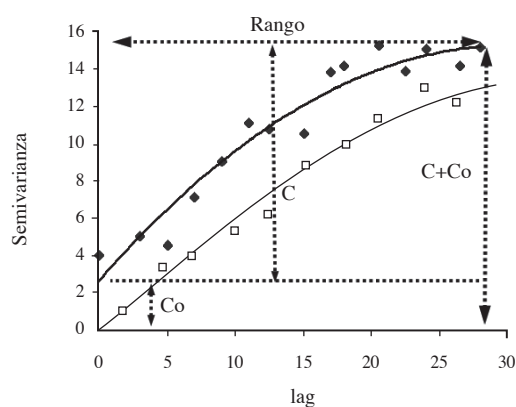


Fig. 1. Semivariograma típico y sus componentes.

El método IDW es más rápido en los cálculos; sin embargo, tiende a producir patrones poco reales conocidos como tipo “ojo de buey” alrededor de los puntos muestreados. Lo anterior expresa el peso que se le puede dar a la variación particular del valor de un punto de muestreo sobre los que están alrededor (Gotway 1996). Kravchenko y Bullock (1999) mencionan que a pesar de que el uso correcto del valor p y del número de valores vecinos más cercanos puede mejorar la interpolación, estos valores no pueden ser obtenidos con base en las propiedades estadísticas de los datos. Kravchenko (2003), recomienda el IDW para bases de datos pequeñas, en donde los parámetros del variograma no son conocidos, también cuando la distancia de muestreo es muy grande e incluso para cuando la distancia de muestreo es mayor al rango de la correlación espacial. Con relación a esto es importante mencionar que la precisión de las estimaciones es influenciada por la varianza aleatoria, la estructura de variación, y por la intensidad de muestreo (Lozano *et al.* 2004).

El objetivo de este trabajo fue determinar la eficiencia de los interpoladores Kriging e IDW, mediante su aplicación a las variables del suelo pH, Ca, CICE y P.

MATERIALES Y MÉTODOS

En un área de 2467 m², ubicada en la localidad de Atirro, Costa Rica, se eligió 61 puntos, distribuidos en un diseño de cuadrícula regular a una distancia promedio de 3,5 m y de los cuales se tomó muestras de suelo a una profundidad de 0-15 cm, cada uno de los puntos fue georeferenciado con la utilización de un GPS. El suelo pertenece al orden Inceptisoles. A las muestras se les determinó pH en agua (relación 1:2,5 suelo:solución), Ca, Mg, y acidez (1:10 KCl 1 N), así como K, P, Cu, Fe, Mn, y Zn (1:10 Olsen modificado). De estas variables solo se utilizó el pH, Ca, CICE y P para la llevar a cabo las interpolaciones, esto debido a que son las variables

típicamente observadas en análisis de fertilidad de suelos.

Con los datos de las variables seleccionadas, utilizando Stata 8.0 el estadístico de asimetría o también llamado grado de asimetría de la distribución de probabilidad, se puede tener un mejor criterio de si los grupos utilizados tienen una distribución normal y así aplicar el Kriging con más confianza. Entre mayor es el valor absoluto del estadístico de asimetría, se asume que los valores provienen de una distribución normal. También se calculó el estadístico de curtosis, el cual mide el grado en que las observaciones están agrupadas en los extremos de la curva de la distribución de datos, los valores cercanos a cero indican normalidad de los datos. Ambas pruebas fueron realizadas con los datos originales y transformados (logaritmo natural) de las variables elegidas. Con relación a la transformación de datos, Benmostefa (2006) menciona que esta normaliza los datos que no tienen este comportamiento, aunque esta no siempre es efectiva. De hecho, no hay garantía de que al hacer la transformación previo a aplicar el Kriging, los estimados vayan a ser los mejores estimados no sesgados.

Se procedió luego a realizar las interpolaciones con los procedimientos Kriging ordinario y el IDW mediante el uso del programa GS+ para windows. En el caso particular del interpolador Kriging, inicialmente se comparó los coeficientes de determinación (R^2) y el residuo de los errores de los modelos, con el fin de determinar cual de ellos se ajustaba mejor a la generación del semivariograma.

