



Colombia Forestal

ISSN: 0120-0739

Proyecto Curricular de Ingeniería Forestal, Facultad
del Medio Ambiente y Recursos Naturales, Universidad
Distrital Francisco José de Caldas.

Andrades, Jesús; Cuesta, Ledyz; Camargo,
Cristopher; López, Juan; Torres, Hugo; Osorio, Arturo
Propuesta metodológica para la construcción y selección
de modelos digitales de elevación de alta precisión
Colombia Forestal, vol. 23, núm. 2, 2020, Julio-Diciembre, pp. 34-46
Proyecto Curricular de Ingeniería Forestal, Facultad del Medio Ambiente
y Recursos Naturales, Universidad Distrital Francisco José de Caldas.

DOI: <https://doi.org/10.14483/2256201X.15155>

Disponible en: <https://www.redalyc.org/articulo.oa?id=423964938003>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UDEM
redalyc.org

Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS



Propuesta metodológica para la construcción y selección de modelos digitales de elevación de alta precisión

Methodological proposal for the construction and selection of digital high-precision elevation models

Jesús Andrades¹, Ledyz Cuesta², Cristopher Camargo³, Juan López⁴,
Hugo Torres⁵, Arturo Osorio⁶

Recepción: 26 de julio de 2019

Aprobación: 21 de mayo de 2020

Andrades, J., Cuesta, L., Camargo, C., López, J., Torres, H. y Osorio, A. (2020). Propuesta metodológica para la construcción y selección de modelos digitales de elevación de alta precisión. *Colombia Forestal*, 23(2), 34-46.

Resumen

Los modelos digitales de elevación mejoran la precisión en el análisis del terreno. Este trabajo tuvo como objetivo comparar diferentes métodos de interpolación utilizando criterios estadísticos en un levantamiento topográfico de 339 puntos levantados con estación total en la zona de San Juan de Lagunillas (Venezuela). Se documentó la existencia de un proceso de tendencia significativa con autocorrelación espacial modelada mediante un semivariograma gaussiano isotrópico. Los kriging consiguieron residuales menores a un error mínimo tolerable igual 0.1 m. El kriging residual fue el mejor evaluado en la validación cruzada con exclusión (error medio absoluto de 0.71 m). El kriging disyuntivo tuvo menor media del error pero un amplio rango del error y el mayor error medio absoluto (12.8 m). Los otros kriging tienen medias de la varianza del error superiores a 1.3 m y errores mínimos absolutos superiores a 0.85 m. Debido a la alta dependencia en la forma de especificación de los modelos se recomiendan las ayudas gráficas para un mejor ajuste.

Palabras clave: geoestadística, MDE, método de interpolación, semivariograma.

Abstract

Digital Elevation Models improve accuracy in terrain analysis. This work compared different interpolation methods using statistical criteria in a topographic survey of 339 points raised with a total station in the "San Juan de Lagunillas" area (Venezuela). The existence of a significant trend process with spatial autocorrelation modeled by an isotropic Gaussian semivariogram was documented. The kriging allowed to obtain residuals less than a minimum tolerable error equal to 0.1 m. The residual kriging was the best evaluated in cross validation with exclusion (absolute mean error of 0.71 m). The disjunctive kriging showed a lower error average but a wider error range and the highest mean absolute error (12.8 m). The other kriging showed mean error variance greater than 1.3 m and absolute minimum errors greater than 0.85 m. Due to the high dependence on the model specification form, graphics are recommended for a better fit.

Keywords: DEM, geostatistics, interpolation method, semivariogram.

- 1 Universidad de Los Andes, Venezuela. andradesjesus@yahoo.com
- 2 Universidad Católica del Maule. Talca, Chile. lecuestaherrera@gmail.com
- 3 Autor para correspondencia cristophercamargo_roa@hotmail.com
- 4 jllopez@ula.ve
- 5 Universidad de Santander. Bucaramanga, Colombia. Al.torres@mail.udes.edu.co
- 6 raosorio@gmail.com

<https://doi.org/10.14483/2256201X.15155>

INTRODUCCIÓN

Los datos espaciales para variables continuas desempeñan un papel preponderante en la planificación forestal, ambiental y territorial. Estos suelen obtenerse de fuentes puntuales tomadas en campo o mediante sensores remotos (Li y Heap, 2008). Una variable continua de especial interés es la altitud, la cual puede ser representada con el modelo digital de elevación (MDE); esta es pieza clave del análisis geomorfológico y equivalente informatizado de la cartografía clásica de elevaciones, tradicionalmente representada mediante curvas de nivel. Esto supone un enorme salto cualitativo frente a sus predecesores, pues al utilizar un modelo numérico en lugar de uno gráfico se permite una diferencia conceptual de análisis superior, en este caso, a la que tiene lugar con otras capas de información (Olaya, 2011).

Formalmente, los MDE son una estructura numérica de datos que representa la distribución espacial de la altitud de la superficie del terreno (Felicísimo, 1994). Sin embargo, estos presentan como problema fundamental la precisión que deben tener para la representación adecuada del espacio geográfico a una determinada escala cartográfica (Wang, 2005), por ello, se necesitan métodos que representen adecuadamente los datos originales (Felicísimo, 1994). La creación de MDE implica, en la mayoría de casos, la utilización de métodos de interpolación (MI), los cuales corresponden a modelos matemáticos que pronostican valores en ubicaciones donde no hay muestra disponible y que se pueden clasificar en tres categorías: no geoestadísticos, geoestadísticos e híbridos (Lloyd, 2010).

Los MI geoestadísticos se basan fundamentalmente en que los datos colectados en puntos cercanos tienden a ser más similares que los colectados a mayor distancia (Mantovani y Magdalena, 2014); esto es la manifestación de la ley geográfica de Tobler (1970), la cual indica que: “Todo está relacionado con todo pero las cosas más cercanas están más relacionadas que las distantes” (p. 3). Lo que

lleva a afirmar que este tipo de método considera un modelo geográfico de variación espacial, el cual contiene al menos tres componentes: 1) una estructura general llamada tendencia; 2) una estructura superimpuesta, relacionada con la correlación espacial y con una variación gradual; y 3) un componente que consiste en una variación al azar causada por errores de muestreo o variaciones espaciales a escalas menores que la red de muestras (Mantovani y Magdalena, 2014). Con ello se puede aseverar que, mediante el uso de la dependencia espacial que tienen los datos de una muestra, la geoestadística emplea un algoritmo que cuantifica una variable dada en lugares no muestreados utilizando métodos probabilísticos, lo cual da una ventaja teórica sobre los MI no geoestadísticos. En la Geoestadística se parte del concepto estocástico de que los datos son un proceso espacial aleatorio $Z(x):x \in R$ donde R corresponde a una colección continua de localizaciones espaciales siendo el espacio de interés (D), un subconjunto fijo de \mathbb{R}^d que contiene un rectángulo d -dimensional con volumen positivo (Matheron, 1970; Cressie, 1992).

