

Landslides susceptibility assessment using discriminant analysis. Application on a regional scale

Juan Pablo Londoño-Linares

Departamento de Ingeniería Civil, Fundación Universitaria Agraria de Colombia, Bogotá, Colombia. londono.juanpa@uniagraria.edu.co

Received: December 6th, de 2016. Received in revised form: April 24th, 2017. Accepted: May 12th, 2017

Abstract

In this article, we present the stepwise application of the quantitative multivariate method known as Discriminant Analysis to obtain models of susceptibility to landslides at regional scales. As a proposal, context variables were incorporated into the analysis, which aim to account for anthropogenic action within the territory. The detailed procedure is presented, applied to a specific area, obtaining five susceptibility functions, which are analyzed in terms of performance to select the best model. In conclusion, adequate values of performance of the models are obtained, confirming the usefulness of the method as an initial step in the calculation of landslide risk at regional scale for land planning.

Keywords: Landslide susceptibility; risk assessment; land planning.

Cálculo de susceptibilidad a deslizamientos mediante análisis discriminante. Aplicación a escala regional

Resumen

En el presente artículo expone la aplicación, paso a paso, del método multivariado cuantitativo conocido como Análisis Discriminante, para la obtención de modelos de susceptibilidad a deslizamientos de tierra a escalas regionales. Como propuesta se incorporan variables de contexto dentro del análisis, que pretenden dar cuenta de la acción antrópica dentro del territorio. Se presenta el procedimiento detallado, se aplica a un área específica, obteniendo cinco funciones de susceptibilidad, las cuales se analizan en términos de rendimiento para seleccionar el mejor modelo. Como conclusión, se obtienen valores adecuados de rendimiento de los modelos, confirmando la utilidad del método como paso inicial en el cálculo de riesgo por deslizamiento a escala regional para la ordenación del territorio.

Palabras clave: Susceptibilidad a deslizamientos; análisis de riesgo; planificación del territorio.

1. Introducción

Pese a la existencia de un marco de referencia robusto sobre Evaluación de Amenazas y Riesgos de Desastre, elaborado en los últimos 30 años, con teorías concretas, es poco común en nuestro medio contabilizar pérdidas, recolectar datos, y evaluar el riesgo sistemáticamente en términos numéricos. Como consecuencia de ello, no se dispone de opciones de priorización de recursos, al no conocer el riesgo o amenaza asociado a cada porción del territorio y por tanto resulta difícil proteger a las comunidades e inversiones, y reducir la vulnerabilidad frente

a futuros impactos.

Una de las amenazas que mayores impactos genera en un entorno Andino como el nuestro, son los procesos de remoción en masa, que presentan un carácter recurrente y producen pérdidas que, si bien son contenidas, presentan una tendencia repetitiva año tras año.

Un posible enfoque para analizar el fenómeno, desde un punto de vista cuantitativo, es en función de procesos sucesivos de cálculo que terminan en la zonificación de riesgo. Se inicia definiendo susceptibilidad como la predisposición espacial, intrínseca, del territorio a sufrir daño, lo cual en un escenario posterior debe relacionarse con

la frecuencia de ocurrencia, dada por factores detonantes (amenaza), para finalmente, a través de un análisis de exposición y vulnerabilidad, obtener cuantificaciones de riesgo, en términos de pérdidas, métricas asociadas, o zonificaciones para la planificación territorial.

La ocurrencia de deslizamientos es el producto de las condiciones geológicas, hidrológicas y geomorfológicas y la modificación de éstas por procesos geodinámicos, vegetación, usos de la tierra y actividades humanas, así como la frecuencia e intensidad de las precipitaciones y la sismicidad.

Bajo este panorama, existe un fuerte componente antrópico, representado en factores como la localización y características de las carreteras, canales, tuberías, oleoductos y demás elementos, producto de la acción humana; presencia o ausencia de zonas duras o pavimentadas; las áreas de zonas verdes; la localización de sumideros y demás elementos de drenaje y los sitios de descarga de los diversos colectores de agua; alteración de la cobertura vegetal [11,15], entre otros importantes condicionantes que reflejan la presencia humana en el territorio y deben considerarse, preferiblemente de manera cuantitativa en un modelo de susceptibilidad.

En general se distinguen cuatro principales métodos utilizados en la evaluación y confección de mapas de susceptibilidad del terreno a producir roturas de las laderas [8,16,14]: 1) determinísticos, 2) heurísticos, 3) estadístico-probabilísticos y 4) geomorfológicos, como se condensa en la Tabla 1.

Entre los métodos más conocidos están los modelos multivariantes. Estos analizan la interacción y dependencia de un conjunto de factores concurrentes, sobre la generación de deslizamientos. Los métodos más utilizados son la regresión múltiple y análisis discriminante [7,10,3,1]. El resultado es una serie de funciones basadas en la presencia/ausencia de deslizamientos que son combinación lineal de los factores de mayor significación estadística en la definición de la inestabilidad y consecuente generación de los movimientos de ladera.