La comparación de las predicciones obtenidas a partir de los interpoladores utilizados, fue efectuada mediante el uso de 2 medidas de precisión usadas por Schloeder *et al.* (2001), que fueron el promedio absoluto del error (PAE) y el promedio del cuadrado del error (PCE).

El PAE es el promedio de la suma absoluta de los residuos (valor observado-valor estimado), que es definido también como el sesgo o error de la predicción y es calculado como:

$$PAE = \frac{1}{n} \sum_{i=1}^n \left| z(x_i) - \hat{z}(x_i) \right| \quad (5)$$

donde $z(x_i)$ es el valor observado en el punto “ i ”, $\hat{z}(x_i)$ es el valor estimado en el punto “ i ” y n es el número de puntos utilizados. Cuando el valor PAE es pequeño, se asocia a un método con pocos errores.

Esta medida aún así no refleja la magnitud de los errores que pueden ocurrir. Para tal efecto se utiliza el valor PCE que es la suma de los residuos al cuadrado (varianza de los residuos) en donde valores pequeños indican predicciones más precisas punto por punto (Mueller *et al.* 2001 y Schloeder *et al.* 2001).

$$PCE = \frac{1}{n} \sum_{i=1}^n \left[z(x_i) - \hat{z}(x_i) \right]^2 \quad (6)$$

También se utilizó una medida de efectividad llamada estimado de efectividad de predicción (E), el cual es un estimado de que tan efectiva fue la predicción, en comparación con el uso solamente del promedio general de los datos (Schloeder *et al.* 2001, Gotway *et al.* 1996). Este parámetro se calcula como se muestra en la ecuación 7.

$$E = \left(1 - \left\{ \frac{\sum_{i=1}^n \left[z(x_i) - \hat{z}(x_i) \right]^2}{\sum_{i=1}^n \left[z(x_i) - \bar{z} \right]^2} \right\} \right) 100 \quad (7)$$

donde \bar{z} es el promedio de la muestra. Cuando E es igual a 100% indica una predicción perfecta y cuando el valor es negativo indica que hubiese sido mejor usar la media general que la predicción. De acuerdo con Kravchenko (2003), si la estructura espacial es pobre y no es viable hacer un muestreo intensivo para crear un buen mapa, es mejor trabajar usando la media general de los datos.

El valor E tiende a ser inversamente proporcional a la estructura espacial expresada como la relación de Co/C (Figura 1), donde una rela-

ción baja indica una estructura espacial pobre (Mueller *et al.* 2001, Kravchenko 2003). Mueller *et al.* (2001) encontraron que el Kriging tuvo una eficiencia marginal (E<48%) cuando el muestreo fue intenso (30 m entre puntos) y pobre cuando el muestreo fue menos intenso (100 m entre puntos) con E<21%. Comparando con el Kriging, el IDW fue superior en la mayoría de los casos, pero con tendencia a disminuir conforme disminuía la distancia de muestreo. Aún así, los autores afirman que el incremento en intensidad de muestreo (<100m) y por consiguiente en costos, no se justifica por el poco mejoramiento en la exactitud de predicción.

Finalmente se utilizó el valor de validación cruzada, el cual es conocido también como el método “jackknife” que remueve consecutivamente un valor de los datos y el valor ausente es interpolado con los datos remanentes para luego ser comparado con el dato removido u observado.

Con los datos interpolados, se graficaron cada uno de los resultados de las interpolaciones y se realizó también una comparación visual entre ambos métodos.

RESULTADOS Y DISCUSIÓN

Los estadísticos estimados a partir de los datos son presentados en el cuadro 1. Como se puede observar, los parámetros que muestran mayor variación son el Ca y el P (46% ambos), mientras que el pH tiene solo 4% de coeficiente de variación. En cuanto a los valores promedio, pH y Ca estuvieron por debajo de 5,5 y 4 cmol(+).kg⁻¹, respectivamente, que son sus niveles críticos; la CICE por encima del nivel crítico que es 5 cmol(+).kg⁻¹ y el P por debajo del mínimo que es 10 mg.kg⁻¹.