Partiendo de estas ideas, cabe preguntarse si es posible mejorar la precisión de los MDE. Para ello, se definieron como objetivos de investigación proponer una metodología que permita obtener un MDE con mayor precisión a partir de métodos estadísticos y determinísticos, que conceda a su vez la estimación de la escala cartográfica de forma indirecta y la disminución del error en los MI.

MATERIALES Y MÉTODOS

Se dispuso de un levantamiento topográfico de 339 puntos, levantados con estación total Geodimeter System 500 de una zona de San Juan de Lagunillas, Estado Mérida, Venezuela (figura 1). Obtenidos los datos en campo, se realizó un análisis exploratorio de los mismos para evaluar el comportamiento de la distribución de frecuencia; lo que implicó la estimación de los cuatro momentos de una variable cuantitativa y continua (media, desviación

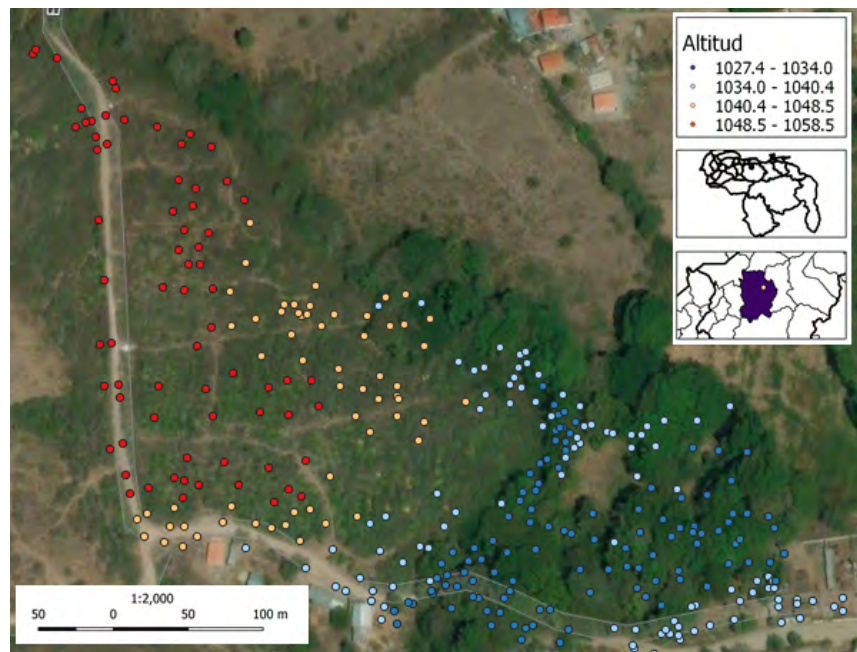


Figura 1. Distribución espacial de las 339 muestras levantadas.

estándar, coeficiente de asimetría y curtosis). Para ello se utilizó el lenguaje de programación con capacidad estadística R y la librería Rcmdr (Field *et al.*, 2012; Fox, 2005; R Core Team, 2019).

Dado que el interés de este estudio involucra aplicar un enfoque geoestadístico predictivo que involucra técnicas de tipo lineal, la normalidad de los datos es una condición deseable. Por el contrario, cuando esta no se alcanza los MI geoestadísticos pueden estar afectados negativamente por datos atípicos aditivos (Giraldo, 2002; Díaz, 2002; Bivand *et al.*, 2013; Emery, 2013). En caso de no cumplir este supuesto una opción es transformar los datos y, posteriormente, invertir la transformación, mediante correcciones para el sesgo y la varianza. Sin embargo, este es un método no tan plausible si la transformación no es eficiente (Díaz, 2002), ya que la normalidad en muchas ocasiones no es una suposición sólida, y no se puede probar en práctica (Diggle y Ribeiro, 2002). En este sentido, se evaluó el comportamiento de la distribución de frecuencia de los datos y posibles transformaciones a los mismos.

Ante la posibilidad de que se manifestara la ley geográfica de Tobler (1970) se procedió a realizar

un análisis exploratorio de datos espaciales. Para ello se usó el concepto sugerido por Bivand *et al.* (2013) de realizar diagramas de dispersión entre los pares $Z(x_i)$ y $Z(x_j)$, agrupados según su separación de distancia $h_{ij} = \|x_i - x_j\|$, con una distancia horizontal euclidiana de 10 m, que determinó de forma exploratoria (no confirmatoria) la naturaleza de la autocorrelación espacial (lineal, no lineal e híbrida) y el aproximado de la distancia hasta donde se percibe autocorrelación espacial.

El problema de interpolación fue abordado desde dos enfoques generales. El primero corresponde al determinístico (no geoestadístico) tradicional y el segundo a un concepto estocástico de una aproximación geoestadística (Giraldo, 2002). El enfoque tradicional aplicado corresponde con la ejecución de la aproximación del MI no geoestadístico, siendo uno de los más utilizados el inverso de la distancia ponderada. Que se define como: $\hat{z} = \sum_{i=1}^n p_i z_i^k / \sum_{i=1}^n p_i^k$, donde z es el valor por predecir, z_i es valor observado o conocido, p_i es el peso asignado al punto i -ésimo, el cual es estimado utilizando el método de la ponderación por distancia inversa, es decir $p_i = 1/d_i^k$, donde d corresponde a

la distancia entre puntos y k el exponente que toma habitualmente el valor 2 (Felcísimo, 1994; Olaya, 2011; Li y Heap, 2008; Lloyd, 2010).

