

# DETERMINACIÓN DE LA DESVIACIÓN ALTIMÉTRICA DE MODELOS DIGITALES DE ELEVACIÓN MEDIANTE MÉTODOS GEOESTADÍSTICOS FASE I

## DETERMINING OF ALTIMETRIC ERROR DEVIATION BY GEOSTATISTICAL METHODS PHASE I

Julián Garzón Barrero\*, Gonzalo Jiménez Cleves, José Joaquín Vila Ortega

Universidad del Quindío, Facultad de Ingeniería, Programa de Topografía  
Grupo de investigación GEOIDE G-62. juliangarzonb@uniquindio.edu.co,  
gjclevess@uniquindio.edu.co, jjvilaortega@uniquindio.edu.co.

Recibido: Octubre 2 de 2013

Aceptado: Noviembre 1 de 2013

\*Correspondencia del autor. CEIFI, Grupo GEOIDE G62 – Universidad del Quindío,  
Av. Bolívar calle 12 norte. Armenia Quindío, Colombia.  
E-mail: juliangarzonb@uniquindio.edu.co

### RESUMEN

Los Modelos Digitales de Elevación se definen como estructuras de datos numéricas que representan la distribución espacial de la elevación sobre la superficie terrestre. En términos de representación, la solución aparentemente más adecuada hasta el presente ha sido el tratamiento del relieve mediante una estructura vectorial espacial: la Red de Triángulos Irregulares o TIN siglas de Triangular Irregular Network propuesta por Poiker en el año 1968, pero por su naturaleza, este algoritmo obedece a una función lineal, mientras que la geometría de la naturaleza no lo es, mostrando un comportamiento caótico; es aquí donde podría lograrse una mejor representación de dichas formas con la implementación del algoritmo de Kriging permitiendo analizar la variación estadística sobre diferentes distancias y direcciones, obteniendo el mínimo error en la elevación estimada; este trabajo busca establecer la relación entre la geometría de la superficie y el algoritmo de interpolación, a través de un análisis de varianzas (ANOVA) de los Errores Medios Cuadráticos (EMC) de una muestra de 80 diferentes modelos clasificados de acuerdo a su rango de pendiente.

**Palabras clave:** ANOVA, error medio cuadrático, geoestadística, Kriging, TIN

### ABSTRACT

Digital Elevation Models (DEM) are defined as numerical data structures that represent the spatial distribution of elevation above the surface. From the 70's was established the representation of surfaces generated by the TIN model as the most appropriate, but by its nature, this algorithm follows a linear function, while the geometry of nature is showing a chaotic behavior, for that reason could achieve a better representation of these forms with the implementation of Kriging algorithm allowing to analyze the statistical variation over different distances and directions, causing the minimum error in estimated elevation; this work seeks to establish the relationship between surface geometry and the interpolation algorithm, through an analysis of variance (ANOVA) of Root Mean Square Error (RMSE) of a sample of 80 different models ranked according to their slope.

**Key words:** ANOVA, Root Mean Square Error, Geostatistics, Kriging, TIN

## INTRODUCCIÓN

La modelación de superficies es un término generalmente usado para describir el proceso de representar un espacio físico natural o artificialmente creado mediante expresiones matemáticas (1). La superficie terrestre es susceptible de ser modelada mediante un Modelo Digital de Elevaciones (DEM) que es una estructura numérica de datos que representan la distribución espacial de la altitud de la superficie del terreno (2). Debido a que la superficie de la tierra posee un número infinito de puntos, se hace un muestreo de aquellos que producen los rasgos más significativos de la geometría de la superficie, estos se procesan mediante algún sistema computarizado, y a partir de la utilización de algún algoritmo de interpolación se puede pasar de un variable matemática de tipo simple (elevación de puntos), a una de tipo continuo conocida como superficie. Con la creciente disponibilidad de los DEM y su uso en una amplia gama de aplicaciones, el problema de estudiar la manera de medir su calidad es un asunto de creciente interés (3). Actualmente los DEM son la forma más efectiva de representación de la superficie terrestre y se han convertido en una herramienta y producto muy demandado para su uso en los ámbitos de la geoinformática como la planificación del manejo de recursos naturales (4); en modelación hidrológica, la ruta del agua superficial está condicionada por la estructura superficial derivada de un DEM (5). La fotogrametría y teledetección usan los DEM como elemento primario en la corrección geométrica de imágenes satelitales en áreas de montaña (6). Debido a que se estableció por Thomas K. Poiker, desde 1968 que la representación que genera el modelo TIN (Triangular Irregular Network) es la más adecuada para representar formas topográficas en un Sistema de Información Geográfica (7), y entendiendo que este algoritmo está basado en la ecuación de un plano que pasa a través de tres puntos no colineales en el espacio, se logra así determinar la elevación de un punto cualquiera que quede dentro de este triángulo: pero esto obedece a un sistema lineal de ecuaciones, mientras que las formas geométricas de una superficie natural no tiene comportamientos lineales sino más bien caóticos, es aquí donde “posiblemente” pueden tener mejores representaciones de las formas otros algoritmos como Kriging.