Tabla 1.
Descripción de los diferentes tipos de análisis

Tipo de análisis	Característica principal
Distribución de procesos	Mapeo directo de los rasgos asociados a los procesos. El mapa generado presenta información de los sitios donde han ocurrido procesos de remoción
Cualitativo	Es un método directo o semidirecto. A partir de un mapa geomorfológico y de la experiencia del investigador, se llega a un mapa de zonificación, o bien, se basa en la combinación de mapas temáticos (por ejemplo, pendiente litología, forma de la ladera) usando algún SIG y mediante reglas de decisión a partir del criterio del investigador
Estadístico	Es un método indirecto que se basa en análisis estadísticos para estimar zonas donde pueden ocurrir procesos
Determinístico	Es un método indirecto basado en la combinación de mapas y el cálculo de la estabilidad de laderas a partir de un principio físico
Frecuencia de procesos	En este tipo de métodos indirectos, los registros de precipitación sismicidad y modelos hidrológicos se correlacionan con los datos de los procesos de remoción en masa históricos para obtener una zonificación

Fuente: Elaboración propia, a partir de [4]

2. Análisis discriminante

El análisis discriminante, trata de dividir las laderas o unidades del terreno en dos poblaciones (estable e inestable) utilizando un conjunto de parámetros (variables) característicos de las mismas (por ejemplo, pendiente, litología, orientación, rugosidad, cubrimiento vegetal, entre otras variables disponibles). Esta técnica persigue la separación óptima de las dos poblaciones, minimizando la clasificación errónea de las laderas previamente identificadas como estables o inestables. Las variables independientes seleccionadas se combinan de forma lineal y la función discriminante adopta la forma: $D = d1V1 + \dots + dnVn$.

Donde V_i son las variables independientes de mayor significación estadística, D_i son los coeficientes de clasificación estimados y D es el valor discriminante de la función.

Actualmente los sistemas de información geográfica SIG facilitan el manejo de grandes volúmenes de información y la aplicación de técnicas cuantitativas en la evaluación de la susceptibilidad [2,5,12]. Mediante los SIG, los factores analizados pueden obtenerse de forma automática, en algunos casos con un menor costo, gracias a los modelos digitales de elevaciones, permitiendo ser almacenados y analizados de forma digital.

En cuanto a metodologías ampliamente aceptadas, los métodos disponibles se dividen en métodos de tipo data-driven, utilizados para evaluación de susceptibilidad a deslizamientos [4] como se presentan en la Tabla 2, los cuales permiten, a partir de la combinación de factores que han generado deslizamientos en el pasado, y tratamientos de tipo estadístico, predecir de manera cuantitativa para áreas aún no afectadas que reúnan condiciones similares. La salida de estos métodos puede expresarse en términos de probabilidad.

Tabla 2.
Métodos recomendados para evaluación de susceptibilidad a deslizamientos de tipo data-driven [4]

Tipo	Método	Referencia
Métodos estadísticos bivariados	Likelihood ratio model (LRM)	Lee (2005)
	Information Value Method	Yin and Yan (1988)
	Weights of evidence modelling	Van Westen (1993) Bonham-Carter (1994) Suzen and Doyuran (2004)
	Favourability functions	Chung and Fabbri (1993) Luzi (2004)
Métodos estadísticos multivariados	Discriminant analysis	Carrara (1983) Gorsevski et al. (2000)
	Logistic regression	Ohlmacher and Davis (2003) Gorsevski et al (2006)
Métodos de caja negra	Artificial Neural Networks	Lee et al (2004) Ermini et al (2005) Kanungo et al. (2006)

Fuente: [4]

3. Aplicación del análisis discriminante: susceptibilidad a deslizamientos en una cuenca hidrográfica

En aras de proporcionar una metodología paso a paso sobre la aplicación del método Análisis Discriminante, apto para estudios de tipo regional, a continuación, se presenta la aplicación de la técnica al área de la Cuenca del río Chinchiná, en el Departamento de Caldas.

a) *Inventario de deslizamientos*

Se cuenta con un set total bruto de 893 puntos georreferenciados donde han ocurrido eventos de deslizamiento previamente, obtenidos a partir de la revisión detallada de registros, la base de datos SIMMA [13]. La Fig.1 muestra los puntos disponibles.

Desde el punto de vista numérico, extraer modelos a partir del inventario disponible, con concentración fuerte de datos en el área urbana, podría generar problemas de sesgo, los cuales llevarían a obtener modelos inapropiados. Adicionalmente, se busca un modelo aplicable a la zona rural de la cuenca; por ello se ha filtrado el conjunto de datos, descartando los registros que se encuentran dentro del perímetro urbano de los dos municipios principales de la cuenca: Manizales y Villamaría, para reducir la marcada influencia que se tiene hacia las áreas urbanas. El número total de puntos considerados, para la totalidad de la cuenca, considerando su situación rural, es de 234 eventos.

b) *Factores de propensividad*

A partir de la información disponible, se han seleccionado 17 factores potenciales de tres grandes grupos: 1. Distancia a elementos clave; 2. Suelos/Geología; 3. Factores derivados de la topografía. Los factores así obtenidos se explican brevemente en la Tabla 3 y su representación gráfica se presenta de la Fig.2 a la Fig.18.