Para la correcta ejecución de un MI geoestadístico de la familia Kriging resulta necesario establecer la estacionariedad en segundo orden como regla simplificadora (medias y varianzas constantes en el espacio geográfico; Matheron, 1970; Cressie, 1992). Por ello, para demostrar si los datos poseen un proceso estacionario por media o en su defecto tienen tendencia, se procedió a realizar diversos ajustes a superficies de tendencias (polinómicas de primer a cuarto orden) y a estimar sus residuales. Lo que conduce en la práctica a la solución de dos problemas: la predicción del orden k del polinomio el cual es desconocido y la estimación a partir de los residuales (datos-deriva) del semivariograma, pues no es advertido (Díaz, 2002).

Tanto para los datos originales como los residuales (datos-deriva) se realizó el análisis estructural de dependencia espacial, el cual permite estimar el grado de dependencia espacial de un proceso (Giraldo, 2011). Esto es estimar la función de semivarianza experimental, que corresponde con una función de las diferencias entre dos lugares y la distancia que los separa (Matheron, 1970; Cressie, 1992; Moreno, 2008):

$$\gamma(h) = \frac{1}{2N(h)} \sum [Z(x) - Z(x+h)]^2$$

Donde $\gamma_{(h)}$ es la función de semivarianza para todas las muestras localizadas en el espacio separado por el intervalo de distancia h , dentro de este parámetro $N_{(h)}$ viene a ser el número total de pares de muestras separados por un intervalo de distancia h , siendo $Z_{(x)}$ el valor de la muestra en una localización x y $Z_{(x+h)}$ el valor de la muestra a la distancia de intervalo h desde x .

Según Royle (1980), Armstrong (1998), Wackernagel (1998), Diggle y Ribeiro (2007), Chilès y Delfiner (2012), Oliver y Webster (2015) y Montoro et al. (2015), para el ajuste de un modelo de semivariograma los parámetros que deben estimarse para describir son: el rango (r): máxima distancia

hasta donde existe dependencia espacial. La meseta (también conocida como saturación o alféizar), (s): máxima variabilidad en ausencia de dependencia espacial, y el *nugget*, pepita o núcleo (n): variabilidad que no puede explicarse mediante la estructura espacial, es decir, la variabilidad estimada a una distancia nula.

Resulta necesario señalar la existencia de casos en los que el semivariograma experimental solo depende de la distancia sin evidencia de cambios con la dirección, ante esto se estará en presencia de una estructura de dependencia isotrópica (Díaz, 2002). En este sentido, para evaluar si los datos poseen varianza no constante se generó una superficie variográfica que indicaría la presencia de un proceso de isotropía o anisotropía, así como el comportamiento de la función de semivarianza en ángulos azimutales.

Luego, se procedió al ajuste de diversos modelos de semivariograma. Dicho método utiliza una regresión no lineal para estimar los coeficientes de forma iterativa, lo que minimiza la suma ponderada de errores al cuadrado (Bivand et al., 2013). Para ello, se usaron los principios planteados por Olaya (2011) y Bivand et al. (2013); estos son que los modelos de semivariogramas ajustados deben cumplir con los siguientes criterios: 1) ser monótono y creciente; 2) tener un máximo constante o asintótico, es decir, un valor definido de la meseta o alféizar (funciones no acotadas superiormente, tales como las exponenciales, indicarían que la zona de estudio no es suficientemente grande, ya que no alcanza la dimensión a partir del cual el efecto de la dependencia espacial deja de existir); y 3) el *nugget* debe ser positivo. Con los modelos de semivariogramas ajustados se procedió a realizar diversos tipos de MI Kriging lineales y no lineales.

Según Olaya (2011), el MI conocido usualmente como Kriging es semejante a un método basado en ponderación por distancia. Esta familia de MI establece las siguientes variantes lineales de solución del Kriging: simple, ordinario, universal, residual, log normal, multigaussiano y disyuntivo (Giraldo, 2002). Los dos primeros definen que el

proceso es estacionario en segundo orden, asumiendo además en el caso del simple que la media poblacional es conocida (Cressie, 1992). A diferencia de estos, el universal y residual son más flexibles, ya que no asume un proceso estacionario por media. Es decir, la variable medida posee tendencia, lo que conlleva a definir el grado de la función polinómica lineal que lo describe (Cressie, 1992; Bivand et al., 2013). En concordancia con esta idea, el Kriging residual, método empleado bajo las mismas circunstancias del universal, cuyo funcionamiento consiste en remover la tendencia y a partir de los residuales estimar una variante del Kriging estacionario (simple u ordinario), la predicción en un sitio no muestreado que es igual a la tendencia estimada más la predicción del error (Giraldo, 2002; Li y Heap, 2008).

Es importante acotar que, en los procesos lineales de los MI geoestadísticos, cuando la función aleatoria $Z_{(x)}$ sigue un proceso Gaussiano el mejor predictor es un estimador lineal, de lo contrario, se puede aplicar una transformación adecuada que convierta los datos originales (no normales) en gaussianos (con posibles datos atípicos aditivos) (Giraldo, 2002). Cuando este tipo de transformación no es clara aparecen propuestas desde numerosos enfoques, tales como: el Kriging log normal y el multigaussiano, en estos dos procedimientos se hacen transformaciones de la variable regionalizada con el propósito de normalizar la distribución de frecuencia (Giraldo, 2011).

Otro enfoque que aparece es el Kriging disyuntivo, considerado cuando para la variable principal las transformaciones convencionales no pueden producir una distribución normal. En este tipo de Kriging se utilizan polinomios de Hermite (Li y Heap, 2008), siendo el resultado un tipo de estimador no lineal (Giraldo, 2002). La estimación Kriging disyuntivo se obtiene mediante un Cokriging simple de las funciones de diferente orden (Ortiz y Deustsch, 2005).

