

Aplicación de los test de equivalencia al control tipo temático de magnitudes asociadas a un Modelo Digital de Elevaciones

Application of an equivalence test to the thematic-type control of magnitudes derived from a Digital Elevation Model

María Virtudes Alba-Fernández¹

Recibido 24 de febrero de 2021; aceptado 23 de mayo de 2021

RESUMEN

Este trabajo propone un método de control de calidad de tipo temático (basado en clases) aplicable a los casos de las pendientes y orientaciones derivadas de Modelos Digitales de Elevaciones (MDE). La novedad del método que se propone es adoptar la perspectiva de detectar “diferencias relevantes” entre el producto y la referencia y no de igualdad exacta, que es lo usual en los contrastes de hipótesis más generalizados. El trabajo con categorías robustece la aproximación a las variables que se controlan, y la perspectiva de “diferencias relevantes” permite controlar el grado de “cercanía” que se desea por medio de la distancia de Hellinger. En el caso práctico que se presenta se comparan los modelos de pendiente y orientaciones derivados de dos productos MDE, el primero con resolución de 2x2 m, que se considera como referencia o patrón para el control, y el segundo con resolución de 5x5 m que es el producto a evaluar. La aplicación a este caso real permite entender mejor cómo se relacionan los aspectos más significativos del método y entender la capacidad de análisis que posee.

Palabras clave: Modelo Digital de Elevaciones, pendiente, orientación, test de equivalencia, distancia de Hellinger.

¹ Universidad de Jaén, España, correo electrónico: mvalba@ujaen.es.
ORCID: <https://orcid.org/0000-0002-4747-740X>

ABSTRACT

This paper proposes a thematic quality control method (based on classes) applicable to the cases of slope and aspect derived from Digital Elevation Models (DEM). The novelty of the proposed method is that it adopts the perspective of detecting “relevant differences” instead of “exact equality”, which is the usual one in the most generalized hypothesis testing problems. Working with categories strengthens the approach to the variables that are controlled, and the perspective of “relevant differences” allows controlling the degree of “closeness” that is desired by means of the Hellinger distance. In the practical case presented, the slope and orientation models derived from two DEM products are compared, the first is a product with a 2x2 m resolution, which is considered as a reference or standard for the control, and the second is a product with 5x5 m resolution, which is the product to evaluate. The application to this case allows us to understand how the most significant aspects of the method are related, and to understand its analytical capacity.

Key words: Digital Elevation Model, slope, aspect, equivalence test, Hellinger distance.

1. Introducción

Un Modelo Digital de Elevaciones (MDE) es un conjunto de datos, en soporte y modelo digital, que representa las elevaciones de la superficie desnuda del terreno (sin vegetación ni características artificiales). Los MDE permiten analizar las características morfológicas del relieve, su topografía (p.ej. colinas, crestas, valles, ríos, etc.), y se pueden derivar una gran cantidad de parámetros (por ejemplo, pendiente, orientación, curvatura, etc.).

Los MDE son productos de datos con aplicaciones en un amplio conjunto de disciplinas como Ingeniería Civil (Stroeven *et al.*, 2015), Hidrología (Saksena y Merwade, 2015), Geomorfología (Gómez *et al.*, 2015), Agricultura (Rekha *et al.*, 2015), Silvicultura (Juel *et al.*, 2015), etc. Dos usos principales de los MDE son el cálculo de la pendiente y la orientación (Mesa-Mingorance y Ariza-López, 2020; Ariza-López *et al.*, 2018). Utilizando algoritmos relativamente simples, los modelos digitales de pendiente y orientación se derivan de los MDE. El cálculo de estos parámetros se basa en un análisis local de la derivada de elevación. Pendiente es el grado de inclinación de una superficie. La pendiente de una ubicación en particular se calcula como la tasa máxima de cambio de elevación entre esa ubicación y sus alrededores. La pendiente se puede expresar en grados o en porcentaje. La orientación de la dirección de la pendiente máxima medida en el sentido de las agujas del reloj en el intervalo [0, 360], donde 0° representa orientación al norte, 90° orientación al este, 180° orientación al sur, etc. La pendiente es relevante para trabajos de ingeniería y riesgos naturales (por ejemplo, deslizamientos de laderas) (Van Westen, 2013). La orientación está relacionada con la insolación y, por tanto, tiene una gran relevancia para las plantas y el uso de la energía solar (Kumar *et al.*, 1997).

En algunos casos puede ser necesario evaluar la similitud entre dos modelos digitales de pendiente u orientación, bien en el sentido de similitud, donde los dos modelos a comparar tienen la misma consideración, o en el sentido de control de calidad, donde uno de los modelos se considera como referencia o patrón (Ariza López, 2002). En cualquiera de los dos contextos, es más robusto trabajar con un conjunto de categorías, que se pueden modelizar mediante ley multinomial, que, con una variable continua como la pendiente, o circular como la orientación. Esto se ve reforzado si las pruebas estadísticas (test de hipótesis) se basan en diferencias relevantes y no en una perspectiva de igualdad exacta (véase Sección 2). Por ello, es relevante proponer controles de calidad para estos dos parámetros y, en este trabajo, se presenta una de las primeras propuestas de control basado en un test estadístico denominado “test de equivalencia”.

El objetivo de este trabajo es proponer un control de calidad de tipo temático (basado en clases), es decir, en clases o categorías de pendiente y orientación definidas a voluntad por parte del usuario, pero con una perspectiva de detectar diferencias relevantes entre el producto y la referencia, no de igualdad exacta, como es usual en los contrastes de hipótesis al uso, si no permitiendo una holgura en una medida de distancia entre el producto a evaluar y el producto-patrón o producto-referencia.