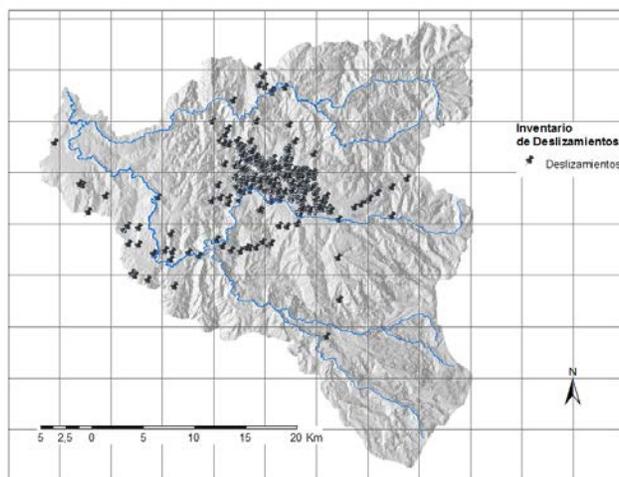


Figura 1. Inventario de deslizamientos previos en la cuenca del río Chinchiná.

Fuente: Elaboración propia

Tabla 3. Factores causales de deslizamiento en la Cuenca del río Chinchiná.

Factor Causal	Tipo	Descripción	
cobe	Mapa de cobertura	Suelos/Geología	Mapa reclasificado de cobertura del terreno
mpar	Material parental (Geología)	Suelos/Geología	Mapa reclasificado de material parental en la cuenca

Factor Causal	Tipo	Descripción	
geom	Procesos geomorfológicos	Suelos/Geología	Mapa reclasificado de procesos geomorfológicos en la cuenca
dem	Modelo digital de elevaciones	Derivados de la topografía	Elevación del terreno en metros sobre el nivel del mar
slop	Pendiente	Derivados de la topografía	Pendiente de cada píxel de terreno en grados
se2p	Pendiente transformada	Derivados de la topografía	Se trata de una transformación senoidal de la pendiente $se2p = \text{sen}(2 \times \text{pend})$. El supuesto es que para taludes, se llega a un "ángulo límite" a partir del cual, si hay estabilidad, se considera que el material es suficientemente resistente y por lo tanto la propensión a deslizarse empieza a reducirse
aspe	Orientación de las laderas	Derivados de la topografía	Orientación de cada celda de terreno respecto al norte geográfico
curv	Curvatura de la superficie	Derivados de la topografía	Curvatura de la superficie. Un valor positivo indica que la superficie es convexa, mientras que uno negativo indica que es cóncava
prft	Curvatura en el sentido perpendicular a la pendiente	Derivados de la topografía	Curvatura (en un plano) perpendicular al sentido de la pendiente de cada píxel
prfl	Curvatura en el sentido de la pendiente	Derivados de la topografía	Es la curvatura del perfil longitudinal del terreno en cada píxel
acca	Área de cuenca vertiente	Derivados de la topografía	Para un píxel dado, es la sumatoria de todos los píxeles cuyas aguas drenan hacia él. (similar al concepto de área aferente)
lcca	Longitud de flujo	Derivados de la topografía	Para un píxel dado, es la distancia máxima que podría recorrer una gota de agua que drenara hacia él
inso	Insolación incidente en el terreno	Derivados de la topografía	Es el promedio de los mapas de insolación para cada hora del día; se calcula según la posición del sol y varía de 0 o nula insolación hasta 255 o máxima insolación
dfal	Distancia a fallas y lineamientos	Distancia a elementos clave	Distancia euclídiana desde cada punto del terreno hasta la falla geológica más cercana
dvia	Distancia a vías	Distancia a	Distancia euclídiana

Factor Causal	Tipo	Descripción
y carreteras	elementos clave	desde cada punto del terreno hasta la vía más cercana
dprob	Distancia a centros poblados	Distancia a elementos clave desde cada punto del terreno hasta el perímetro de la población o asentamiento más cercano
drio	Distancia a ríos y cauces	Distancia euclidiana desde cada punto del terreno hasta el cauce, quebrada o río más cercano