Luego de realizar las pruebas de normalidad, de los datos (curtosis y simetría a p<0,05), solamente el P necesitó ser transformado con logaritmo natural. A diferencia del trabajo realizado por Henríquez *et al.* (2005), la variación de los datos de este estudio no fue tan alta, pues ellos reportan valores de hasta un 100%, en el caso del Mg pero en un área mucho mayor. El coefi-

Cuadro 1. Estadísticos preliminares de pH, Ca, CICE, P, de las 61 muestras de suelo utilizadas.

	pH	Ca cmol(+).kg ⁻¹	CICE cmol(+).kg ⁻¹	P Mg.kg ⁻¹
Valor mínimo	4,70	2,34	6,28	3,00
Promedio	5,03	6,41	10,92	6,90
Valor máximo	5,80	11,70	17,82	15,10
Error estándar	0,20	2,97	3,22	3,15
% Coeficiente de variación (CV)	4%	46%	29%	46%
Probabilidad de las pruebas Curtosis y simetría simultáneas	0,0030	0,0000	0,0009	0,0631*

* Para ln(P) la probabilidad de la prueba de curtosis y simetría simultáneas fue de 0,0354.

ciente de variación del pH es bajo en este estudio como también fue observado por otros autores (Henríquez *et al.* 2005, Muñoz *et al.* 2006). Esta tendencia es comprensible en vista de la capacidad buffer del suelo.

Los semivariogramas son mostrados en las figuras de la 2 a la 5. Se encontró una leve tendencia a un comportamiento anisotrópico en algunos de los datos de los semivariogramas realizados; sin embargo, todos fueron usados como isotrópicos por la débil tendencia. En el cuadro 2 se presenta los parámetros de los semivariogramas, como se observa el modelo esférico fue el predo-

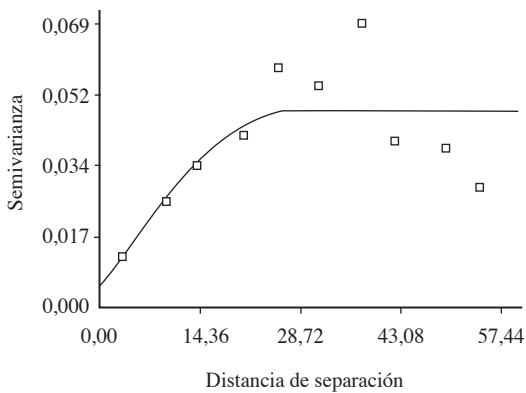


Fig. 2. Semivariograma de los datos de pH, utilizando un modelo esférico para el ajuste.

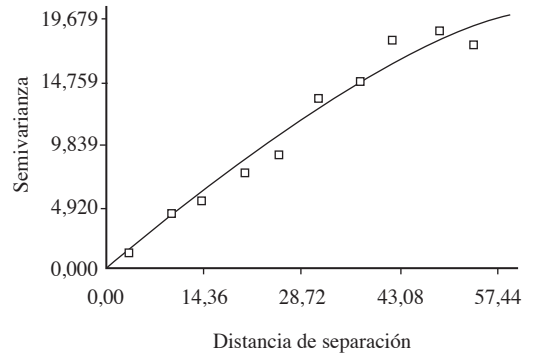


Fig. 3. Semivariograma de los datos de Ca, utilizando un modelo esférico para el ajuste.

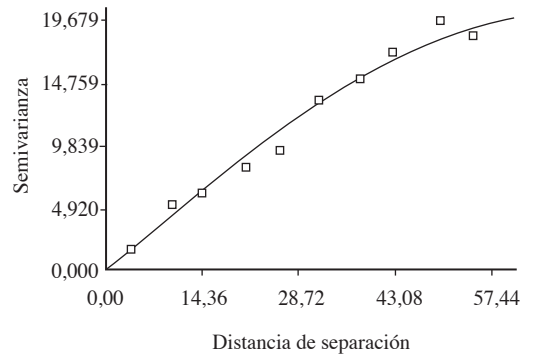


Fig. 4. Semivariograma de los datos de CICE, utilizando un modelo esférico para el ajuste.