El documento se organiza de la siguiente manera, tras esta introducción, la sección 2 presenta el método estadístico que se propone, en el que se incluyen los resultados de un estudio de simulación para el estudio de la significación y potencia del test propuesto. La sección 3 ofrece un ejemplo de aplicación al caso de unos datos reales relativos a la zona de las Bárdenas Reales (Navarra, España). Se analizará lo que ocurre con dos clasificaciones de las pendientes y con las orientaciones. La sección 4 presenta una discusión centrada en el método, sus aspectos más significativos y en los resultados de la aplicación desarrollada. Finalmente, la sección 5 se refiere a las conclusiones.

2. Metodología

Consideraremos que dada una variable genérica Y (pendiente u orientación) se va a categorizar atendiendo a un conjunto de tolerancias o umbrales que determinarán la definición de cada categoría. Supondremos que el número de categorías es k por lo que dada una muestra aleatoria simple de n observaciones de la variable Y , el vector que recoge conjuntamente el número de observaciones en cada categoría, $X = (X_1, X_2, \dots, X_k)'$, sigue una distribución multinomial de parámetros n y $P = (p_1, \dots, p_k)'$, $p_i > 0, 1 \leq i \leq k, \sum_{i=1}^k p_i = 1$, donde p_i representa la probabilidad de ocurrencia asociada a la categoría i -ésima, $X \rightarrow M(n, P)$.

Por otro lado y dado el conjunto de referencia, la clasificación anterior aplicada sobre él nos proporciona el vector $P_0 = (p_{01}, \dots, p_{0k})'$ que representa los valores de referencia que determinarán los criterios de calidad para la clasificación de la variable Y . Nótese que los porcentajes indicados en P_0 son fijos y se obtienen como la probabilidad de que la variable Y tome valores en el intervalo que define cada categoría sobre el MDE de referencia.

Nótese que el vector P_0 puede incluso venir definido por el usuario (experto) a partir de las proporciones observadas sobre la categorización utilizada en otros modelos digitales (anteriores en el tiempo o en otras áreas de estudio).

Para determinar si los niveles de calidad indicados por la categorización en un conjunto de datos de referencia se alcanzan por parte del producto, existen una variedad de test de hipótesis que permiten dar una respuesta. Sin embargo todos ellos tienen en común la definición de una hipótesis nula formulada en términos de igualdad, $P = P_0$ y una hipótesis alternativa en términos de desigualdad, $P \neq P_0$. En este tipo de test en el caso de que no existan evidencias para rechazar la hipótesis, la conclusión no debe ser que la hipótesis nula sea cierta. En otras palabras, puesto que el diseño de los test clásicos no está enfocado a la validación de la hipótesis nula, que no tengamos evidencia suficiente para rechazar que los niveles de calidad no se cumplen ($P \neq P_0$), no significa que dichos niveles se hayan alcanzado ($P = P_0$).

El objetivo de este trabajo es proponer un procedimiento formal (test estadístico) para determinar si los niveles de calidad de un producto MDE se alcanzan en el sentido de que las diferencias entre P y P_0 son suficientemente pequeñas como para considerar que ambas distribuciones multinomiales son iguales excepto "desviaciones irrelevantes".

Para cuantificar el grado de similitud entre P y P_0 vamos a considerar como medida de distancia el cuadrado de la distancia de Hellinger. Esta opción no es casual, esta medida de distancia se ha utilizado previamente con buenos resultados en problemas inferenciales relacionados con la distribución multinomial. Destacamos su uso en el estudio de similaridad de patrones espaciales de puntos (Alba-Fernández y Ariza-López, 2018; Alba-Fernández *et al.*, 2016) o de la similaridad temática mediante matrices de confusión (García-Balboa *et al.*, 2018).

Consideremos $H(P, P_0)$ como el cuadrado de la distancia de Hellinger entre P y P_0 definida como

$$H(P, P_0) = \sum_{i=1}^k (\sqrt{p_i} - \sqrt{p_{0i}})^2. \quad (1)$$

Para un valor $\varepsilon > 0$, se plantea el siguiente contraste de hipótesis, conocido también como test de equivalencia

$$\begin{aligned} H_{0\varepsilon}: H(P, P_0) &\geq \varepsilon \\ H_{1\varepsilon}: H(P, P_0) &< \varepsilon. \end{aligned} \quad (2)$$

Puesto que $0 \leq H(P, P_0) \leq 2$ y $H(P, P_0) = 0$ si y solo si $P = P_0$, rechazar H_0 supone que las distribuciones multinomiales son iguales excepto desviaciones irrelevantes dado el valor residual de ε .

Destacamos que el planteamiento de las hipótesis nula y alternativa en un test de equivalencia contrastan con la formulación de un test "clásico" de una hipótesis nula de "igualdad" frente a una alternativa de "no igualdad". Ahora, en

el test de equivalencia planteado, rechazar H_{0e} supondrá que las distribuciones con vector de probabilidades P y P_0 son prácticamente iguales atendiendo a la distancia considerada entre ellas. Por lo tanto, la conclusión del rechazo es confirmatorio sobre la verdadera distribución que representa los datos observados, es decir, es una prueba de validación de que los niveles de calidad marcados por el conjunto de referencia a través del modelo multinomial, $M(n, P_0)$, se alcanzan por parte del producto, $M(n, P)$.