Fuente: Elaboración propia

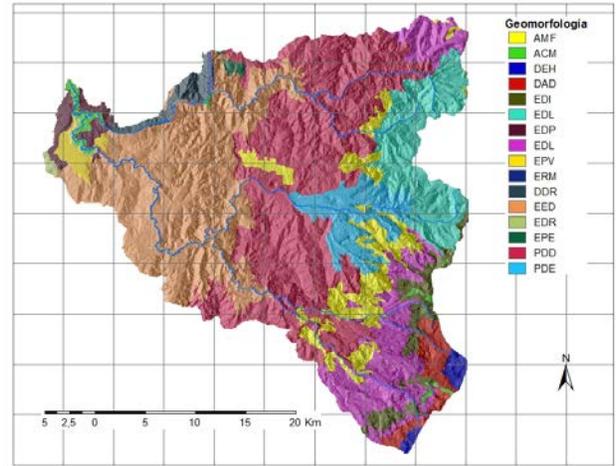


Figura 4. Mapa de procesos geomorfológicos.
Fuente: Elaboración propia

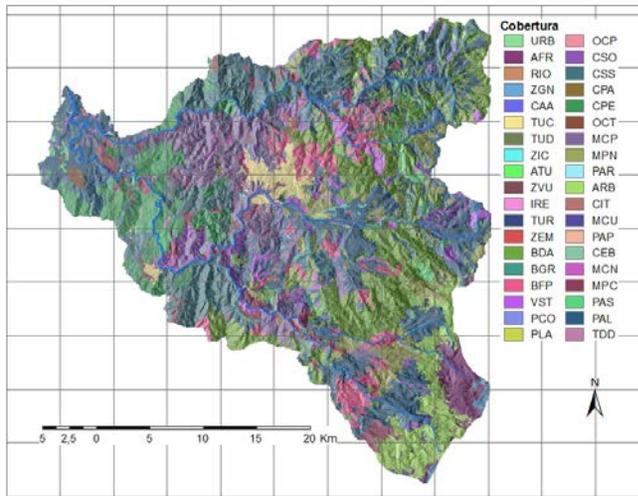


Figura 2. Mapa de coberturas.
Fuente: Elaboración propia

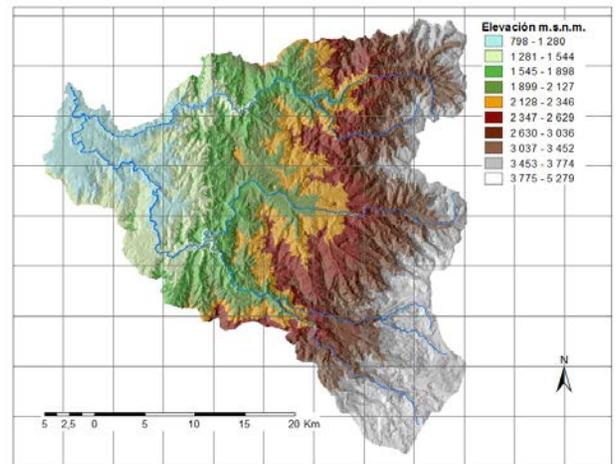


Figura 5. Modelo digital de elevaciones.
Fuente: Elaboración propia

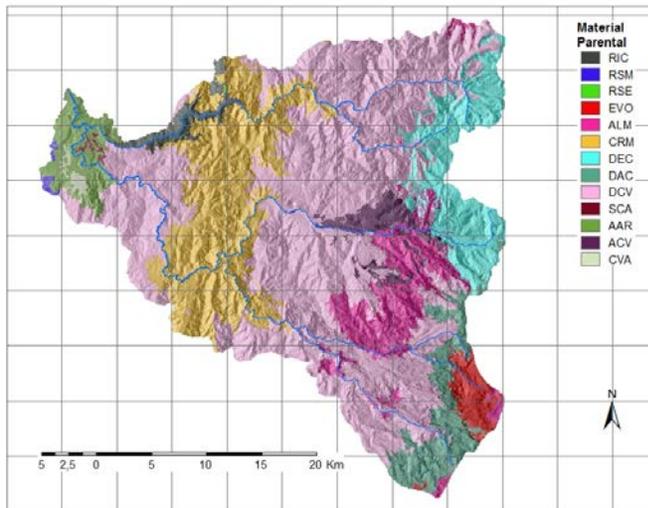


Figura 3. Mapa de características geológicas.
Fuente: Elaboración propia

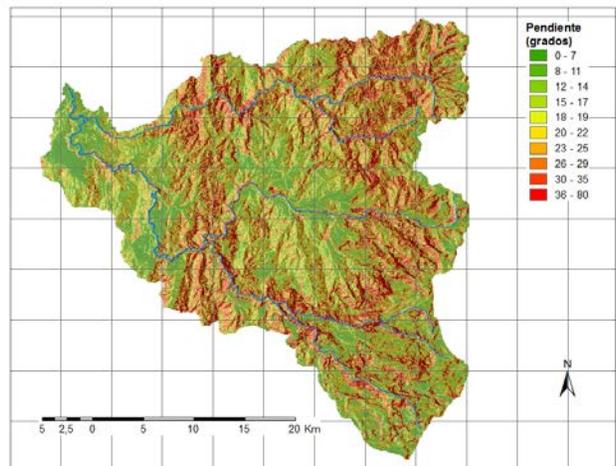


Figura 6. Pendiente del terreno.
Fuente: Elaboración propia

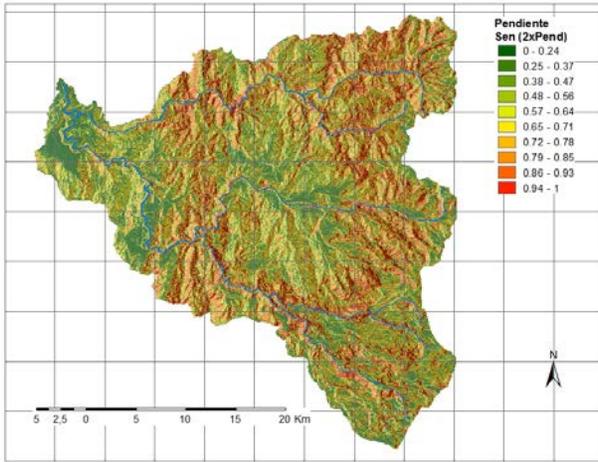


Figura 7. Pendiente transformada del terreno.
Fuente: Elaboración propia

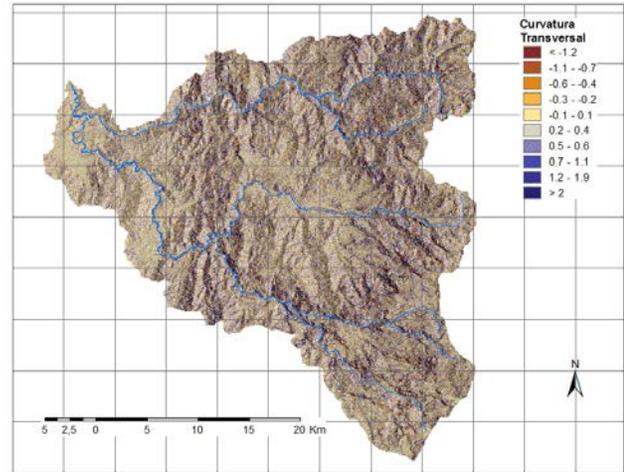


Figura 10. Curvatura transversal del terreno.
Fuente: Elaboración propia

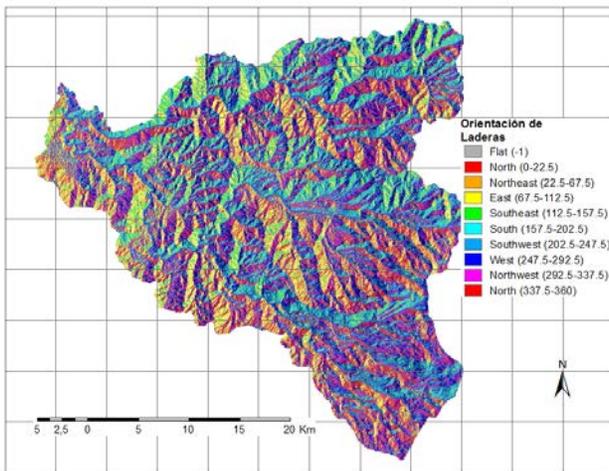


Figura 8. Orientación de las laderas.
Fuente: Elaboración propia

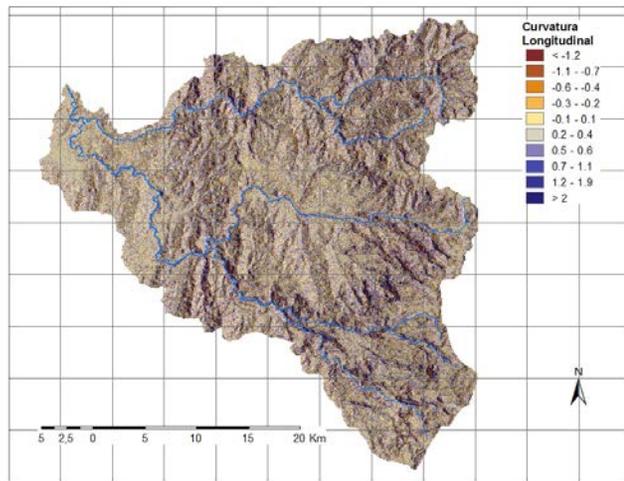