Como proceso de validación se utilizó, en primer lugar, el método de validación cruzada con exclusión (LOOCV, por sus siglas en inglés), este

es un método iterativo que realiza estimaciones a partir del modelo ajustado pero utilizando k-1 observaciones. De esta manera, la observación restante omitida es estimada con el modelo y con ello se puede estimar el error, el cual difiere en su descripción de la validación cruzada tradicional porque repite otras k-1 veces con una observación diferente omitida (Lachenbruch y Mickey, 1968); aunque es más demandante computacionalmente por tener una aproximación más cercana a la máxima verosimilitud del error. El otro método de validación utilizado fue el error medio absoluto (MAE, por sus siglas en inglés) cuyo estimador es $MAE = \sum |estimado - observado|/n$ (Wilmott y Matsuura, 2005). Estos resultados fueron comparados con el error máximo tolerable (EMCT), el cual se estimó utilizando el criterio planteado por el Laboratorio de Fotogrametría y Sensores Remotos (2007). Adicionalmente, para los modelos estocásticos de interpolación debe cumplirse con los diagnósticos tradicionales de un modelo de esta naturaleza, esto es que los residuales sigan los siguientes comportamientos: normalidad, homocedasticidad e independencia (Guajarati y Porter, 2010). Para ello se realizaron diversos gráficos exploratorios, el test de Shapiro-Wilk de normalidad, el test de Breusch-Pagan de homocedasticidad y la verificación del semivariograma residual, es decir, la comprobación de que los supuestos básicos de un modelo paramétrico se cumplen.

RESULTADOS

Los datos de elevación del terreno obtenidos a través de la estación total denotaron una media de 1039.73 m de altitud, una desviación estándar de 7.99 m y un coeficiente de asimetría positivo (0.5697), configurados en una distribución no normal (figura 2a) de tipo peine, refrendado también por el diagrama Q-Q (figura 2b), el cual indica que los mismos se encuentran por fuera de los límites de los intervalos de confianza de una distribución normal ($p < 0.05$) (líneas punteadas), sin presencia de outliers.

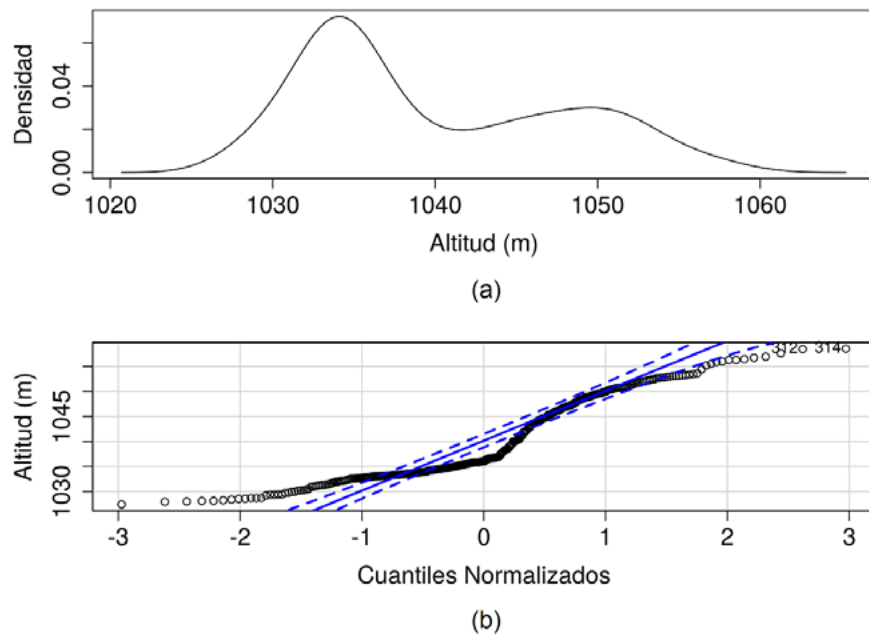


Figura 2. Comportamiento de la distribución de frecuencia de la altitud.

Los resultados de los diagramas de dispersión establecidos a cada 10 m (figura 3) indicaron un proceso de autocorrelación espacial lineal hasta el rango de distancia de 70 m, con un coeficiente de correlación lineal que llega hasta un valor R de 0.71, indicativo de la existencia de un fuerte proceso de autocorrelación hasta esta distancia, en un rango de variabilidad de 0.97 a 0.31, a partir del cual el proceso se convierte aparentemente en un proceso no lineal. Del mismo modo, el semivariograma experimental indicó un proceso monótono y creciente hasta los 400 m, adquiriendo un máximo asintótico donde la meseta (o alféizar) también se alcanza, es decir, se obtiene el punto máximo y a partir de este se asume que la dependencia espacial no es significativa. Sin embargo, se evidencia que el proceso entre 300 m y 400 m tiene un comportamiento aparentemente heterocedástico. Con relación al *nugget* del semivariograma, este fue muy cercano a cero, resultado esperado dado que los datos fueron obtenidos con un levantamiento de alta precisión, lo que permite asumir la variabilidad no explicada por el muestreo cercano a 0.

Con miras al semivariograma experimental se ajustó un modelo gaussiano con y sin corrección por anisotropía (figuras 4a y 4b). Ambos resultados conllevan a inferir que se está ante la presencia de un proceso isotrópico, dado que se denota una mejora en el ajuste del semivariograma cuando se asume este proceso.

Los resultados del modelo gaussiano isotrópico mostraron un *nugget* muy cercano a cero unidades de semivarianza (1.41), indicativo de que la variabilidad no explicada por el modelo es muy baja y una meseta (alféizar) que posee una semivarianza de 226.99 y un rango de 205.68 m. Con este modelo de semivariograma se logró el mejor ajuste inicial (error de ajuste de 34.61 unidades de semivarianza), alcanzado con la modificación de los parámetros del intervalo estimado.

En cuanto al análisis de los residuales obtenidos mediante los MI aplicados (tabla 1) se desprende que el Kriging Simple Residual Isotrópico fue el más acertado, pues no solo posee uno de los errores más bajos (LOOVC de 0.008 m y MAE de 0.70 m), sino que también su máximo y mínimo

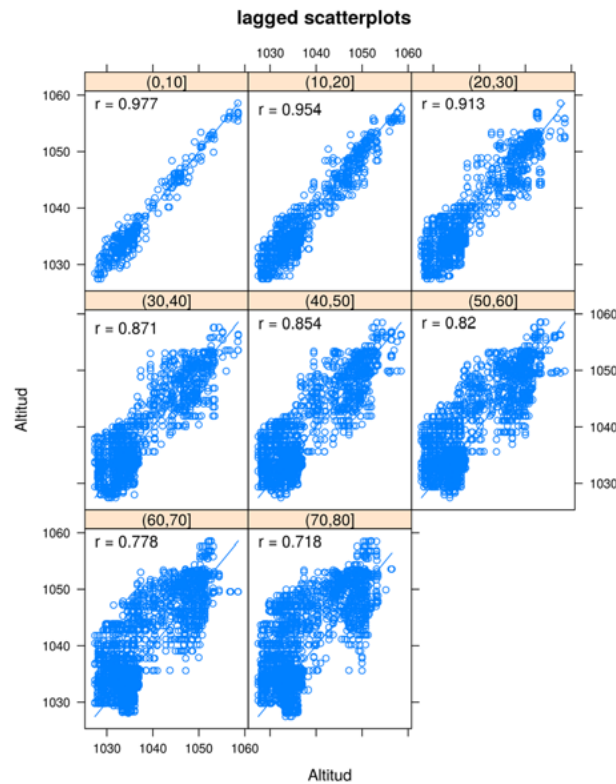


Figura 3. Gráficos de dispersión con rezago espacial.