En este contexto, dado el vector de frecuencias relativas $\hat{P} = (\hat{p}_1, \dots, \hat{p}_k)'$, $\hat{p}_i = \frac{x_i}{n}, \forall i$, que es el estimador de máxima verosimilitud de P , y a partir de $H(\hat{P}, P_0)$, estimador consistente de $H(P, P_0)$, el resultado del Corolario 3.1 en Zografos *et al.* (1990) para $P \in H_{0e}$, nos permite obtener la distribución asintótica de $H(\hat{P}, P_0)$,

$$\sqrt{n}(H(\hat{P}, P_0) - H(P, P_0)) \xrightarrow{\mathcal{L}} N(0, \sigma^2) \tag{3}$$

cuando $n \rightarrow \infty$, donde:

$$\sigma^2(P) = 1 - \left(\sum_{i=1}^k \sqrt{p_i p_{0i}} \right)^2 \tag{4}$$

y $\xrightarrow{\mathcal{L}}$ identifica convergencia en distribución.

Tomando $\hat{\sigma}^2 = \sigma^2(\hat{P})$ como estimador consistente de $\sigma^2(P)$, se verifica que

$$\sqrt{n} \frac{(H(\hat{P}, P_0) - H(P, P_0))}{\hat{\sigma}} \xrightarrow{\mathcal{L}} N(0, 1), \tag{5}$$

cuando $n \rightarrow \infty$.

Por tanto, dado un nivel de significación $\alpha \in (0, 1)$, rechazaremos la hipótesis nula H_{0e} si

$$T_n = \sqrt{n} \frac{(H(\hat{P}, P_0) - \varepsilon)}{\hat{\sigma}} \leq Z_\alpha \tag{6}$$

siendo Z_α el percentil α de la distribución normal estándar.

A partir de este resultado, el test que rechaza H_{0e} según (6) tiene asintóticamente nivel de significación α y es consistente frente a cualquier alternativa fija.

2.1. Estudio del nivel de significación

El test de equivalencia presentado se basa en el comportamiento asintótico del test estadístico T_n . Es decir, para un valor de ε , y para un nivel de significación α , la distribución asintótica, para tamaños de muestra grandes, es una distribución normal estándar.

Sin embargo, para tamaños de muestra pequeños o moderados se hace necesario analizar el comportamiento del test. En concreto, se ha diseñado un conjunto de simulaciones para evaluar sobre datos simulados el nivel de significación y la potencia del test teniendo en cuenta la relación existente entre el número de categorías k , el valor de ϵ y el tamaño de muestra para que la aproximación de T_n a la distribución normal sea adecuada, todo ello, dado un nivel un significación fijado $\alpha \in (0,1)$.

Con este objetivo, describimos brevemente las condiciones del estudio de simulación:

1. Se ha considerado $P_0 = (1/k, \dots, 1/k)'$ para valores de k oscilando entre 3 y 7. La elección de este caso de estudio no es azarosa puesto que se ha utilizado ya previamente para estudiar el comportamiento de diversos test estadísticos vinculados a problemas de inferencia relacionados con la distribución multinomial (Alba-Fernández y Jiménez-Gamero, 2009; Basu *et al.*, 2002; Pardo, 1999). Con respecto al número de categorías, creemos que se cubre el rango más habitual observado en las clasificaciones sobre la pendiente y orientaciones en un MDE (Elewa y Qaddah, 2011; Mikulska *et al.*, 2014; Mogaji *et al.*, 2015; Sheng, 1990; CEN prEN 1998-1, 2003; NTC, 2008).
2. Para el estudio del nivel de significación, debemos considerar casos incluidos en H_{0e} . En particular, se han considerado un conjunto de configuraciones para P de forma que den lugar a unos valores de ϵ entre 0.1 y 0.2, alterando los porcentajes equiprobables de P_0 en alguna o varias categorías. Los casos concretos utilizados en la simulación se presentan en la Tablas 1, Tabla 2 y Tabla 3, para los tres valores de k considerados.
3. Para cada configuración, P , se han generado muestras de la correspondiente multinomial, $M(n, P)$, para varios tamaños de muestra $n=50, 100, 200, 500$ y se ha aplicado el criterio de decisión descrito en (6) para un nivel de significación $\alpha=0.05$.
4. Repetido el procedimiento 100000 veces se ha determinado el porcentaje de rechazo de la hipótesis nula, lo que nos proporciona una estimación del error tipo I o nivel de significación α . La Tabla 4 y la Tabla 5, resumen los resultados obtenidos para cada valor de k .
5. Para el estudio de la potencia del test, debemos generar ahora bajo la hipótesis alternativa, por tanto, repetiremos los pasos 3) y 4) tomando como configuración para P el caso P_0 para $n=50, 100, 200$. La Tabla 6 muestra los valores estimados de la potencia tomando $\epsilon=0.15$ y $\alpha=0.05$.

Puesto de dichos porcentajes de rechazo presentados en la Tabla 4 y la Tabla 5 representan las probabilidades estimadas de tipo I (nivel de significación), podemos observar cómo para valores pequeños de k , dicho nivel se alcanza para tamaños de muestra bajos ($n=50$ para $k=3$, independientemente del valor de ϵ), mientras que la aproximación a la distribución normal en (5) requiere tamaños de muestra superiores conforme el valor de k va aumentando o el valor de ϵ va disminuyendo. Para tamaños de muestra para los que las probabilidades estimadas de tipo I son inferiores a 0.05, el test es conservativo, es decir, rechaza

en menor medida que el nivel de significación. Sin embargo, para este nivel de significación (5%), la potencia del test es muy elevada.