Figura 11. Curvatura longitudinal del terreno.
Fuente: Elaboración propia

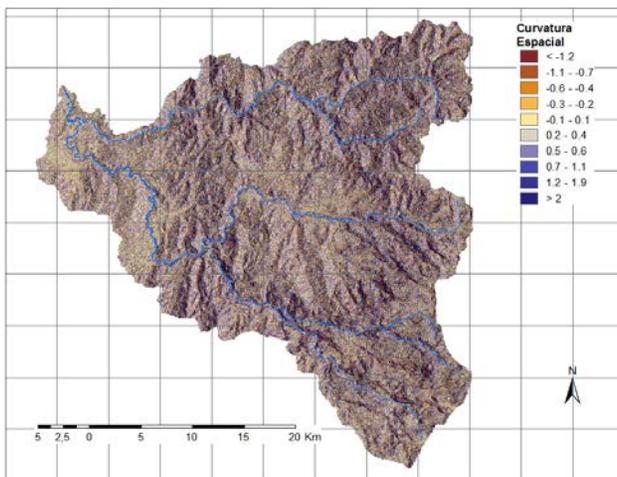


Figura 9. Curvatura del terreno.
Fuente: Elaboración propia

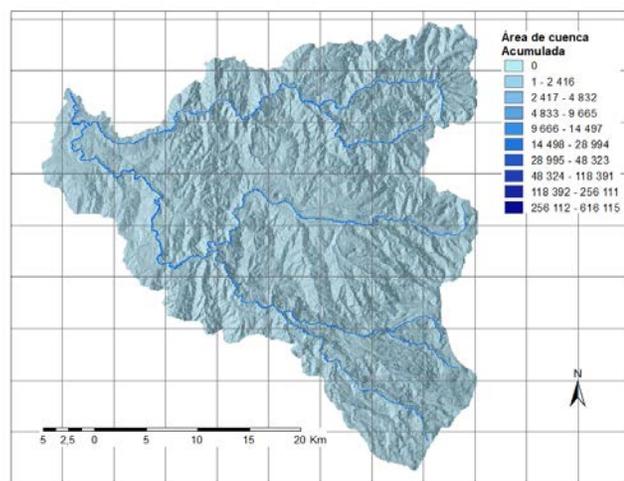


Figura 12. Área de cuenca vertiente.
Fuente: Elaboración propia

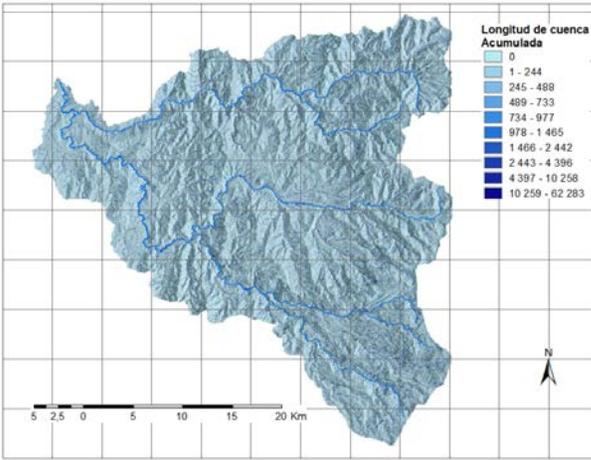


Figura 13. Longitud de cuenca vertiente.
Fuente: Elaboración propia

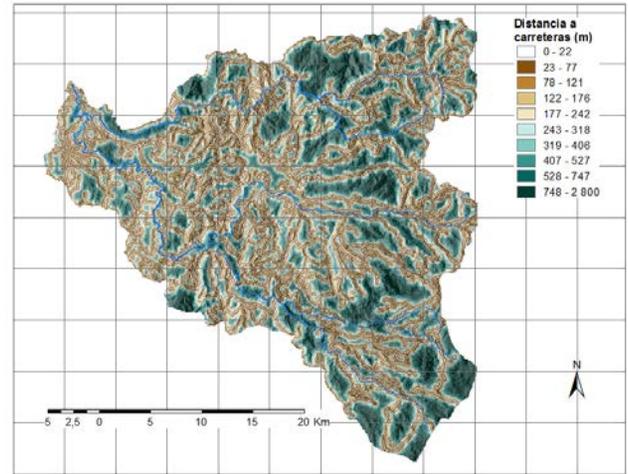


Figura 16. Distancia a vías.
Fuente: Elaboración propia

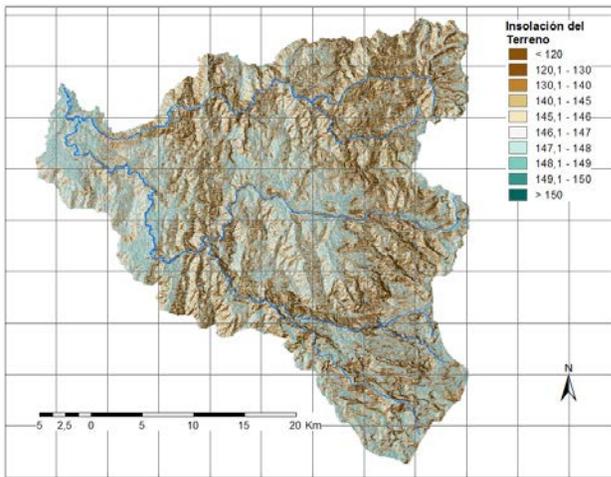


Figura 14. Insolación del terreno.
Fuente: Elaboración propia

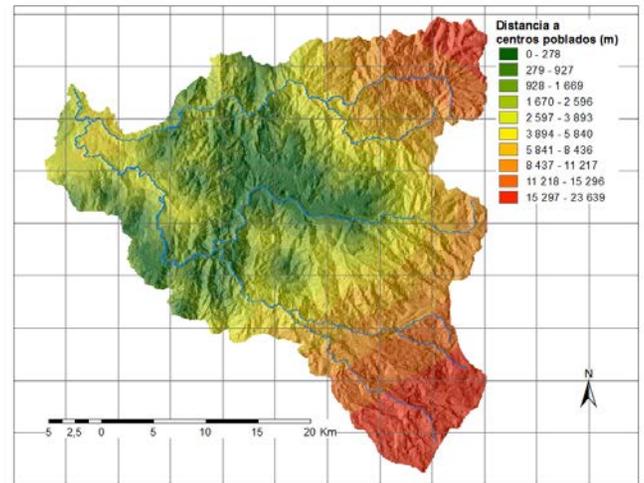


Figura 17. Distancia a centros poblados.
Fuente: Elaboración propia

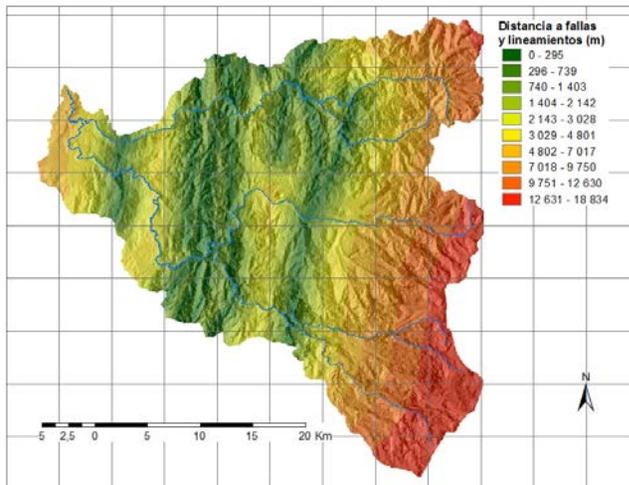


Figura 15. Distancia a fallas y lineamientos.
Fuente: Elaboración propia

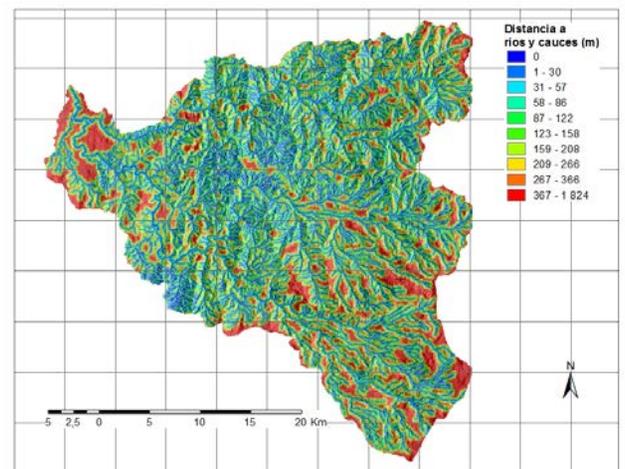


Figura 18. Distancia a ríos y cauces.
Fuente: Elaboración propia

c) Muestra de análisis