Tradicionalmente los DEM no han recibido en el ámbito geográfico una atención especial en los procesos de evaluación de la calidad altimétrica, eligiendo el algoritmo de interpolación más por una apariencia visual que por conocimiento de las limitaciones que ellos pre-

sentan, causando grandes desviaciones altimétricas que ocasionan análisis espaciales erróneos. De acuerdo con Morillo Barragán et al, (8) la precisión del modelo generado se puede considerar dependiente de una serie de parámetros implicados en su generación, que van desde las características orográficas de la zona, los datos fuente hasta el método de elaboración. Ley (9) fue uno de los primeros en sugerir que las medidas de error de los DEM deben basarse en una consideración de más que simplemente elevación, y sugirió la exactitud de los valores en función de la pendiente como una alternativa, ya que reúne errores tanto horizontales como verticales. En este trabajo se hace una revisión de los dos algoritmos en estudio: TIN y Kriging, interpolando los datos originales de 80 mapas topográficos levantados por métodos convencionales (estación total), los mapas han sido clasificados de acuerdo a los rangos de pendiente propuestos por Ley (9): plano, ondulado, escarpado y montañoso; cada uno de las cuales será interpolado por los dos algoritmos en mención. En el ámbito de los DEM el error se conoce como la desviación altimétrica del modelo con respecto a la superficie real (10); hay que tener en cuenta que la información siempre será imprecisa debido a la simplificación que está inmersa dentro de la modelación, pues la superficie real siempre tendrá un infinito número de puntos para modelarse con todas sus características morfométricas, es así como se hace un muestreo de los puntos más significativos que representen la superficie con la mayor cantidad de características posibles que hablen de su forma. Muchos trabajos sobre los errores en la creación de los DEM muestran tendencias a centrarse en la precisión métrica de los valores de elevación. Una práctica normal en este entorno es comparar una muestra de elevaciones desde el DEM con una muestra de elevaciones medidos a un mayor nivel de precisión (11). Mediante la metodología propuesta se busca evaluar la calidad de cada modelo a partir del cálculo del Error Medio Cuadrático, en función del número de puntos de control que dependen del modelo a evaluar, posteriormente se calculará el error cualitativa y cuantitativamente mediante de un Análisis de Varianza (ANOVA) que determine la calidad de un conjunto de medidas en términos de las medias de sus varianzas, definiendo la significancia o no de la variable dependiente en función de una posible afectación causada por una variable de tipo independiente. La precisión de estos modelos se calculará mediante la comparación del mismo conjunto de datos interpolados con dos algoritmos diferentes TIN y Kriging, tomando como libre de error (en términos de representación) TIN, pues presenta una fidelidad superior.

### Planteamiento del problema

La mayoría de especialistas que calculan volúmenes de movimientos de tierra, utilizan los DEM para tal fin, haciéndolo de una manera indiscriminada y sin el conocimiento necesario para la creación de dicho modelo, pues no se tiene conciencia de la utilización de elementos fundamentales como líneas de rotura, zonas de recorte, puntos acotados, y menos aún de los algoritmos de interpolación haciendo una elección de este último, más por razones de presentación en términos de su apariencia visual, que por razones que obedezcan a la relación existente entre la geometría de la forma que se busca representar y su algoritmo de interpolación. Para ocultar un poco esta falencia, en el mercado laboral se tiene el mito que los cálculos de volúmenes son aproximados, motivo por el cual se llega a perder bastante dinero a la hora de comparar el volumen cubicado con el volumen extraído. El motivo principal entre estas diferencias se encuentra radicado en la evaluación subjetiva que tradicionalmente se hace a un DEM, estos modelos son construidos con datos espaciales, que por su naturaleza poseen errores intrínsecos a las mediciones y su propagación en el modelo, tales errores deben ser tratados con un procedimiento geoestadístico que permita cuantificar su error y buscar una respuesta a la incertidumbre que en nuestro medio se presenta ante la tradicional evaluación subjetiva de los DEM.

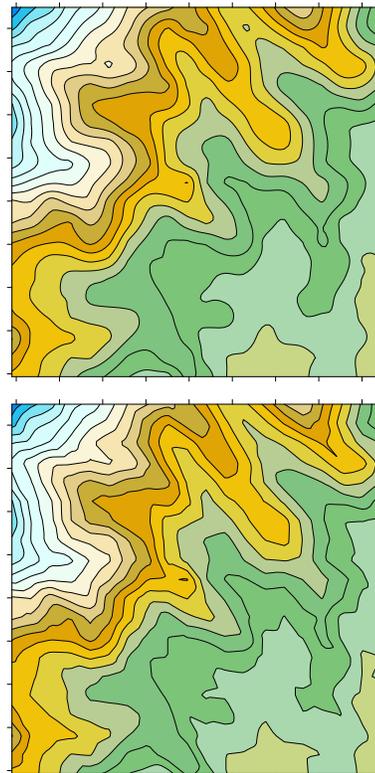
Los parámetros extraídos de los Modelos Digitales de Elevación (DEM) permiten caracterizar el relieve de forma cuantitativa, así como establecer las relaciones de estos con los procesos y características del suelo, parte fundamental en la evaluación de aptitud de las tierras (12). Pequeños errores en elevación pueden producir grandes errores en los valores derivados de la superficie, especialmente en los derivados de segundo orden como la curvatura (13,14).

Imhof (15) expone la dificultad de calcular la precisión de la cartografía topográfica: “para calcular el error cartográfico se pretende comparar dos superficies diferentes y complejas, que no pueden ser expresadas mediante fórmulas matemáticas. Una de ellas es la actual superficie de la tierra, reducida mentalmente a la escala del mapa, la otra superficie es la formada por isolíneas que conforman el mapa bajo estudio. Como resultado de las deficiencias de la cartografía, ambas superficies no coinciden perfectamente, y se desea ser capaz de juzgar la naturaleza y extensión de tales variaciones. La primera superficie (la real) nunca es exactamente conocida, ya que cualquier medición viene acompañada de errores. La segunda superficie, aparte de los errores cartográ-

ficos inherentes nunca está totalmente definida, ya que las isolíneas, debido a la equidistancia, dejan grandes espacios a los lados que no son cubiertos. Entre cada dos isolíneas la posición de la superficie es desconocida, aunque se presume que las une una línea recta, lo que es menos probable a medida que aumenta la equidistancia. Desde este punto de vista: ¿Cómo podemos comparar una superficie incompletamente definida (la real) con otra que no puede ser exactamente determinada (el mapa)? ”

### MATERIALES Y METODOS

La utilización de técnicas diferentes sobre el mismo conjunto de datos permite determinar si existe mayor o menor grado de concordancia entre datos e hipótesis y también cómo esa concordancia varía irregularmente con la distribución espacial de la información.