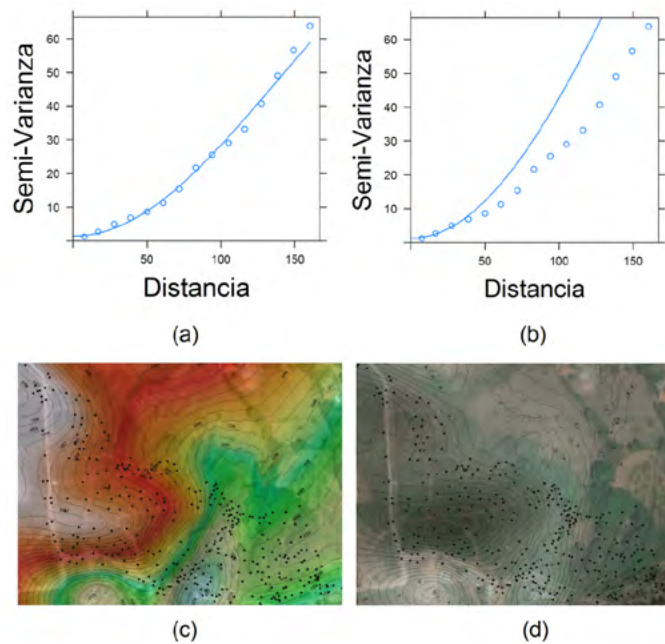


Figura 4. a) Semivariograma experimental con intervalo de 0.5 m con modelo gaussiano isotrópico; b) semivariograma experimental con intervalo de 0.5 m con modelo gaussiano anisotrópico ajustado; c) predicción de Kriging Simple Residual.; d) varianza de Kriging Simple Residual.

oscila entre 2 y 3 m, siendo menor que los resultados presentados por el Kriging Universal Isotrópico. Por ello, se podrían generar curvas de nivel a una equidistancia de 0.1 m, lo cual es indicativo que el proceso que presentan los datos es de tipo isotrópico y con tendencia. Esto va en concordancia con los resultados obtenidos en el ajuste del semivariograma en las figuras 4a y 4b, ya que es evidente que el modelo mejora sustancialmente en la medida en que se asume la isotropía de los datos. Así mismo, debe descartarse el Kriging disyuntivo, ya que a pesar de que posee el indicador LOOVC es también el que más rango tiene, esto es refrendado por el valor estimado del MAE que manifiesta un valor extremadamente elevado. En las figuras 4c y 4d se muestra la predicción y las varianzas respectivamente del Kriging simple

residual isotrópico, siendo evidente que el proceso posee varianzas mayores donde no hay muestras, o en su defecto donde la densidad de las observaciones es menor. Adicionalmente, el diagnóstico de este MI manifiesta un proceso homocedástico sin autocorrelación espacial residual, lo que indica que este presenta un buen rendimiento.

Finalmente, tanto las versiones del Kriging disyuntivo como el inverso de la distancia fueron los que menor desempeño presentaron, siendo un indicativo de que la autocorrelación es lineal y modelable desde el punto de vista estadístico, y que el buen desempeño del Kriging simple residual isotrópico se debe a que la cantidad de muestras es tan elevada que permite que la media muestral se acerque a la media poblacional, información suficiente para ser ajustado de forma correcta.

Tabla 1. Resultados de los métodos de interpolación.

Método de Interpolación	Condición				Error (m)		Varianza (m ²)			MAE (m)
	Tendencia	Media	Var	Med	Min	Max	Med	Min	Max	
Kriging Ordinario	No	No Conocida	Iso	-0.04	-3.61	3.97	1.30	1.37	1.98	0.9369
			Aniso	0.04	-3.22	3.76	1.38	1.31	2.50	0.8527
	Si (Residual)	No Conocida	Iso	-0.013	-2.72	2.88	1.16	1.01	2.73	0.7248
Kriging Simple	No	Conocida	Iso	-0.02	-3.78	4.21	1.36	1.30	1.96	10.076
			Aniso	0.03	-3.07	3.25	1.37	1.30	2.06	0.8887
	Si (Residual)	Conocida	Iso	0.008	-2.64	3.14	1.14	1.01	2.56	0.7091
Kriging Universal	Lineal Primer Orden	Con Tendencia	Iso	-0.027	-4.64	3.19	1.37	1.31	2.00	0.9342
			Aniso	0.02	-3.75	3.82	1.39	1.31	2.84	0.8598
	Polinomial de Segundo Orden X^2+Y^2	Con Tendencia	Iso	-0.006	-3.49	3.83	1.37	1.31	2.16	0.9255
			Aniso	0.04	-3.34	3.68	1.39	1.31	2.18	0.8669
Kriging Disyuntivo	Si (Residual)	Conocida	Iso	-0.0006	-6.69	6.19	0.61	0.29	1.5	1039.73
Inverso de la Distancia $p=2$	NA	NA	NA	0.13	NA	NA	NA	NA	NA	NA
Inverso de la Distancia $p=1$	NA	NA	NA	0.24	NA	NA	NA	NA	NA	NA

DISCUSIÓN

Un primer elemento de discusión es la escala cartográfica del levantamiento, que se estableció en aproximadamente 1:400 por medio del semivariograma. Esta afirmación se basa en el criterio planteado por [Gallardo \(2006\)](#) y [Kerry et al. \(2010\)](#), quienes manifiestan que es posible utilizar el semivariograma y las ecuaciones del Kriging para cuantificar el grado, la escala cartográfica y de variación espacial, el intervalo de muestreo y el error tolerable específico. El método descrito es un método indirecto, ya que según [Skøien y Blöschl \(2006\)](#) las escalas de modelado deberían ir con los objetivos del proyecto de trabajo. Sin embargo, esto no suele ser posible, porque la verdadera estructura de covarianza de la variable de interés es desconocida *a priori*, así como por razones logísticas (conocimiento del relieve, accesibilidad, entre otras).