Con la información disponible se conforma una muestra de datos de 234 eventos de deslizamiento, que constituyen uno de los dos grupos sobre los que se hará el análisis. El otro grupo está constituido por otros 234 puntos elegidos aleatoriamente en el área donde no hay eventos registrados. En total se tienen 468 puntos que conforman la muestra de análisis constituida en dos grupos: estable ($desl=0$) e inestable ($desl=1$).

d) Función discriminante de susceptibilidad a deslizamientos

Es posible elaborar varios modelos a partir de diferentes combinaciones de variables. Un camino razonable es utilizar inicialmente todas las variables y generar un primer modelo para posteriormente valorar la exclusión de algunas de ellas en modelos posteriores. A continuación, se describen los pasos para construir un primer modelo por este método.

Un paso preliminar es hacer un análisis de correlaciones. Con el propósito de remover variables redundantes (altamente correlacionadas) de los datos, se calcula la matriz

de correlaciones que aparece en la Tabla 4 La matriz ofrece una visión de la interrelación entre las variables de entrada.

Como puede notarse, existen algunos factores que presentan altas correlaciones; en este caso debe analizarse por pares qué variable es conveniente conservar y cual debe descartarse. Las variables con más fuerte correlación se muestran en la Tabla 5.

El test de igualdad de medias entre grupos mide el potencial de cada variable independiente antes de que el modelo sea creado. Cada test muestra el resultado de un análisis one-way ANOVA para la variable independiente usando la variable de agrupación como factor. Si el valor de significancia es mayor a 0.10, la variable probablemente no contribuye al modelo. Los valores Lambda de Wilks son otra medida del potencial de la variable. Valores pequeños indican que la variable es mejor discriminando entre grupos. Los valores de Lambda de Wilks y el F de Fisher para cada variable se presentan en la Tabla 6.