**Figura 1.** Arriba: interpolación Kriging, Abajo: Interpolación TIN.

Las figura 1 corresponden a dos modelos de interpolación diferentes de un mismo conjunto de datos muestrales, dejando ver que se pueden generar errores como consecuencia de la selección del algoritmo.

El objeto de la metodología propuesta es la implementación de un modelo experimental que permita determinar bajo qué rangos de pendiente (tabla 1) es posible

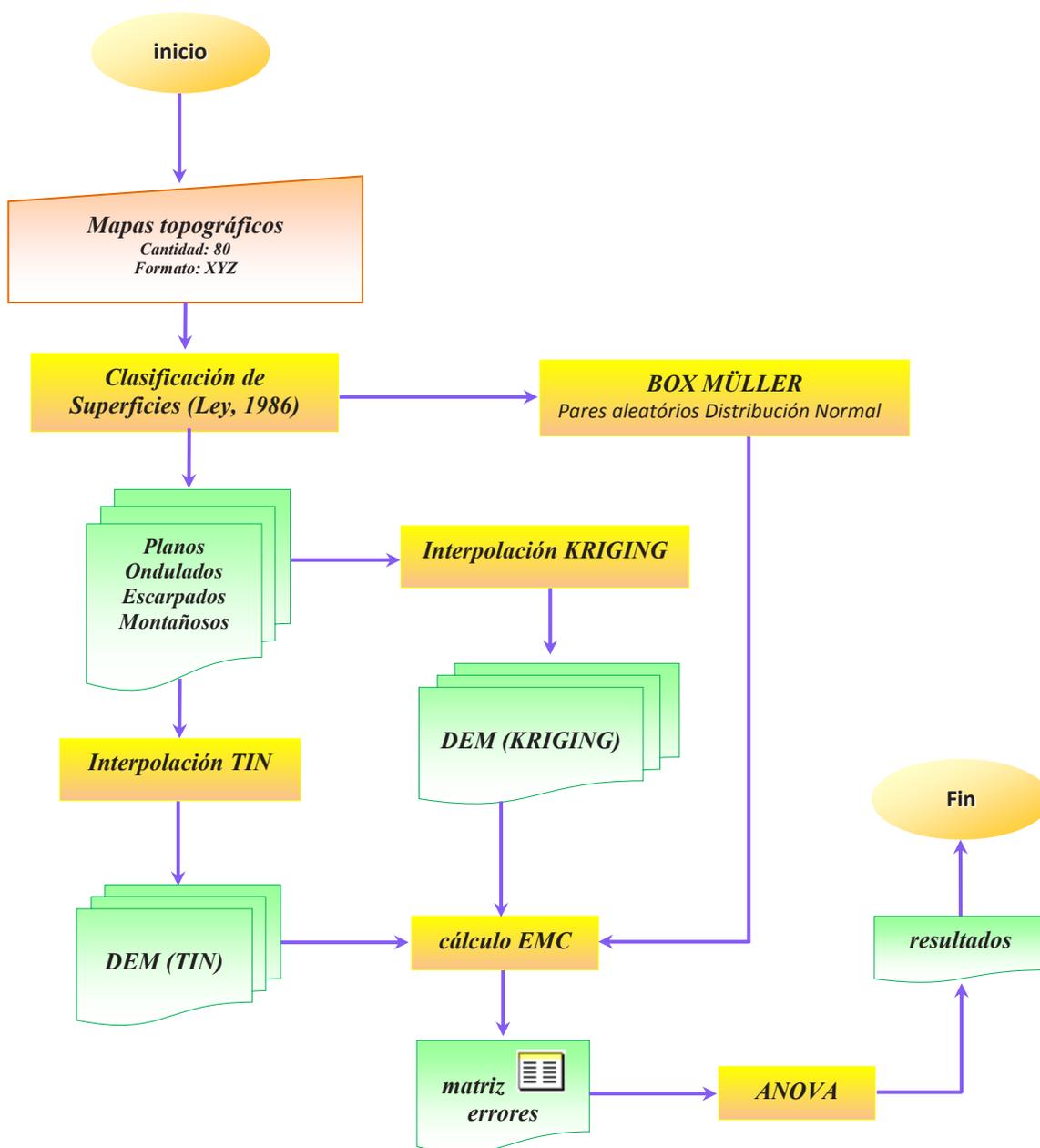


Figura 2. Diagrama de flujo de la metodología.

utilizar el algoritmo de interpolación de Kriging, sin comprometer la calidad altimétrica de la representación de la superficie en los Modelos Digitales de Elevación.

Por disponibilidad se supone que los valores de pendiente deben estar disponibles o ser estimados antes de que se haga el muestreo, a fin de ayudar en la determinación de sus intervalos. Si existe un DEM de la zona de trabajo, entonces los valores de pendiente para los puntos que lo conforman se pueden calcular y el promedio se puede usar como representativo (18). Por el contrario, la pendiente podría estimarse desde un modelo estéreo,

o por un par estereoscópico de fotografías aéreas con traslapeo, o desde mapas con curvas de nivel. El método propuesto por Wentworth (19) se discute aún ampliamente para estimar el promedio de pendiente de un área desde mapas con curvas de nivel. Por variabilidad se entiende que los valores de pendiente podrían alterarse de un lugar a otro, así que la estimación de pendiente que es representativa para un área podría no adaptarse a otra. En este caso, los valores promedios podrían usarse tal como lo sugirió Ley (9). Si la pendiente varía notablemente en un área, entonces el área deberá dividirse en partes más pequeñas para su estimación.