Tabla 1. Configuraciones de P para $k=3$

<i>Casos</i>	p_1	p_2	p_3
Caso 1	0.1758	0.1758	0.6483
Caso 2	0.1433	0.1433	0.7133
Caso 3	0.1167	0.1167	0.7667

Tabla 2. Configuraciones de P para $k=5$

<i>Casos</i>	p_1	p_2	p_3	p_4	p_5
Caso 1	0.1350	0.1350	0.1350	0.1350	0.4600
Caso 2	0.1400	0.1500	0.1400	0.1500	0.4200
Caso 3	0.1535	0.1535	0.1535	0.1535	0.3860

Tabla 3. Configuraciones de P para $k=7$

<i>Casos</i>	p_1	p_2	p_3	p_4	p_5	p_6	p_7
Caso 1	0.1300	0.1300	0.1300	0.1300	0.1300	0.1300	0.2200
Caso 2	0.1229	0.1229	0.1229	0.1229	0.1229	0.1229	0.2629
Caso 3	0.1155	0.1155	0.1155	0.1155	0.1155	0.1155	0.3071

Tabla 4. Porcentajes de rechazo para el nivel de significación del 5% ($k=3$ y $k=5$)

<i>n</i>	$k=3$			$k=5$		
	<i>Caso 1</i>	<i>Caso 2</i>	<i>Caso 3</i>	<i>Caso 1</i>	<i>Caso 2</i>	<i>Caso 3</i>
ϵ	0.1	0.15	0.2	0.1	0.15	0.2
50	0.057	0.047	0.052	0.035	0.034	0.038
100	0.056	0.056	0.049	0.039	0.039	0.046
200	0.057	0.053	0.049	0.042	0.043	0.050
500	0.054	0.052	0.050	0.046	0.045	0.052

Tabla 5. Porcentajes de rechazo para el nivel de significación del 5% para $k=7$

n	Caso 1	Caso 2	Caso 3
ε	0.1	0.15	0.2
50	0.019	0.022	0.023
100	0.027	0.028	0.029
200	0.033	0.034	0.033
500	0.038	0.040	0.040
1000	0.044	0.041	0.043
2000	0.046	0.044	0.044

Tabla 6. Estimaciones de la potencia para un nivel de significación del 5%

n	$k=3$	$k=5$	$k=7$
50	0.999	0.995	0.977
100	1	1	1
200	1	1	1

3. Aplicación

Para demostrar la aplicación del enfoque y métodos propuestos, se va a trabajar con datos correspondientes a dos productos MDE de los que se derivarán sendos modelos de pendientes y orientaciones. El área de trabajo se corresponde con una zona de las Bárdenas Reales (provincia de Navarra, España), que está delimitada por la hoja 0245 del Mapa Topográfico Nacional de España. Los productos originales son:

- MDE02. Modelo digital de elevaciones de paso de malla de 2x2 m procedente de un levantamiento LiDAR del año 2017, dentro del proyecto PNOA (Plan Nacional de Ortofotografía Aérea, <https://pnoa.ign.es/el-proyecto-pnoa-lidar>).
- MDE05. Modelo digital de elevaciones de paso de malla de 5x5 m procedente de un levantamiento LiDAR del año 2012, dentro del proyecto PNOA.

Ambos conjuntos de datos están disponibles en el centro de descargas del Centro Nacional de Información Geográfica (España) (<http://centrodedescargas.cnig.es/CentroDescargas>). La Figura 1 muestra el MDE sobre la zona de estudio.

De los modelos anteriores, y por medio de operaciones convencionales de análisis SIG de la caja de herramientas de ArcGIS™, se han derivado los modelos de pendientes (PTE02 y PTE05) y orientaciones (ORI02 y ORI05). En este trajo los modelos de menor paso de malla (2x2 m) son utilizados como referewncia o patrón y los de mayor paso de malla (5x5 m) como producto a evaluar.



Figura 1. Modelo de elevaciones (Bárdenas Reales, Navarra, España).
Fuente: Elaboración propia a partir de datos del IGN.

3.1 Pendiente

En esta zona de estudio, se han considerado dos clasificaciones para la pendiente. Cuando Dichas clasificaciones se aplican al modelo PTE02 dan lugar a los valores “patrón” que serán los consideraremos como valores fijos en el contraste (2) y que el producto PTE05 debe cumplir. En la Tabla 7 y la Tabla 8 se resumen dichos porcentajes, cuya representación espacial se muestran en la Figura 2.

Tabla 7. Clasificación #1

<i>Grados</i>	<i>Clasificación</i>	<i>P₀</i>
0-1	Llano	0.23
-1-7.5	Moderado	0.54
>7.5	Elevado y muy elevado	0.23

Fuente: elaboración propia.

Tabla 8. Clasificación #2

<i>Grados</i>	<i>Clasificación</i>	P_0
0-2	Llano casi llano	0.40
2-8	Ligeramente inclinado	0.38
8-15	Fuertemente inclinado	0.10
15-30	Muy inclinado	0.09
>30	Empinado	0.03

Fuente: elaboración propia.