Lo expuesto está condicionado por el muestreo espacial, ya que la toma de las muestras establecerá parcialmente el comportamiento del semivariograma y, como consecuencia, la autocorrelación espacial y el módulo escalar ([Skøien y Blöschl, 2006](#)). En este orden de ideas [Plant \(2012\)](#) ha indicado que para comparar los métodos de muestreo espacial y la estimación del semivariograma es necesario definir estándares de comparación; una posible manera de realizarlo es considerar alguna medida de intervalos de confianza. Para alimentar estas ideas la eficiencia del muestreo se define como el inverso de la varianza muestral; por ello, el diseño más eficiente conduce a una estimación más precisa y bajo un proceso de autocorrelación espacial la densidad de los puntos de las muestras debería aumentarse en las ubicaciones que expresan una mayor variabilidad espacial (declives de altas pendientes o formas abruptas que generen heterogeneidad espacial). Los valores de las muestras con fuertes similitudes pueden evidenciar redundancia y sobre-muestreo en esas áreas, con ello debería disminuirse la varianza del estimador del Kriging ([Aubry, 2000](#); [Delmelle, 2014](#)).

Los elementos antes descritos parten del principio de que se posee información en ciertas posiciones espaciales, que faciliten los análisis exploratorios y la caracterización de la distribución de frecuencia ([Skøien y Blöschl, 2006](#); [Plant, 2012](#)). Una forma para sortear el problema de la covarianza espacial sin poseer datos sobre el atributo de trabajo es utilizar la aproximación de los patrones puntuales; según [De la Cruz \(2006\)](#), en esta aproximación las medidas de la covarianza no necesariamente involucran un atributo para la estimación de la misma, la idea principal es plantear el tipo de diseño de muestreo que se desea aplicar en el área de trabajo (aleatorio, estratificado, regular, entre otros) y, a partir de esta aproximación, simular medidas de la covarianza para el proceso de muestreo y escala(s) deseada(s).

Un método para la optimización de muestreo para datos topográficos de alta precisión es el presentado por [Mao et al. \(2016\)](#), quienes parten de un muestreo espacial preliminar, mediante métodos de clasificación del error y desde el supuesto de la distribución normal en los mismos, para luego aplicar diferentes filtros gaussianos con el fin de eliminar los errores en cada uno de ellos y posteriormente poder obtener la diferencia entre los datos topográficos originales y los datos topográficos libres de errores. Esto permite configurar una superficie de error en la cual se incluyen los puntos extremos, con ello el proceso es equivalente a seleccionar la menor cantidad de puntos de muestra.

Otros autores como [Barrios et al. \(2005\)](#), [Pérez y Mas \(2009\)](#), [Moreno et al. \(2010\)](#), [Lorenzo et al. \(2012\)](#), [Hernández et al. \(2013\)](#) y [Camargo et al. \(2014\)](#) han abordado el tema de la precisión en modelos digitales de elevación con enfoques determinísticos, estadísticos, fotogramétricos y Lidar, reportando errores entre 7.3 m y 10 m, con errores mínimos de 0.94 m, 0.18 m y 0.62 m; más adelante, este estudio muestra un método más efectivo para el control del error. Estos enfoques poseen diferencias sustanciales con respecto a la metodología aquí planteada, el primero es el conocimiento de la escala, con la cual pudieron establecer una

medida de la precisión mínima del mismo (Congalton y Green, 2008; FAO, 2010). El segundo es la aproximación estocástica sugerida por Bivand *et al.* (2013) que fue aplicada en el presente trabajo, la cual indica que el valor predeterminado del intervalo da como resultado una visión general inicial de la correlación espacial, y con ello la modificación de los parámetros del intervalo resultó en más detalles y como consecuencia en la disminución del error. En tercer término, la medida del error utilizada en el presente trabajo de investigación (LOOV) tiene varias ventajas sobre estimaciones más simples de error predictivo, como es el criterio de información de Akaike (AIC) y métodos máximo verosímiles, pero que se usan menos en la práctica porque implican pasos computacionales adicionales (Vehtari *et al.* 2017); estos pasos fueron sorteados efectivamente en este trabajo.

CONCLUSIONES

Tras el análisis exploratorio y al comparar los distintos modelos de Kriging se hace evidente que la tendencia es significativa y correspondería a una función de al menos primer orden, que fue identificada mediante el análisis exploratorio de datos y ello implica una carga para el usuario. En concordancia, una forma automatizada de análisis no es recomendable debido a la presencia de un componente de tendencia no homogénea que hace indispensable la juiciosa aplicación del análisis exploratorio de datos, así como también la especificación de los parámetros. Por ello, se recalca que esta metodología requiere de una carga interactiva e iterativa por parte del usuario, dado que es quien toma la decisión de las características del modelo a utilizar, el método de estimación de los parámetros del semivariograma y la decisión de si se cumplen los supuestos estadísticos fundamentales de cada uno de los métodos, lo que le permite al usuario la posibilidad de disminuir el error.

Dado que una de las características de este tipo de levantamientos es que se dispone de datos

altamente voluminosos en áreas de pequeña envergadura, es probable que alguna variante del Kriging simple sea el más adecuado; debido posiblemente a que la cantidad de datos sea elevada y con ello la media de los valores muestrales se aproximen sustantivamente a la media poblacional. Con los modelos que involucraban una remoción de tendencia se consiguió que los residuales se ajustaran a una distribución normal y tuvieran menor varianza del error.

Ante la alta carga que corresponde al usuario y la dificultad que implica la implementación de esta metodología en el lenguaje de programación R u otros lenguajes estadísticos, se hace necesario el desarrollo de una herramienta gráfica con documentación teórico-operativa que auxilie al usuario para la toma de decisiones, para así tener la capacidad de discernir sobre las características inherentes de sus datos.

Finalmente, el Kriging disyuntivo se considera menos adecuado para su aplicación con los datos utilizados. Esto es debido a que a pesar de que tiene menor error medio y menor varianza de los errores, presenta un rango demasiado amplio en los mismos. Adicionalmente, en obediencia al principio de parsimonia y debido a la alta dependencia de este método a la especificación por parte del usuario, no se recomienda su uso por parte de profesionales no entrenados en su aplicación.