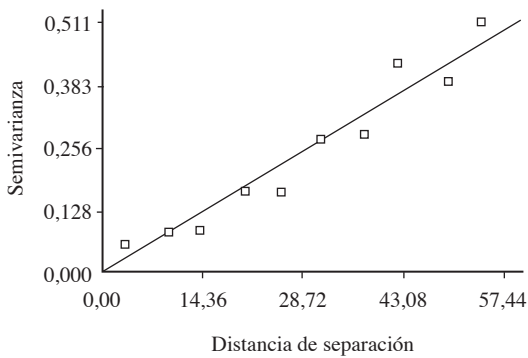


Fig. 5. Semivariograma de los datos transformados de P, utilizando un modelo lineal para el ajuste.

minante, no obstante el lineal se ajustó mejor para el P transformado.

Los mejores modelos fueron seleccionados con base en la suma residual de los cuadrados más baja, la cual está asociada con una mejor estimación de los semivariogramas. Los rangos variaron de 27 hasta 114 m, para pH y ln(P), respectivamente. Este ámbito bajo (27 m) para pH, refleja una baja autocorrelación espacial también encontrada en otros estudios (Muñoz *et al.* 2006, Lozano *et al.* 2004). También se observa una dependencia espacial o estructural aún a 115 m de un punto a otro al medir el P, lo cual sugiere un efecto de autocorrelación fuerte que puede deberse a algún factor de gran influencia (mayor

a la distancia de muestreo) en el sitio. Esta distancia decrece a 74 m aproximadamente con el Ca y a 66 m con la CICE. En cuanto a la proporción $C/(Co+C)$ en todos los casos es bastante alta, reflejando una dependencia espacial alta, que también se observa con un efecto nugget o pepita muy bajo (varianza aleatoria).

Estos datos estarían reflejando que, en el caso del pH, los muestreos podrían ser más distanciados y el cambio en las estimaciones no variaría mucho, como ocurriría si se distancia los puntos muestreados para P, el cual tiene una alta autocorrelación espacial y depende de más cantidad de observaciones a su alrededor para predecir los puntos no muestreados.

En el cuadro 3 se resume el resultado de las interpolaciones utilizando el IDW y el Kriging. En el caso del IDW, se puede observar como al aumentar el exponente que da peso a los valores cercanos (p), el promedio de las estimaciones aumenta al igual que el error estándar. El coeficiente de determinación (R^2) tiende a aumentar, pero más relevante aún, la predicción del error estándar tiende a decrecer, con excepción del ln(P) el cual tiene un comportamiento opuesto. Al comparar los resultados del Kriging con los del IDW, las predicciones del error estándar son menores en el Kriging con excepción del ln(P).

El cuadro 4 muestra los estadísticos calculados para determinar cuál interpolador es más

Cuadro 2. Parámetros de los variogramas para pH, Ca, CICE y ln(P).

	pH	Ca	CICE	ln (P)
Modelo de ajuste	esférico	esférico	esférico	lineal
Efecto pepita (Co: nugget)	0,0049	0,0100	0,0400	0,0010
Meseta (Co+C: sill)	0,048	18,770	20,070	0,958
Rango	27,100	74,400	66,770	114,000
Proporción $C/(Co+C)$	0,897	0,999	0,998	0,999
Coficiente de determinación: R^2	0,560	0,970	0,981	0,944
Suma residual de los cuadrados	0,0011	8,5700	7,0300	0,0115

C: varianza estructural.

Cuadro 3. Resultados de las interpolaciones y validación cruzada con los métodos Kriging y el IDW.