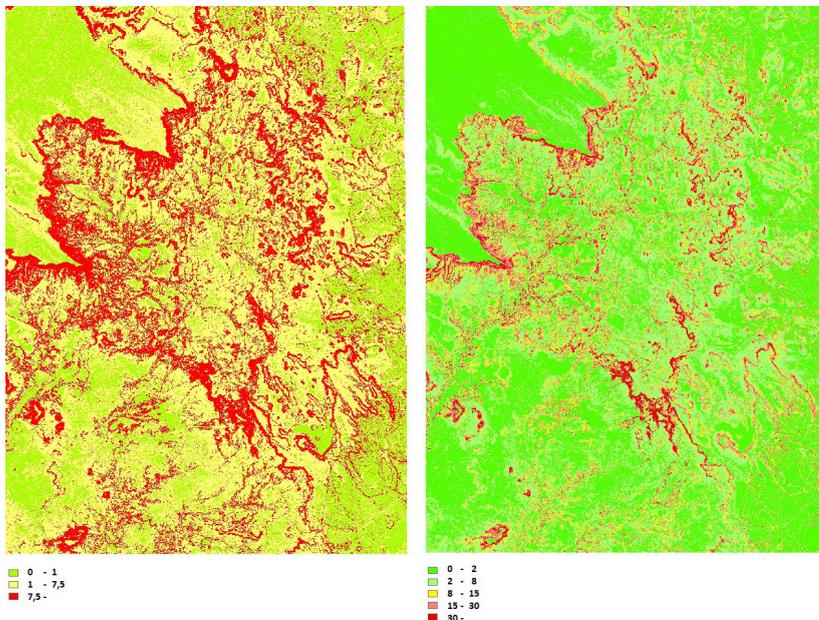


Figura 2. Modelos de pendientes sobre DEM02 según las clasificaciones #1 (izquierda) y #2 (derecha).

Recordemos que para un valor ε positivo y suficientemente pequeño, rechazar $H_{0\varepsilon}$ supondrá que los valores “patrón” se han alcanzado. En nuestro caso, aplicamos (2) para $\varepsilon=0.1$, teniendo en cuenta que $Z\alpha=-1.645$ para $\alpha=0.05$. Para ello, hemos extraído muestras aleatorias simples de tamaños $n=100, 200$ y 500 para asegurarnos la correcta aplicación de la clasificación #2 independientemente del valor de ε . La Tabla 9 y la Tabla 10, muestran los porcentajes observados en el producto para cada clasificación, el estadístico de contraste T_n para cada caso y el resultado de la regla de decisión. Los resultados reflejados en ambas tablas nos llevan a la conclusión de que efectivamente los niveles de calidad

representados por los valores “patrón” en el producto PTE05 se cumplen, tanto para la clasificación #1 como para la clasificación #2. Este nivel de cumplimiento debe entenderse como que la multinomial observada en PTE05 y la definida por el “patrón” (PTE02) son iguales, salvo diferencias mínimas en términos de la distancia de Hellinger.

Tabla 9. Aplicación del test de equivalencia sobre la clasificación #1

<i>Tamaño de muestra</i>	\hat{p}	T_n	<i>Decisión</i>
100	(0.280, 0.540, 0.180)	-12.73	Rechazar H_{0e}
200	(0.235, 0.550, 0.215)	-78.11	Rechazar H_{0e}
500	(0.240, 0.528, 0.232)	-167.23	Rechazar H_{0e}

Fuente: elaboración propia.

Tabla 10. Aplicación del test de equivalencia sobre la clasificación #2

<i>Tamaño de muestra</i>	\hat{p}	T_n	<i>Decisión</i>
100	(0.38, 0.44, 0.09, 0.07, 0.02)	-13.46	Rechazar H_{0e}
200	(0.405, 0.400, 0.110, 0.070, 0.015)	-20.54	Rechazar H_{0e}
500	(0.382, 0.398, 0.112, 0.082, 0.026)	-67.12	Rechazar H_{0e}

Fuente: elaboración propia.

3.2 Orientaciones

En el caso de las orientaciones, hemos considerado la clasificación usual en ocho sectores (N, NW, W, SW, S,...). Los valores de las proporciones correspondientes a estas categorías de orientación en ORI02 (datos de referencia), se muestran en la Tabla 11, en ella no aparece la categoría de terreno llano (sin orientación), pues es muy minoritaria y casi nula. Dichos valores vuelven a considerarse como el “patrón” que el producto ORI05 debe cumplir. La Figura 3 representa el modelo de orientaciones de esta zona.

Tabla 11. Clasificación de la orientación

<i>Grados</i>	<i>Clasificación</i>	P_o
337.5-360, 0-22.5	Norte (N)	0.10
22.5-67.5	Nordeste (NE)	0.11
67.5-112.5	Este (E)	0.12
112.5-157.5	Sureste (SE)	0.12
157.5-202.5	Sur (S)	0.14
202.5-247.5	Suroeste (SW)	0.16

Continuación Tabla 11

Grados	Clasificación	P_0
247.5-292.5	Oeste (W)	0.14
292.5-337.5	Noroeste (NW)	0.11

Fuente: elaboración propia.

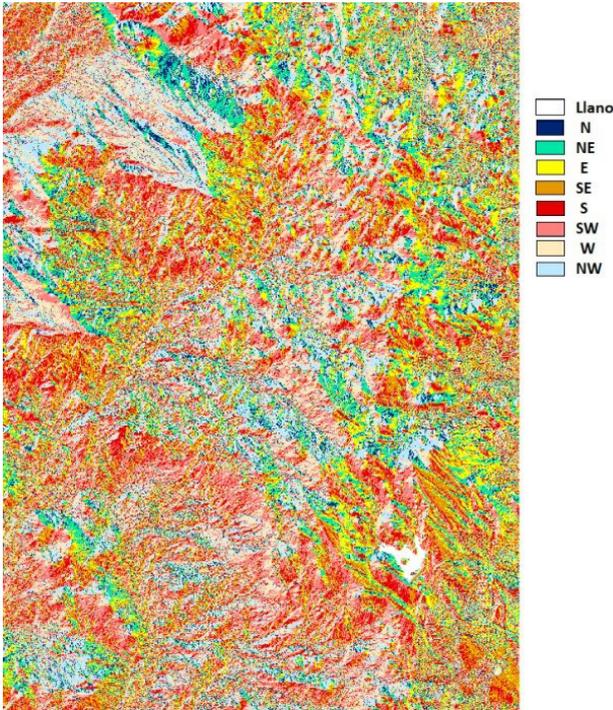


Figura 3. Modelo de orientaciones sobre DEM02.