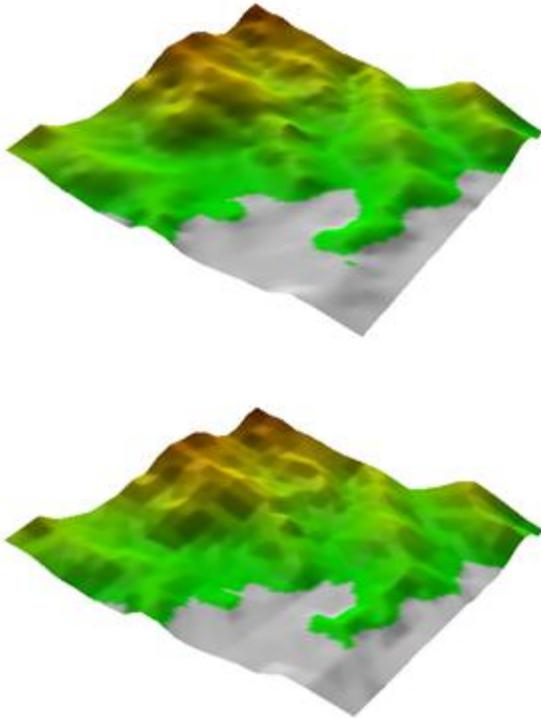
**Tabla 1.** Rangos de Pendiente (Ley, 1986).

Tipo de terreno	Pendiente (°)	Pendiente (%)	Relieve (rango de alturas)
Plano	< 2	<4,4	< 80
Ondulado	2-6	4,4 – 13,3	80-300
Escarpado	6-25	13,3 – 55,6	300-600
Montañoso	> 25	>55,6	> 600

Fuente: Ley, R, 1986, Accuracy assessment of digital terrain models

### Algoritmos de interpolación evaluados: TIN vs Kriging

En las ciencias de la tierra es muy usual encontrar variables que se manifiestan distribuidas espacialmente. Para el estudio de estas variables se utilizan diferentes procedimientos de estimación. Esto quiere decir que, a partir de un conjunto de muestras tomadas en localizaciones del dominio en que se manifiesta un fenómeno a estudiar y consideradas representativas de su realidad, que por lo general es siempre desconocida, éstos procedimientos permiten su caracterización con dos fines diferentes: el primero proporcionar valores estimados en localizaciones de interés y el segundo interpolar valores que en conjunto presenten iguales características de dispersión que los datos originales.

**Figura 3.** Resultados DEM: Kriging (arriba) TIN(abajo).

En términos de Geoestadística es necesario que estos valores se auto-correlacionen, y el estudio de esta correlación se llama análisis estructural y es representado gráficamente mediante el variograma, luego de este análisis, se hacen estimaciones a los sectores no muestreados usando el algoritmo de Kriging. De acuerdo con este modelo metodológico se deben interpolar las 80 muestras de superficies topográficas por los algoritmos anteriormente expuestos: TIN y Kriging, que presentan diferencias en los DEM como se aprecia en la figuras 3.

### Puntos de control

Para determinar el Error Medio Cuadrático se hace necesario tener un conjunto limitado de puntos de control, y las diferencias de altura entre el modelo que se toma como *Valor Más Probable* (libre de error) que para este caso es la superficie generada con el algoritmo de TIN, y el modelo que se está evaluando: Kriging, se puede determinar el EMC para cada Modelo Digital de Elevación. De acuerdo con Särndal et al, (20) el tamaño que debe poseer una muestra  $n$  para obtener una determinada variación de la media  $S$ , está dado por la siguiente ecuación de muestreo aleatorio simple:

$$n = \frac{k^2 \cdot p \cdot q \cdot N}{(e^2 \cdot (N - 1)) + k^2 \cdot p \cdot q} \quad [1]$$

Donde  $n$  corresponde al número de puntos de control,  $N$  es el tamaño de la población,  $k$  es una constante que depende del nivel de confianza,  $e$  es el error de la muestra,  $p$  es la proporción de individuos que poseen en la población la característica de estudio, y  $q$  es la proporción de individuos que no poseen esa característica, es decir, es  $1-p$ .

### Variabes aleatorias Normales

El método de Box-Müller (nombrado así por sus inventores George Edward Pelham Box y Mervin Edgar Müller, 1958) es un método de generación de pares de números aleatorios independientes con distribución normal “estándar” (esperanza cero y varianza unitaria), a partir de una fuente de números aleatorios uniformemente distribuidos.

$$Z_0 = R \cos \Theta = \sqrt{-2 \ln U_1 \cos(2\pi U_2)} \quad [2]$$

$$Z_1 = R \sin \Theta = \sqrt{-2 \ln U_1 \sin(2\pi U_2)} \quad [3]$$

Donde  $Z_0$  y  $Z_1$  corresponden a variables aleatorias independientes con una distribución normal con desviación standard 1,  $U_1$  y  $U_2$  son variables aleatorias independientes que están uniformemente distribuidas en el intervalo (0, 1],  $R$  y  $\Theta$  corresponden a las coordenadas polares del vector.