	pH	Ca	CICE	ln (P)
Promedio observado	5,03	6,41	10,92	1,83
Desviación estándar	0,20	2,97	3,22	0,46
Kriging				
Promedio estimado	5,054	7,276	11,837	1,745
Desviación estándar del promedio estimado	0,024	8,922	9,759	0,104
Validación cruzada				
Coefficiente de la regresión	1,056	1,005	1,032	1,161
Error estándar	0,121	0,060	0,068	0,117
R ²	0,563	0,826	0,795	0,626
intercepto	-0,277	-0,007	-0,330	-0,262
Predicción del error estándar	0,135	1,236	1,465	0,278
IDW (p=1)				
Promedio estimado	5,022	6,699	11,172	1,749
Desviación estándar del promedio estimado	0,012	5,691	5,923	0,123
Validación cruzada				
Coefficiente de la regresión	1,352	1,078	1,103	1,100
Error estándar	0,201	0,101	0,116	0,104
R ²	0,434	0,658	0,604	0,656
intercepto	-1,737	-0,135	-0,665	-0,159
Predicción del error estándar	0,154	1,735	2,035	0,266
IDW (p=2)				
Promedio estimado	5,033	6,858	11,347	1,753
Desviación estándar del promedio estimado	0,019	6,705	7,178	0,147
Validación cruzada				
Coefficiente de la regresión	1,134	1,043	1,063	0,970
Error estándar	0,156	0,080	0,089	0,094
R ²	0,473	0,741	0,708	0,643
intercepto	-0,650	-0,040	-0,392	0,073
Predicción del error estándar	0,149	1,508	1,748	0,271
IDW (p=3)				
Promedio estimado	5,039	6,955	11,449	1,755
Desviación estándar del promedio estimado	0,024	7,485	8,139	0,161
Validación cruzada				
Coefficiente de la regresión	0,974	0,989	0,990	0,859
Error estándar	0,134	0,069	0,075	0,090
R ²	0,471	0,778	0,746	0,607
intercepto	0,146	0,231	0,307	0,277
Predicción del error estándar	0,149	1,395	1,631	0,284

Cuadro 4. Valores de PAE, CME y E de las interpolaciones efectuadas por los métodos Kriging y el IDW a 3 diferentes exponentes (p).

	pH	Ca	CICE	ln (P)
Kriging				
PAE	0,094	0,91	1,06	0,28
PCE	0,02	1,51	2,05	0,15
E	56,40	82,59	93,33	99,33
IDW (p=1)				
PAE	0,103	1,38	1,60	0,20
PCE	0,03	3,10	4,30	0,07
E	38,9	64,10	86,04	99,65
IDW (p=2)				
PAE	0,10	1,12	1,29	0,19
PCE	0,02	2,30	3,11	0,07
E	45,72	73,40	89,91	99,65
IDW (p=3)				
PAE	0,10	0,97	1,13	0,20
PCE	0,02	1,94	2,66	0,08
E	46,27	77,53	91,37	99,60

preciso y con menos errores. El método Kriging es más preciso (valores más bajos de PAE y PCE) y eficiente (valores más altos de E) para las interpolaciones de pH, Ca, y CICE. Aún así, el IDW tiende a mejorar conforme el valor de p aumenta hasta 3, pero siempre por debajo del Kriging. En el caso del ln(P), a pesar de tener un coeficiente de variación alto (46%), igual que el Ca, las interpolaciones son más precisas y eficientes con el método IDW, pero los valores de E son muy similares entre ambos interpoladores.

Las figuras 6 y 7 muestran los mapas generados por el Kriging ordinario y el IDW (p=2) para la variable de pH. Es evidente como la interpolación de los datos es menos brusca con el Kriging ordinario que con el IDW. El IDW tiende a crear puntos de concentración (“ojo de buey”) en los extremos Sur y Este del área muestreada.

Las áreas con estimaciones de pH bajas, visualmente, tienden a ser las mismas entre ambos interpoladores, no obstante el IDW tiene una mayor tendencia a formar islas. Esta tendencia es común en el IDW al utilizarse exponentes (p) bajos y tener estimaciones más simples, donde tienen menor influencia los datos cercanos al punto estimado.

En las figuras 8 y 9 las estimaciones de Ca tienden a presentar el mismo efecto cuando el IDW es utilizado. El Kriging vuelve a crear un mapa con áreas de transición menos abruptas que las del IDW. Sin embargo, las áreas de mayor y de menor concentración de Ca son muy similares. Un fenómeno muy parecido se observa en las figuras 10 y 11, con el CICE.

En el caso del P, ambos mapas tienden a ser similares (Figuras 12 y 13). Las transiciones

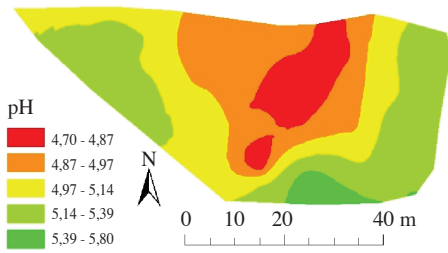


Fig. 6. Mapa obtenido mediante el interpolador Kriging ordinario de la variable pH, a partir de 61 puntos.