Aplicamos nuevamente el test (2) con $\varepsilon=0.1$, y para asegurarnos la correcta aplicación del mismo, se ha seleccionado una muestra aleatoria simple de tamaño 500. Para dicha muestra, las frecuencias relativas observadas en cada categoría fueron 0.092, 0.116, 0.098, 0.122, 0.124, 0.160, 0.150 y 0.120, respectivamente. Para estos valores, el valor del estadístico de contraste fue $T_n = -15.40$ y puesto que $Z_{0.05} = -1.645$, el criterio de decisión (6) en este caso da lugar al rechazo de la hipótesis nula, es decir, el cuadrado de la distancia de Hellinger entre el vector de frecuencias relativas y los valores “patrón” es inferior a 0.1 por lo que se considera que el producto ORI05 ha alcanzado los niveles de referencia para la clasificación de la orientación y que se han establecido por medio de ORI02.

4. Discusión

La discusión se va a desarrollar en tres líneas, la primera sobre el método, la segunda sobre sus aspectos más significativos, y la tercera sobre los resultados.

Consideramos que la aplicación del método es directa. Desde un punto de vista estadístico supone la realización de un contraste de hipótesis donde la hipótesis nula y la alternativa están establecidas de manera diferente a las hipótesis nula y alternativa convencionales que proponen una igualdad y una no igualdad. Esta diferencia supone por un lado un cambio conceptual en cuanto a lo que supone rechazar o no una hipótesis nula, aunque desde el punto de vista de los test estadísticos, no supone ningún tipo de cambio con respecto a la interpretación del nivel de significación y potencia asociados a todo test de hipótesis, pero este cambio en la formulación de las hipótesis, si introduce matices distintos sobre lo que se contrasta y, por ende, sobre lo que se acepta y rechaza desde el punto de vista del control de la calidad. La Tabla 12 resume estos cambios.

Tabla 12. Interpretación de los errores tipo I y II

<i>Errores</i>	<i>Test clásico</i>	<i>Test de equivalencia</i>
$\alpha = P[\text{Rechazar } H_0 / H_0 \text{ cierta}]$ (riesgo del productor)	Probabilidad de rechazar un producto cuando es bueno	Probabilidad de aceptar un producto malo cuando efectivamente lo es
$\beta = P[\text{No rechazar } H_0 / H_0 \text{ es falsa}]$ (riesgo del usuario)	Probabilidad de aceptar un producto cuando es malo	Probabilidad de aceptar un producto malo cuando el producto es bueno
Potencia = $1 - \beta$	Probabilidad de rechazar un producto cuando es malo	Probabilidad de aceptar el producto cuando el producto es bueno

Fuente: elaboración propia.

En relación a los aspectos más significativos consideramos que éstos son el número de categorías, el tamaño de muestra y la distancia de Hellinger, los cuáles se encuentran muy interrelacionados en el método y necesitan un ajuste conjunto. El método se basa en la comparación de distribuciones multinomiales, por lo cual le es connatural el trabajar con categorías. El método se puede aplicar a variables categorizadas, pero en este trabajo se ha aplicado a variables continuas que se han discretizado. Estas categorías las puede definir el usuario a voluntad, pero su número no es un aspecto neutro. Un mayor número de categorías lleva a mayores tamaños de muestra, lo cual es lógico desde la perspectiva de los procesos de estimación sobre multinomiales.

Recordemos que la regla de decisión (6) se basa en la aproximación asintótica bajo H_0 del estadístico de contraste T_n , y por tanto, el nivel de significación α se alcanzará de manera adecuada a partir de un determinado tamaño de

muestra. Dicha aproximación se alcanzará para tamaños de muestra menores para valores de k menores (independientemente del valor de ε), y se hace más costosa en muestreo si k aumenta. Esta consideración es usual a otros test de hipótesis vinculados a la distribución normal, como por el ejemplo el test χ^2 de Pearson, cuya aplicación se basa en aproximaciones asintóticas.

Por ello se recomienda tomar siempre el menor número de categorías posibles. En cualquier caso, consideramos que, con los ejemplos de aplicación mostrados, donde $k=3, 5$ y 8 , se demuestra que se puede aplicar a los números de categorías más usuales.

Por otro lado, el test de hipótesis no se realiza sobre una medida natural al problema (p.ej. grados, radianes, metros, etc.), si no por medio de la distancia de Hellinger, la cual es desconocida para la que la mayoría de los usuarios y, aunque se entienda su base matemática, no se tiene demasiada sensibilidad sobre ella. Es decir, ¿ $\varepsilon = 0.1$ es mucho o es poco? Para dar respuesta a esta pregunta debemos tener en cuenta la definición de $H(P, P_0)$, relacionada en cierta forma con la distancia euclídea sobre las raíces cuadradas de las frecuencias observadas y patrón, junto con el interés que el usuario ponga en que las “discrepancias” mencionadas sean o no de una determinada magnitud. En cualquier caso, esta limitación es muy usual en la aplicación de herramientas estadísticas a datos geoespaciales (p.ej. análisis cluster, etc.), donde aparecen diversas distancias (p.ej. Minkowski, Chebyshev, Manhattan, Mahalanobis, etc.) que tampoco son distancias “naturales” al problema. Esto conlleva la necesidad de alcanzar la suficiente sensibilidad en el manejo de la distancia que se aplica.