### Cálculo de errores medios cuadráticos

De acuerdo a la teoría de Errores desarrollada por Gauss el Error Medio Cuadrático es el modelo matemático más exigente para calcular el error de un DEM, por lo que su uso facilita la comparación de los valores obtenidos; lo que se pretende es establecer y cuantificar las diferencias más significativas encontradas entre la aplicación de estos dos algoritmos usados para la construcción de superficies.

$$EMC = \pm \sqrt{\frac{\sum_{i=1}^n \vartheta_t^2}{n-1}} \quad [4]$$

El procedimiento a seguir desde el punto de vista teórico es el siguiente:

- En primer lugar hay que calcular, para cada punto de control, la diferencia altimétrica existente entre el modelo interpolado por TIN y el interpolado por Kriging.
- Luego se elevan dichas diferencias al cuadrado  $\vartheta_t^2$  para eliminar el valor del signo.
- Después se obtiene el promedio de las diferencias para todo el conjunto de puntos de control

$$\frac{\sum_{i=1}^n \vartheta_t^2}{n-1}$$

- Por último se halla la raíz cuadrada del promedio obteniendo el EMC

El resultado final es la cobertura puntual que contiene los registros para caracterizar los modelos de cada área de estudio, en función de su EMC para las distintas pendientes en las cuales se han caracterizado los terrenos.

### Análisis de varianza (ANOVA)

La validez de este análisis está supedita al cumplimiento de algunos requisitos de tipo matemático y probabilístico en las observaciones. El modelo ANOVA es apto para comparar efectos de variables en un fenómeno si los datos experimentales se ajustan al modelo, es decir,

si las condiciones subyacentes al modelo están implícitas en los datos. Para este caso específico se realizó un ANOVA de un factor, lo que quiere decir que se permite comparar varios grupos de una variable cuantitativa, se definió como variable nominal la pendiente del terreno categorizada de acuerdo a la tabla 1, y la variable cuantitativa, es decir en la cual se desean comparar los grupos es el Error Medio Cuadrático de cada una de las 80 superficies, resumiendo se tiene:

#### Análisis No 1

Variable independiente: Tipo de superficie (plana, ondulada, escarpada, montañosa)

Variable dependiente: Error Medio Cuadrático

#### Análisis No 2

Variable independiente: Tipo de superficie (plana, ondulada, escarpada, montañosa)

Variable dependiente: Porcentaje de error en el cálculo del volumen.

La hipótesis que se pone a prueba en el ANOVA de un factor es que las medias poblacionales (las medias del EMC en cada nivel del tipo de superficie) son iguales. Si las medias poblacionales son iguales, esto significa que los grupos no difieren en el Error Medio Cuadrático, y que en consecuencia, el tipo de superficie es independiente del Error Medio Cuadrático.

$$h_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_a \quad [6]$$

$$h_1: \text{No es cierto } h_0 \quad [7]$$

Donde  $h_0$  corresponde a la hipótesis nula;  $\mu_1, \mu_2, \mu_3$  y  $\mu_a$  representan las medias correspondientes a cada nivel del factor; y  $h_1$  corresponde a la hipótesis alternativa.

La estrategia para poner a prueba la hipótesis nula  $h_0$  de igualdad de las medias consiste en obtener el valor F (de la distribución de Fisher) que refleja el grado de parecido existente entre las medias que se están comparando. Si las medias de la muestra son parecidas el coeficiente F tomará un valor cercano a 1. Si las medias muestrales son distintas dicho coeficiente tomará un valor mayor que 1, y cuanto más diferentes sean las medias, mayor será el valor de F. Si el nivel crítico asociado al valor F es menor que 0,05 se rechaza la hipótesis de igualdad de medias y se concluye que no todas las medias poblacionales son iguales. En el caso contrario no se puede rechazar la hipótesis de igualdad y no se puede afirmar que los grupos que se están comparando difieran en sus promedios poblacionales.

**Tabla 2.** Matriz de errores para ANOVA superficies planas, onduladas, escarpadas y montañosas

Tipo	muestra	datos activos	datos a evaluar	EMC	desnivel	coeficiente de variación (z)	% Variación Volumen
pl	1	62	54	0.013	0.54	0.51	0.15
	2	54	48	0.019	0.50	0.02	0.04
	3	49	44	0.014	0.51	0.69	1.10
	4	42	38	0.012	0.48	0.46	1.74
	5	47	42	0.019	0.78	0.72	0.58
	6	88	72	0.023	0.75	0.66	2.42
	7	67	57	0.041	1.49	0.33	0.26
	8	91	74	0.024	0.69	0.39	1.69
	9	88	72	0.017	0.60	0.86	1.42
	10	89	73	0.026	1.70	0.48	0.86
	11	97	78	0.012	0.82	0.39	0.52
	12	73	62	0.011	0.50	0.43	0.08
	13	75	63	0.019	1.02	0.28	0.51
	14	91	74	0.022	0.60	0.75	0.94
	15	82	68	0.016	0.99	0.46	0.00
	16	72	61	0.029	1.95	0.45	0.86
	17	72	61	0.024	0.75	0.48	0.24
	18	65	56	0.019	1.19	0.42	0.85
	19	70	60	0.021	0.40	0.57	4.37
	20	87	72	0.020	1.25	0.52	0.43
on	21	146	106	0.150	12.07	0.45	0.89
	22	93	76	0.205	3.96	0.41	2.46
	23	112	87	0.096	10.98	0.43	0.23
	24	102	81	0.079	8.76	0.50	0.19
	25	88	71	0.123	6.32	0.65	0.38
	26	70	60	0.168	9.58	0.45	0.21
	27	71	61	0.164	10.00	0.49	0.30
	28	79	66	0.245	10.51	0.32	0.64
	29	69	59	0.143	7.12	0.59	0.26
	30	67	58	0.119	7.12	0.59	0.13
	31	63	55	0.152	5.78	0.63	0.66
	32	92	75	0.124	5.49	0.77	1.05
	33	70	60	0.184	7.79	0.50	0.43
	34	76	64	0.184	13.43	0.70	1.13
	35	83	69	0.076	7.65	0.47	0.03
	36	112	87	0.141	6.54	0.97	0.50
	37	93	76	0.074	8.89	0.62	0.29
	38	79	66	0.087	9.05	0.78	0.18
	39	87	71	0.091	6.45	0.44	0.16
	40	87	71	0.149	7.78	0.43	0.48
esc	41	253	152	0.394	8.57	0.44	2.02
	42	88	72	0.668	18.00	0.76	3.06
	43	63	55	0.403	30.00	0.72	0.01
	44	65	56	0.272	16.00	0.40	0.33
	45	68	48	0.218	9.56	0.55	1.07
	46	68	48	0.158	14.97	0.64	0.12
	47	72	51	0.126	14.97	0.64	0.07
	48	70	49	0.271	15.01	0.48	0.19
	49	72	51	0.179	20.92	0.70	0.27
	50	86	60	0.307	29.29	0.69	0.14
	51	96	67	0.493	18.07	0.65	1.23
	52	157	110	0.678	35.39	0.64	0.14
	53	132	92	0.423	32.46	0.41	0.06
	54	84	59	1.083	23.64	0.31	0.75
	55	145	102	1.375	42.00	0.59	1.05
	56	132	92	0.625	13.04	0.73	0.63
57	112	79	0.527	13.25	0.75	0.31	
58	148	104	0.575	45.21	0.60	0.22	
59	102	72	0.648	37.42	0.67	0.47	
60	107	75	1.247	31.74	0.63	1.24	
mont	61	113	79	1.636	65.02	0.55	0.06
	62	116	81	1.119	36.00	0.74	0.21
	63	123	86	1.432	39.90	0.57	0.88