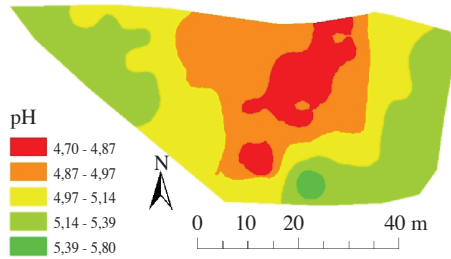


Fig. 7. Mapa obtenido mediante el interpolador IDW ($p=2$) de la variable pH, a partir de 61 puntos.

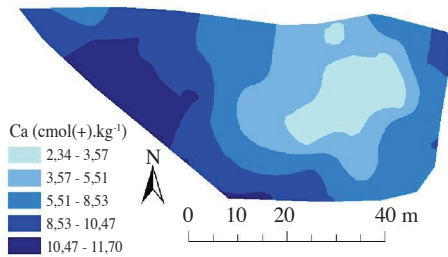


Fig. 8. Mapa obtenido mediante el interpolador Kriging ordinario de la variable Ca, a partir de 61 puntos.

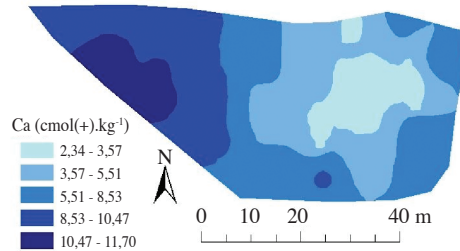


Fig. 9. Mapa obtenido mediante el interpolador IDW ($p=2$) de la variable Ca, a partir de 61 puntos.

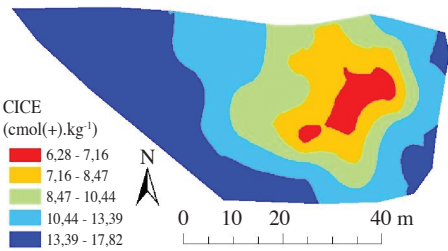


Fig. 10. Mapa obtenido mediante el interpolador Kriging ordinario de la variable CICE, a partir de 61 puntos.

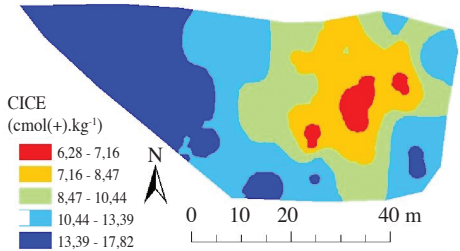


Fig. 11. Mapa obtenido mediante el interpolador IDW ($p=2$) de la variable CICE a, partir de 61 puntos.

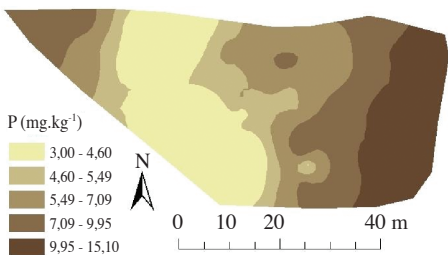


Fig. 12. Mapa obtenido mediante el interpolador Kriging ordinario de la variable P, a partir de 61 puntos.

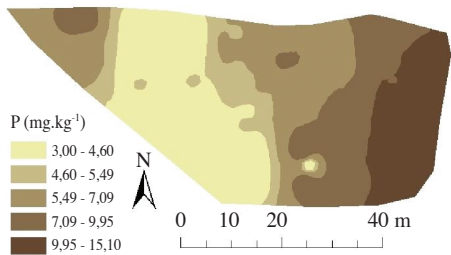


Fig. 13. Mapa obtenido mediante el interpolador IDW ($p=2$) de la variable P, a partir de 61 puntos.

estimadas por el Kriging son más marcadas que en los mapas anteriores, lo que indica que el IDW empieza a ser ligeramente mejor. Como se observó anteriormente, los coeficientes de determinación de las validaciones cruzadas, y los valores de PAE, CME y E son ligeramente mejores para el IDW que el Kriging ordinario en este caso particular.