En el apartado 2.1 se ha desarrollado un estudio del nivel de significación que puede dar orientaciones para los casos que se han considerado en él, pero que también puede servir de guía para que los interesados en el método realicen sus ajustes entre número de categorías y ε .

No hay que perder de vista que el resultado es general para cualquier “patrón” P_0 y que se puede adaptar a otro contexto distinto del planteado en el manuscrito (equiprobabilidad), para ello, basta repetir los pasos 1)-5) indicados en la sección 2.1 para el valor P_0 objeto de estudio y otras configuraciones de P de forma que $H(P, P_0) = \varepsilon$, para un $\varepsilon > 0$ dado.

Finalmente, en relación a los resultados de los tres casos prácticos presentados, consideramos que se ha evidenciado la posibilidad de aplicación de esta metodología a las variables pendiente y orientación derivadas de MDE. Dado que los modelos MDE02 y MDE05 son de calidad alta, de fechas cercanas y que no han existido grandes cambios territoriales, el resultado de la aplicación es que los modelos derivados para las pendientes (PTE02 y PTE05), y para las orientaciones (ORI02 y ORI05), no muestran entre ellos diferencias superiores al valor de ε considerado. Un aspecto importante de la aplicación práctica es el tamaño de muestra necesaria para este control de calidad, el cual tiene repercusiones directas sobre el coste. En los ejemplos para la pendiente se han considerado tres tamaños de muestra ($n=100, 200, 500$) y, como se puede observar, en todos ellos se llega a la misma conclusión (rechazo de H_{0e}). A mayor

tamaño de n , el estadístico T_n obtenido también es mayor, lo que significa que se tiene más confianza en esta decisión. Por otro lado, también se evidencia que el crecimiento de T_n es menor con n en el caso de $k=5$ que en el caso de $k=3$, lo cual es lógico, pues mayor número de categorías requiere más esfuerzo muestral. En cualquier caso, el menor valor de T_n supera con creces el valor de Z_α considerado, lo que permite considerar que se podrían utilizar tamaños de muestra algo menores, especialmente para $k=3$ donde es suficiente un tamaño de muestra de $n=50$ para obtener resultados fundamentados. Finalmente, en relación a la ejecución de estos muestreos en campo, no conviene olvidar que la medición de la pendiente y orientación en campo suele realizarse de manera simultánea o subrogada a otro tipo de trabajos (p.ej. catas edáficas), por lo que su disponibilidad estará en la mayoría de las veces condicionadas por este tipo de actuaciones.

5. Conclusión

En este trabajo se ha presentado una metodología para controlar algunas magnitudes derivadas de un MDE por medio de un proceso de categorización y del modelo multinomial resultante. Para realizar dicho control se propone un test de equivalencia cuya peculiaridad es el planteamiento de las hipótesis nula y alternativa a contrastar. En concreto se definen en términos de una medida de distancia entre la multinomial determinada por el producto y la correspondiente al conjunto de referencia, de forma que se considera que ambas distribuciones son equivalentes si el cuadrado de la distancia de Hellinger entre ellas es inferior a un umbral positivo y suficientemente pequeño ε .

Por otro lado, dado que los resultados del test se basan en la aproximación asintótica del estadístico de contraste a la distribución normal, hay que tener en cuenta el número de categorías y ε para determinar el tamaño de muestra necesario para su correcta aplicación, consideración por otra parte usual en test de hipótesis cuya regla de decisión se base en aproximaciones asintóticas.

El método se plantea de manera general para cualquier magnitud susceptible de ser categorizada, de tal manera que el punto de partida sea una distribución multinomial. Esto es, su uso es general y aplicable en otros contextos, como por ejemplo en control de la calidad temática en mapas de cobertura del terreno (en el contexto de las referencias García-Balboa *et al.*, 2018), o en el estudio de la similaridad de patrones espaciales de puntos (en el contexto de las referencias Alba-Fernández *et al.*, 2016 y Alba-Fernández y Ariza-López, 2018), entre otros posibles usos.

Se ha evidenciado la posibilidad de aplicación de esta metodología a las variables pendiente y orientación derivadas de MDE, no mostrando grandes diferencias entre los modelos derivados para las pendientes (PTE02 y PTE05) y para las orientaciones (ORI02 y ORI05).

Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto de investigación “Calidad Funcional en Modelos Digitales de Elevaciones del Terreno en Ingeniería” (https://coello.ujaen.es/investigacion/web_giic/funquality4dem/) de la Agencia Estatal de Investigación. PID2019-106195RB-I00/AEI/10.13039/501100011033.