**Tabla 2.** Matriz de errores para ANOVA superficies planas, onduladas, escarpadas y montañosas (Continuación)

Tipo	muestra	datos activos	datos a evaluar	EMC	desnivel	coeficiente de variación (z)	% Variación Volumen
	64	123	86	1.621	65.65	0.52	0.04
	65	109	76	2.824	55.05	0.57	0.19
	66	108	76	1.946	43.96	0.51	1.10
	67	71	50	2.591	90.89	0.56	1.03
	68	131	92	2.631	38.48	0.55	2.46
	69	93	65	3.701	202.90	0.62	0.91
	70	136	95	0.593	37.57	0.60	0.97
	71	152	106	1.207	23.99	0.37	0.08
	72	148	104	1.034	47.00	0.45	0.30
	73	118	82	2.201	28.00	0.81	0.04
	74	113	79	0.990	57.00	0.40	0.65
	75	157	110	1.703	47.00	0.53	0.17
	76	116	81	0.627	40.00	0.69	1.64
	77	171	120	1.520	66.00	0.37	0.16
	78	129	90	1.803	44.00	0.55	2.30
	79	113	79	1.470	53.29	0.53	1.69
	80	168	118	1.259	49.44	0.47	1.24

pl: superficie plana, on: superficie ondulada, esc: superficie escarpada, mont: superficie montañosas.

**RESULTADOS Y DISCUSION**

Se obtuvieron cuatro matrices de error, una por cada tipo de superficie (plana, ondulada, escarpada y montañosas). Esta matriz concreta la información básica que se usó en el análisis estadístico de varianzas, en ella las columnas se ocupan para las clases de referencia (ver tabla 2) y las filas para el número de identificación de la muestra procesada digitalmente.

Analizando la Tabla 3, La variación entre grupos es de 63,895, para saber si los resultados de este análisis son significativos (es decir que la probabilidad tiene un valor menor a 0,05 el valor de F necesita ser al menos 2,72 (o sea el valor crítico para F). Entonces, como el valor de F (63,895) es mucho mayor que el valor crítico para F (2,720) con seguridad los resultados de la prueba son significativos, dicho de otra forma: sí existe una relación significativa en cuanto que cada tipo de terreno posee relación directa con su Error Medio Cuadrático, se han graficado la dispersión de puntos por categoría (Figura 4) donde se corrobora de forma gráfica la significancia entre el conjunto de datos.

**Análisis 1**

Variable independiente: Tipo de superficie (plana, ondulada, escarpada, montañosas)

Variable dependiente: Error Medio Cuadrático

Como se obtuvo un valor significativo, se acepta que existen diferencias significativas entre las varianzas, y surgen las siguientes preguntas: ¿cuáles son dichas diferencias?, ¿Son todos los grupos significativamente diferentes entre sí o sólo algunos de ellos respecto de los demás?

Para responder a esos cuestionamientos se hace necesario realizar un Test de Rango Múltiples

**Pruebas de Múltiple Rangos para EMC por categoría**

Método: 95,0 porcentaje de la Diferencia Mínima Significativa LSD (en inglés, *Least Significant Difference*)

En la tabla 4 se aplica un procedimiento de comparación múltiple para determinar cuáles medias son significativamente diferentes de otras. El asterisco que se encuentra al lado de los cinco pares indica que estos

**Tabla 3.** Resultados ANOVA

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	35,086	3	11,695	63,895	9,939E-21	2,724
Dentro de los grupos	13,911	76	0,183			
Total	48,998	79				