CONCLUSIONES

El método Kriging fue más preciso y eficiente que el IDW en el proceso de interpolación respecto del pH, el Ca, y el CICE, aunque la diferencia entre ambos métodos no fue muy amplia. Posiblemente, la alta densidad de muestreo benefició a ambos métodos haciendo difícil determinar cuál fue mejor, lo cual pudo haber sido más marcado a una menor densidad de muestreo.

El Kriging proporciona un análisis más elaborado y con un fundamento estadístico, por lo que este método puede ser el preferido cuando la intensidad de muestreo es mayor, aunque esto significaría costos mayores.

Cuando el distanciamiento es muy grande, los variogramas no son posibles de obtener, entonces el Kriging deja de ser una opción y comparativamente el IDW se perfila como el mejor.

La incorporación de más parámetros del suelo sería de interés, para observar cuál es el comportamiento general entre ambos interpoladores. Para ampliar estudios futuros, el análisis podría consistir en la comparación entre diferentes unidades de suelo y determinar si el tipo o el manejo del suelo pueden afectar la variación espacial de las variables y por ende la eficiencia del uso de interpoladores matemáticos o geoestadísticos.

LITERATURA CITADA

- BENMOSTEFA F.Z. 2006. Estimating soil contamination with Kriging interpolation method. *American Journal of Applied Sciences* 3(6):1894-1898.
- BURROUGH P.A., MCDONNELL.1998. Principles of geographical information systems. Oxford University Press. 333 p.
- DEMERS M.N. 1999. Fundamentals of geographic information systems. 2 ed. Wiley. 498 p.
- GOTWAY C.A., FERGUSON R.B., HERBERT G.W., PETERSON, T.A. 1996. Comparison of Kriging and Inverse Distance Methods for mapping soil parameters. *Soil Science Society of American Journal* 60:1237-1247.
- HENRIQUEZ C., KILLORN R., BERTSCH F., SANCHO F. 2005. La Geoestadística en el estudio de la variación espacial de la fertilidad del suelo mediante el uso del interpolador Kriging. *Agronomía Costarricense* 29(2):73-81.
- JOHNSTON K., VER HOEF J.M., KRIVORUCHKO K., LUCAS N. 2001. Using ArcGis Geostatistical Analyst. ESRI. 300 p.
- KRAVCHENKO A. 2003. Influence of spatial structure on accuracy of interpolation methods. *Soil Science Society of American Journal* 67:1564-1571.
- KRAVCHENKO A., BULLOCK D.G. 1999. A Comparative study of interpolation methods for mapping soil properties. *Agronomy Journal* 91:393-400.
- LOZANO Z., BRAVO C., OVALLES F., HERNÁNDEZ R.M., MORENO B., PIÑANGO L., VILLANUEVA J.G. 2004. Selección de un diseño de muestreo en parcelas experimentales a partir del estudio de la variabilidad espacial de los suelos. *Bioagro* 16(1):1-17.
- MUELLER T.G., PIERCE F.J., SCHABENBERGER O., WARNEDE D.D. 2001. Map quality for site-specific fertility management. *Soil Science Society of American Journal* 65:1547-1558.
- MUELLER T.G., PUSULURI N.B., MATHIAS K.K., CORNELIUS P.L., BARNHISEL R.I. 2004. Site-specific soil fertility management: a model for map quality. *Soil Science Society of American Journal* 68:2031-2041.
- MUÑOZ J.D., MARTÍNEZ L.J., GIRALDO R. 2006. Variabilidad espacial de propiedades edáficas y su relación con el rendimiento en un cultivo de papa (*Solanum tuberosum* L.). *Agronomía Colombiana* 24(2):1-16.
- OLIVER M. 1987. Geostatistics and its applications to soil science. *Soil and Management* 3(1):8-19.
- SCHLOEDER C.A., ZIMMERMAN N.E., JACOBS, M.J. 2001. Comparison of methods for interpolating soil properties using limited data. *Soil Science Society of American Journal* 65:470-479.