AGRADECIMIENTOS

Al Laboratorio de Fotogrametría y Sensores Remotos de la Facultad de Ciencias Forestales y Ambientales de la Universidad de Los Andes (Mérida, Venezuela), a través de los cuales se pudo procesar parte de la información. Al programa de Doctorado en Modelamiento Matemático Aplicado de la Universidad Católica del Maule (Talca, Chile), por su vinculación y aportes a este estudio. Esta investigación es resultado de la iniciativa de los investigadores sin financiamiento alguno.

CONFLICTO DE INTERESES

El manuscrito fue preparado y revisado con la participación de los autores, quienes declaran no tener algún conflicto de interés que ponga en riesgo la validez de los resultados aquí presentados.

CONTRIBUCIÓN POR AUTOR

Todos los autores planearon y diseñaron la investigación, coordinaron y llevaron a cabo el trabajo en campo y laboratorio, analizaron los datos y redactaron el manuscrito.

REFERENCIAS

- Armstrong, M.** (1998). *Basic Linear Geostatistics*. Nueva York: Springer-Verlag.
<https://doi.org/10.1007/978-3-642-58727-6>
- Aubry, P.** (2000). *Le Traitement des Variables Régionalisées en Ecologie: Apports de la Géomatique et de la Géostatistique* (tesis doctoral). Université Claude Bernard, Lyon 1, Francia.
<https://tel.archives-ouvertes.fr/tel-00003736/>
- Barrios A., Pacheco C. y López, J.** (2005). Modelo digital de elevación con fines hidrológicos a partir de curvas de nivel en una cuenca de montaña: evaluación de factores de escala. *Revista Forestal Venezolana*, 49(1), 7-15.
- Bivand, R., Pebesma, E. y Gómez- Rubio, V.** (2013). *Applied spatial data analysis with R*. Nueva York: Springer-Verlag.
<https://doi.org/10.1007/978-1-4614-7618-4>
- Camargo, C., Andrades, J. y Vidal, R.** (2014). Evaluación de MDE para la estimación de volúmenes por procesos de erosión aplicando análisis de autocorrelación espacial en una cuenca de los andes venezolanos caso de estudio: microcuenca torrencial La Machirí. *Colombia Forestal*, 17(1), 40-59.
<https://doi.org/10.14483/udistrital.jour.colomb.for.2014.1.a03>
- Chilès, J. P. y Delfiner, P.** (2012). *Geostatistics: modeling spatial uncertainty*. Nueva York, John Wiley & Sons.
<https://onlinelibrary.wiley.com/doi/book/10.1002/9781118136188>
- Cressie, N.** (1992). *Statistics for spatial data*. *Terra Nova*, 4(5), 613-617.
<https://doi.org/10.1111/j.1365-3121.1992.tb00605.x>
- Congalton, R. y Green, K.** (2008). *Assessing the accuracy of remotely sensed data: Principles and Practices*. Boca Ratón: CRC Press Taylor and Francis group.
<https://www.crcpress.com/Assessing-the-Accuracy-of-Remotely-Sensed-Data-Principles-and-Practices/Congalton-Green/p/book/9781498776660#-googlePreviewContainer>
- De La Cruz, M.** (2006). Introducción al análisis de datos mapeados o algunas de las (muchas) cosas que puedo hacer si tengo coordenadas. *Revista Ecosistemas*, 15(3), 19-39.
<https://www.revistaecosistemas.net/index.php/ecosistemas/article/view/159>
- Delmelle, E. M.** (2014). Spatial sampling. En M. M. Fisher y P. Nijkamp (eds.), *Handbook of regional science* (pp. 1385-1399). Berlín: Springer.
https://doi.org/10.1007/978-3-642-23430-9_73
- Díaz, M.** (2002). *Geoadística aplicada*. México D. F.: Instituto de Geofísica, UNAM; Instituto de Geofísica y Astronomía, CITMA.
<http://mmc2.geofisica.unam.mx/cursos/geoest/GeoEstadistica.pdf>
- Diggle, P. J. y Ribeiro, P. J.** (2007). *Model-based geostatistics. Springer series in statistics*. Nueva York: Springer.
<https://doi.org/10.1007/978-0-387-48536-2>
- Emery, X.** (2013). *Geoestadística*. Santiago: Universidad de Chile.
https://www.ucursos.cl/usuario/f697544fb-48474f4153ef2aafbe4c963/mi_blog/r/Geoestadistica.pdf
- FAO** (2010). *Planos y mapas topográficos*.
http://www.fao.org/tempref/FI/CDrom/FAO_Training/FAO_Training/General/x6707s/x6707s09.htm