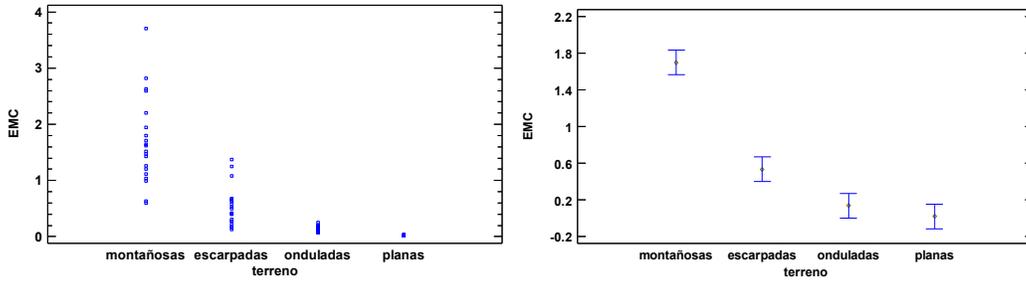


Figura 4. Izquierda: dispersión por tipo de terreno, Derecha: media de Fisher 95% LSD

Tabla 4. Resultados ANOVA

Categoría	Casos	Media	Grupos Homogéneos
plana	20	0,020	X
ondulada	20	0,137	X
escarpada	20	0,533	X
montañosa	20	1,695	X

Tabla 5. Test de rangos múltiples

Contraste	Sig.	Diferencia	+/- Límites
escarpada - montañosa	*	-1,161	0,2694
escarpada - ondulada	*	0,395	0,2694
escarpada - plana	*	0,5134	0,2694
montañosa - ondulada	*	1,557	0,2694
montañosa - plana	*	1,6753	0,2694
ondulada - plana		0,1176	0,2694

pares muestran diferencias estadísticamente significativas con un nivel del 95,0% de confianza. No existen diferencias estadísticamente significativas entre aquellos niveles que compartan una misma columna de X's. El método empleado actualmente para discriminar entre las medias es el procedimiento de diferencia mínima significativa (LSD) de Fisher. Con este método hay un riesgo del 5,0% al decir que cada par de medias es significativamente diferente, cuando la diferencia real es igual a cero.

Fuente: Los autores

\* indica una diferencia significativa.

**Análisis 2**

Variable independiente: Tipo de superficie (plana, ondulada, escarpada, montañosa)

Variable dependiente: Porcentaje de error en el cálculo del volumen

La tabla 6 (resultados de ANOVA) indica que se descompone la varianza de del volumen en dos componentes: un componente entre-grupos y un componente dentro-de-grupos. La razón-F, que en este caso es igual a 1,042, es el cociente entre el estimado entre-grupos y el estimado dentro-de-grupos. Puesto que el valor-P de la razón-F es mayor o igual que 0,05, no existe una diferencia estadísticamente significativa para la media de 'variación de volumen' entre un nivel de categoría y otro, con un nivel del 95,0% de confianza.

Tabla 6. Resultados ANOVA

Grupos	Cuenta	Suma	Promedio	Varianza
planas	20	19,081	0,954	1,064
onduladas	20	10,606	0,530	0,299
escarpadas	20	13,374	0,668	0,599
montañas	20	16,110	0,805	0,572

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de los cuadrados	F	Probabilidad	Valor crítico para F
Entre grupos	1,983	3	0,661	1,042	0,378	2,724
Dentro de los grupos	48,180	76	0,633			
Total	50,163	79				

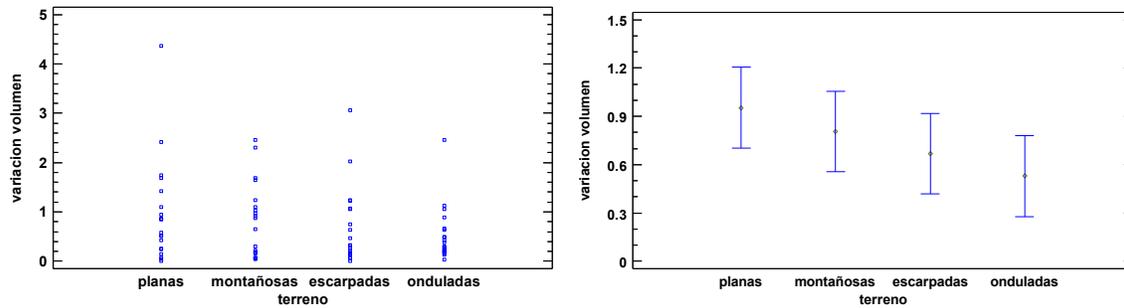


Figura 5. Izquierda: dispersión por tipo de terreno, Derecha: media de Fisher 95% LSD

La figura 5 muestra que la dispersión de puntos por categorías al tomar como variable dependiente el volumen es similar en los 4 tipos de superficies (escarpada, montañosa, ondulada y plana), puesto que las proyecciones de los puntos medios alcanzan a reflejarse en sus categorías vecinas; ratificando que no hay una diferencia significativa para la variable volumen.

## CONCLUSIONES

La elección del algoritmo de interpolación en la construcción de un DEM debe estar fuertemente vinculada con el empleo final al cual éste vaya destinado, así mismo esta selección exige el conocimiento del tipo de estructura que almacena la información altimétrica, la densidad de los puntos (en caso de las estructuras vectoriales), la tolerancia en la magnitud del error final de representación, y finalmente la forma de medir este error. De acuerdo con los resultados obtenidos en esta investigación es necesario medir la magnitud de los errores producidos por los algoritmos interpolantes sobre las distintas geometrías que se presentan en la naturaleza de la superficie terrestre, esto con el fin de poder tener una luz en la propagación de su comportamiento, ya que se comprueba que existe un alto índice autocorrelación espacial del error, así como vínculos muy fuertes entre éste (el error) y la pendiente a través del procedimiento de modelación.