Bibliografía

- Alba-Fernández, V., & Jiménez-Gamero, M. D. (2009). Bootstrapping divergence statistics for testing homogeneity in multinomial populations. *Mathematics and Computers in Simulation*, 79, 3375-3384. <https://doi.org/10.1016/j.matcom.2009.04.002>
- Alba-Fernández, M. V., & Ariza-López, F. J. (2018). A homogeneity test for comparing gridded-spatial-point patterns of human caused fires. *Forests*, 9, 454. <https://doi.org/10.3390/f9080454>
- Alba-Fernández, M. V., Ariza-López, F. J., Jiménez-Gamero, M. D., & Rodríguez-Avi, J. (2016). On the similarity analysis of spatial patterns. *Spatial statistics*, 18, 352-362. <http://dx.doi.org/10.1016/j.spasta.2016.07.004>
- Ariza López, F. J. (2002). *Calidad en Producción Cartográfica*, Ra-Ma, Madrid.
- Ariza-López, F. J., Chicaiza-Mora, E. G., Mesa-Mingorance, J. L., Cai, J., & Reinoso-Gordo, J. F. (2018). DEMs: An Approach to Users and Uses from the Quality perspective. *International Journal of Spatial Data Infrastructures Research*, 13, 131-171. Recuperado de <https://ijsdir.sadl.kuleuven.be/index.php/ijsdir/article/view/469>
- Basu, A., Ray, S., Park, C., & Basu, S. (2002) Improved power in multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society. Series D*, 51 (3), 381-393. <https://doi.org/10.1111/1467-9884.00325>
- CEN rEN 1998-1- Eurocode 8: design of structures for earthquake resistance. Part 1: General rules, seismic actions and rules for resistance. Part 1: General rules, seismic actions and rules for buildings. Draft No 6, Doc CEN/TC250/SC8/N335, 2003, Brussels.
- Elewa, H., & Qaddah, A. (2011). Groundwater potentiality mapping in the Sinai Peninsula, Egypt, using remote sensing and GIS-watershed-based modeling. *Journal of Hydrology*, 19(2), 613-628. <https://doi.org/10.1007/s10040-011-0703-8>
- García-Balboa, J. L., Alba-Fernández, M. V., Ariza-López, F. J., & Rodríguez-Avi, J. (2018). Analysis of Thematic Similarity Using Confusion Matrices. *ISPRS International Journal of Geo-Information*, 7(6), 233. <https://doi.org/10.3390/ijgi7060233>.
- Gómez, C., Hayakawa, Y., & Obanawa, H. (2015). A study of Japanese landscapes using structure from motion derived DSMs and DEMs based on historical aerial photographs: New opportunities for vegetation monitoring and diachronic geomorphology. *Geomorphology*, 242, 11-20. <https://doi.org/10.1016/j.geomorph.2015.02.021>.
- Juel, A., Groom, G. B., Svenning, J. C., & Ejrnaes, R. (2015). Spatial Application of Random Forest models for fine-scale coastal vegetation classification using object based analysis of aerial orthophoto and DEM data. *International Journal of Applied Earth Observation and Geoinformation*, 42, 106-114. <https://doi.org/10.1016/j.jag.2015.05.008>

- Kumar, L., Skidmore, A. K., & Knowles, E. (2017). Modelling topographic variation in solar radiation in a GIS environment. *International Journal of Geographical Information Science*, 11(5), 475–497. <https://doi.org/10.1080/136588197242266>
- Mesa-Mingorance, J. L., & Ariza-López, F. J. (2020). Accuracy Assessment of Digital Elevation Models (DEMs): A Critical Review of Practices of the Past Three Decades. *Remote Sensing*, 12(16), 2630. <https://doi.org/10.3390/rs12162630>
- Mikulski, K.S, Sajnyg, N., & Krupowicz, W. (2014). Methodology of validation of agricultural real properties in Poland with the use of geographic information system tools. *Proceeding of the 14th SGEM GeoConference on Informatics, Geoinformatics and Remote Sensing*, 2, 345-356. <https://doi.org/10.5593/SGEM2014/B22/S9.044>
- Mogaji, K. A., Lim, H. S., & Abdullah, K. (2015). Regional prediction of groundwater potential mapping in a multifaceted geology terrain using GIS-based Dempster-Shafer model. *Arab. J. Geosci.*, 8, 3235-3258. <https://doi.org/10.1007/s12517-014-1391-1>
- NTC Nuove Norme Tecniche per le Costruzioni. Part 3: Categorie di sottosuolo e condizioni topografiche, *Gazzetta Ufficiale n. 29 del 4 febbraio 2008*.
- Pardo, M. C. (1999). On Burbea-Rao divergence based Goodness-of-fit tests for multinomials models. *Journal of Multivariate Analysis*, 69, 65-87. <https://doi.org/10.1006/jmva.1998.1799>
- Rekha, P. N., Gangadharan, R., Ravichandran, P., Mahalakshmi, P., Panigrahi, A., & Pillai, S.M. (2015). Assessment of impact of shrimp farming on coastal groundwater using Geographical Information System based Analytical Hierarchy Process. *Aquaculture*, 448, 491-506. <https://doi.org/10.1016/j.aquaculture.2015.06.025>
- Saksena, S., & Merwade, V. (2015). Incorporating the effect of DEM resolution and accuracy for improved flood inundation mapping. *Journal of Hydrology*, 530, 180-194. <https://doi.org/10.1016/j.jhydrol.2015.09.069>
- Sheng, T. C. (1990). *Watershed management field manual: Watershed survey and planning*. FAO Conservation guide 13/6. Food and Agriculture Organization of the United Nations, Rome.
- Stroeven, P., Li, K., Le, N.L.B., He, H., & Stroeven, M. (2015). Capabilities for property assessment on diferent levels of the micro-structure of DEM-simulated cementitious materials. *Construction and Building Materials*, 88, 105-117. <https://doi.org/10.1016/j.conbuildmat.2015.04.012>
- Van Westen, C. J. (2013). Remote sensing and GIS for natural hazards assessment and disaster risk management. In Shroder, J. (Editor in Chief), Bishop, M. P. (Ed.), *Treatise on Geomorphology*. Academic Press, San Diego, CA, 2013, vol. 3, *Remote Sensing and GIScience in Geomorphology*, 259-298.
- Zografos, K., Ferentinos, A., & Papaioannou, T. (1990). Divergence statistics: sampling properties and multinomial goodness of fit and divergence tests. *Communication in Statistics- Theory and Methods*, 19, 1785-1802. <https://doi.org/10.1080/03610929008830290>