- Felícísimo, A.** (1994). *Modelos digitales del terreno. Introducción y aplicaciones en las ciencias ambientales*. Oviedo: Pentalfa Ediciones.
<http://www6.uniovi.es/~feli/index2.html>
- Field, A., Miles, J. y Field, Z.** (2012). *Discovering Statistics using R*. Londres: Sage Publications.
https://doi.org/10.1111/insr.12011_21
- Fox, J.** (2005). The R commander: a basic-statistics graphical user interface to R. *J Journal of Statistical Software*, 14, 1-42.
<https://www.jstatsoft.org/article/view/v014i09>
- Gallardo, A.** (2006). Geostadística. *Revista ecosistemas*, 15(3), 48-58.
<https://www.revistaecosistemas.net/index.php/ecosistemas/article/view/161>
- Giraldo, H.** (2002). *Introducción a la geoestadística. Teoría y aplicación*. Bogotá: Universidad Nacional de Colombia.
- Giraldo, H.** (2011). *Estadística espacial, notas de clase*. Bogotá: Universidad Nacional de Colombia.
- Gujarati, D. y Porter, D.** (2010). *Econometría*. México D. F.: Mc Graw Hill.
- Hernández, C., Castillo, W., Becerra, X. y Hernández, S.** (2013). Evaluación y comparación de métodos de interpolación determinísticos y probabilísticos para la generación de modelos digitales de elevación. *Investigaciones Geográficas, Boletín del Instituto de Geografía*, 82, 118-130.
<http://doi.org/10.14350/rig.35906>
- Kerry, R., Oliver, M. A. y Frogbrook, Z. L.** (2010). Sampling in precision agriculture. En M. A. Oliver (ed.), *Geostatistical applications for precision agriculture* (pp. 35-62). Heidelberg: Springer-Verlag.
https://doi.org/10.1007/978-90-481-9133-8_2
- Laboratorio de Fotogrametría y Sensores Remotos** (2007). *Sistemas de información geográfica, ejercicio número 04: escaneo y georeferenciación*. Mérida: Universidad de los Andes, Laboratorio de Fotogrametría y Sensores Remotos.
- Lachenbruch, P. A. y Mickey, M. R.** (1968). Estimation of Error Rates in Discriminant Analysis. *Technometrics*, 10(1), 1-11.
<https://doi.org/10.1080/00401706.1968.10490530>
- Li, J. y Heap, A.** (2008). *A review of spatial interpolation methods for environmental Scientists*. Canberra: Geoscience Australian.
- Lloyd, C.** (2010). *Spatial Data Analysis: An Introduction for GIS Users*. Oxford: Oxford University Press.
<https://global.oup.com/academic/product/spatial-data-analysis-9780199554324?cc=cl&lang=en#>
- Lorenzo, A., Iseburg, M., Arbelo, M. y Alonso-Benito, A.** (2012). Comparación de modelos digitales del terreno obtenidos mediante LIDAR y técnicas fotogramétricas en una zona forestal de la isla de Tenerife.. En *XV Congreso Nacional de Tecnologías de la Información Geográfica* (pp. 2-7). Madrid: AGE-CSIC.
http://tig.age-geografia.es/2012_Madrid/ponencia4/Lorenzo_final_imp.pdf
- Mao, H. X., Shi, W. Z. y Tian, Y.** (2016). A Preliminary Study on Spatial Sampling for Topographic Data. En *In Quality Aspects in Spatial Data Mining* (pp.: 73-86). CRC Press.
<https://www.crcpress.com/Quality-Aspects-in-Spatial-Data-Mining/Stein-Shi-Bijker/p/book/9780367386320>
- Mantovani, E. C. y Magdalena, C.** (2014). *Manual de agricultura de precisión*. Montevideo, Uruguay: IICA; Prociur.
<http://opac.biblioteca.iica.int/cgi-bin/koha/opac-detail.pl?biblionumber=37>
- Matheron, G.** (1970). La teoría de las variables regionalizadas y sus aplicaciones. En, *Los Cuadernos del Centro de Morfología Matemática de Fontainebleau*. Fascículo 5. París: Centro de Geoestadística de la Escuela de Minas de París.
http://cg.ensmp.fr/bibliotheque/public/MATHERON_Ouvrage_00536.pdf
- Montero, J., Fernández, G. y Mateu, J.** (2015). *Spatial and spatio-temporal geostatistical modeling and kriging*. West Sussex: John Wiley & Sons.
<https://www.wiley.com/en-us/Spatial+and+Spatio+Temporal+Geostatistical+Modeling+and+Kriging-p-9781118413180>

- Moreno, A.** (2008). *Sistemas y análisis de información geográfica. Manual de autoaprendizaje con Arc GIS*. México D. F.: Alfaomega Grupo Editor.
<https://dialnet.unirioja.es/servlet/libro?codigo=504377>
- Moreno, J., Alonso, F., Gomariz, F. y Alonso, D.** (2010). Análisis y validación de modelos digitales de elevaciones mediante datos LiDAR. En *XIV Congreso Nacional de Tecnologías de la información Geográfica* (pp. 254-271).
<https://dialnet.unirioja.es/servlet/libro?codigo=435989>
- Olaya, V.** (2011). *Sistemas de información geográfica*.
https://wiki.osgeo.org/wiki/Libro_SIG.
- Oliver, M. A. y Webster, R.** (2015). *Basic steps in geostatistics: the variogram and kriging*. Nueva York: Springer.
<https://www.springer.com/gp/book/9783319158648>
- Ortiz, J., Oz, B. y Deutsch, C.** (2005). A step by step guide to bi-Gaussian disjunctive kriging. En O. Leuangthong y C. Deutsch (eds.), *Geostatistics Banff 2004 Quantitative Geology and Geostatistics*, vol. 14 (pp. 1097-1102). Dordrecht: Springer.
http://doi.org/10.1007/978-1-4020-3610-1_114
- Pérez, A. y Mas, J.** (2009). Evaluación de los errores de modelos digitales de elevación obtenidos por cuatro métodos de interpolación. *Investigaciones Geográficas*, 69, 53-67.
<http://www.investigacionesgeograficas.unam.mx/index.php/rig/article/view/18004>
- Plant, R. E.** (2012). *Spatial data analysis in ecology and agriculture using R*. CRC Press.
<https://www.crcpress.com/Spatial-Data-Analysis-in-Ecology-and-Agriculture-Using-R/Plant/p/book/9780815392750>
- Royle, A. G.** (1980). *Geostatistics*. Nueva York: McGraw-Hill.
https://www.abebooks.com/servlet/BookDetails-PL?bi=30256852226&tab=1&searchurl=an%3Dh.i.r.m%2Bj.m%26sortby%3D17&cm_sp=snippet_-_srp3_-_seller17
- R Core Team** (2019). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
<http://www.R-project.org>.
- Skøien, J. O. y Blöschl, G.** (2006). Scale effects in estimating the variogram and implications for soil hydrology. *Vadose Zone Journal*, 5(1), 153-167.
<https://doi.org/10.2136/vzj2005.0069>
- Tobler, W.** (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, suppl., 46, 234-240.
<https://www.tandfonline.com/doi/abs/10.2307/143141>
- Vehtari, A., Gelman, A. y Gabry, J.** (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing*, 27(5), 1413-1432.
<https://doi.org/10.1007/s11222-016-9696-4>
- Wackernagel, H.** (1998). *Multivariate Geostatistics*. Berlín: Springer-Verlag.
<https://www.springer.com/gp/book/9783540441427>
- Wang, F. (ed.)** (2005). *Geographic information systems and crime analysis*. Idea Group Inc: Hershey, PA, EE. UU.
<https://www.igi-global.com/book/geographic-information-systems-crime-analysis/411>
- Willmott, C. y Matsuura, K.** (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79-82.
www.jstor.org/stable/24869236