En la representación de los DEM sus errores son inherentes a la información espacial. Generalmente los usuarios no tienen información sobre la naturaleza y el alcance de error en sus datos. En esta investigación se ha abordado la temática de los errores aleatorios en los DEM; se evaluaron dos métodos de interpolación para la representación de error aleatorio, de acuerdo al modelo estadístico usado (ANOVA), no existen diferencias significativas entre las varianzas de las superficies categorizadas como planas y onduladas, por lo cual para

este par de clases se cumple la hipótesis nula:  $h_0: \mu_1 = \mu_2$ . En función del resultado de aceptación de la nulidad de esta hipótesis, la ausencia de diferencias entre las varianzas de las superficies *planas* y *onduladas*, se puede usar el algoritmo de interpolación de Kriging sin que afecte altimétricamente la geometría de las formas de la superficie terrestre, y en cambio sí presenta formas más suavizadas asemejándose a su verdadera forma.

No se recomienda utilizar el algoritmo de Kriging para representar formas de la superficie terrestre mediante DEM para territorios que presenten formas topográficas categorizadas como: *escarpadas* y *montañosas*, esto de acuerdo con el criterio de evaluación del Análisis de Varianza de los Errores Medios Cuadráticos de las superficies modeladas comparadas con el algoritmo TIN. Para estos tipos de terreno, muy usuales en la zona Andina Colombiana, se sugiere utilizar el interpolador lineal TIN.

Estadísticamente no existen diferencias significativas entre el cálculo de volúmenes de movimiento de tierra en relación a los dos algoritmos evaluados en este documento: TIN vs Kriging, esto quedó demostrado mediante el análisis No 2 donde la variable independiente fue cambiada al  $\Delta$  (delta) volumen.

Este estudio es limitado por las características de la fuente de información, sus errores inherentes y el software utilizado para su desarrollo, de modo que las conclusiones no puedan extrapolarse más allá de estos términos.

## BIBLIOGRAFÍA

1. Petrie G, Kennie TJM. *Terrain Modelling in surveying and civil engineering*. London, Whittles Publishing Services; 1990.
2. Felicísimo AM. *Modelos digitales de terreno*. Oviedo, Pentalfa Ediciones; 1994.
3. Fisher PE, Tate NJ. Causes and consequences of error in digital elevation models. *Prog. in Phys. Geogr.* 2006; 30 (4), 467-489.
4. Fisher PF. Extending the applicability of Viewsheds in landscape planning. *Photogramm. Eng. Remote Sens.* 1996; 62 (11): 1279-1302.
5. Wood J. *The Geomorphological Characterization of Digital Elevation Models*. Ph.D. Disertación, Department of Geography, University of Leicester, Leicester. UK. 1996.
6. Collins SH, Moon GC. 1981. Algorithms for dense digital terrain models. *Photogramm. Eng. Remote Sens.* 1981; 47, pp. 71-76.
7. Mark DM. 1997. The history of geographic information systems: invention and re-invention of triangulated irregular networks (tins) Proceedings, [Internet]. 1997. 2013 [citado 14 Nov 2012] GIS/LIS'97. NCGIA, Department of Geography University at Buffalo Buffalo, NY 14261-0023. Disponible en: <http://www.geog.buffalo.edu/~dmark/>
8. Morillo Barragán J, Pozo Baselga J, Pérez Preciado F, Rodríguez Gordillo MC, Rebollo Castillo FJ. Análisis de calidad de un modelo digital de elevaciones generado con distintas técnicas de interpolación. En: XIV Congreso Internacional de Ingeniería Gráfica. Santander, España; 2002.
9. Ley R. *Accuracy assessment of digital terrain models*. London: 1986.
10. Márquez Joaquín. *Modelos digitales de elevaciones, métodos de interpolación y parámetros morfológicos del relieve [tesis doctoral]*. Sevilla España: Universidad de Sevilla, Facultad de Geografía e Historia, Departamento de geografía física y análisis geográfico regional; 2004.
11. Weibel R, Heller M. Digital terrain modeling, in Maguire DJ, Goodchild MF, and Rhind DW (eds), *Geographical Information Systems: Principles and Applications*, London: Longman, 1991.
12. Munar Óscar. Determinación del potencial de los modelos digitales de elevación como fuente de datos para la evaluación de la aptitud de las tierras. Caso del cultivo de mango / Evaluation of digital elevation models as a source of data for the land suitability evaluation. Case of mango. [tesis de maestría]. Universidad Nacional de Colombia, Facultad de ciencias agrarias; 2010.
13. Florinsky Igor V. Combined analysis of digital terrain models and remotely sensed data in landscape investigations. *Prog. Phys. Geogr.* 1998; (22): 33-60.
14. Wise SM, Lane SN, and Richard KS. The effect of interpolation errors on the use of DEMs in geomorphology. *Landform modeling and analysis.* 1998; (12): 139-164.
15. Imhof E. *Cartographic relief presentation*. New York: Walter de Gruyter; 1982.
16. Balace AE. Determination of optimum sampling intervals in grid digital elevation models (DEM) data acquisition. *Photogramm. Eng. Remote Sens.* 1987; (53): 323-330.
17. Li Zhilin. Theoretical models of the accuracy of digital terrain models: an evolution and some observations. *Photogramm. Rec.* 1993; 14(82): 651-660.
18. Zhu H, and Schneider K. Flat feature processes from triangulated irregular networks for hydrological modeling. En: Proceedings of the 4th International Conference on GeoComputation; Virginia USA 1999 Jul 25-28. Virginia: Mary Washington College Fredericksburg; 2000.
19. Wentwort C. A simplified method for determining the average slope of land surface. *Am. J. Sci.* 1930; (20): 184-194.
20. Särndal, Swensson, and Wretman. *Model Assisted Survey Sampling*. Ed. Springer-Verlag; 1991.