Tabla 4. Matriz de correlaciones.

	asp	dvia	curv	prft	prfl	dem	dfal	dpob	drío	flac	flle	geom	inso	se2p	slop	mpar	cobe
asp	1,00	0,25	0,10	0,14	-0,05	0,24	-0,04	0,03	0,03	0,07	0,04	0,08	0,01	0,23	0,19	0,14	-0,01
dvia	0,25	1,00	-0,10	-0,04	0,13	0,39	0,29	0,17	0,18	-0,05	-0,05	-0,24	0,04	0,08	0,07	0,46	-0,03
curv	0,10	-0,10	1,00	0,86	-0,91	-0,03	-0,15	0,05	-0,03	-0,07	-0,09	-0,10	0,05	-0,19	-0,17	-0,19	-0,06
prft	0,14	-0,04	0,86	1,00	-0,58	-0,04	-0,17	0,03	-0,09	0,00	-0,03	-0,04	0,09	-0,24	-0,26	-0,15	-0,03
prfl	-0,05	0,13	-0,91	-0,58	1,00	0,01	0,10	-0,06	-0,03	0,11	0,13	0,13	-0,01	0,11	0,06	0,19	0,07
dem	0,24	0,39	-0,03	-0,04	0,01	1,00	0,48	0,44	0,26	-0,05	-0,07	-0,14	0,08	0,11	0,07	0,40	0,42
dfal	-0,04	0,29	-0,15	-0,17	0,10	0,48	1,00	0,53	0,28	-0,04	-0,04	-0,41	0,06	-0,04	-0,04	0,33	0,33
dpob	0,03	0,17	0,05	0,03	-0,06	0,44	0,53	1,00	0,21	-0,06	-0,06	-0,58	0,08	-0,08	-0,07	0,06	0,18
drío	0,03	0,18	-0,03	-0,09	-0,03	0,26	0,28	0,21	1,00	-0,13	-0,14	-0,19	0,20	-0,03	-0,04	0,23	0,15
flac	0,07	-0,05	-0,07	0,00	0,11	-0,05	-0,04	-0,06	-0,13	1,00	0,99	0,00	0,07	-0,12	-0,11	0,04	0,04
flle	0,04	-0,05	-0,09	-0,03	0,13	-0,07	-0,04	-0,06	-0,14	0,99	1,00	0,00	0,07	-0,13	-0,11	0,07	0,03
geom	0,08	-0,24	-0,10	-0,04	0,13	-0,14	-0,41	-0,58	-0,19	0,00	0,00	1,00	-0,06	0,15	0,11	0,17	0,07
inso	0,01	0,04	0,05	0,09	-0,01	0,08	0,06	0,08	0,20	0,07	0,07	-0,06	1,00	-0,71	-0,79	0,13	0,13
se2p	0,23	0,08	-0,19	-0,24	0,11	0,11	-0,04	-0,08	-0,03	-0,12	-0,13	0,15	-0,71	1,00	0,97	0,02	0,01
slop	0,19	0,07	-0,17	-0,26	0,06	0,07	-0,04	-0,07	-0,04	-0,11	-0,11	0,11	-0,79	0,97	1,00	0,00	-0,02
mpar	0,14	0,46	-0,19	-0,15	0,19	0,40	0,33	0,06	0,23	0,04	0,07	0,17	0,13	0,02	0,00	1,00	0,37
cobe	-0,01	-0,03	-0,06	-0,03	0,07	0,42	0,33	0,18	0,15	0,04	0,03	0,07	0,13	0,01	-0,02	0,37	1,00

Fuente: Elaboración propia

Tabla 5. Variables con correlaciones altas.

Variables	Correlación
curv - prft	0.864
curv - prfl	-0.909
prft - prfl	-0.575
dfal - dpob	0.528
dpob - geom	-0.584
flac - flle	0.992
inso - se2p	-0.706
inso - slope	-0.790
se2p - slope	0.973

Fuente: Elaboración propia

Los gráficos correspondientes a los valores de la Tabla 6 permiten visualizar de mejor manera el grado de influencia de los factores de entrada en el modelo. En el caso del F de Fisher (Fig.19), para valores bajos, empiezan a perder influencia los factores; el caso del Lambda de Wilks (Fig.20) es en sentido inverso, a partir de cierto valor (alto) los factores ya no aportan. El punto de corte se establece en valores de significancia por encima de 0.10, en este caso, a partir de la variable *prft*.

Tabla 6. Test de igualdad de Media entre grupos.

var	Test de igualdad de media entre grupos				
	Wilks' Lambda	F	df1	df2	sig
dpob	0.703	197.289	1	466	0.000
dfal	0.823	100.087	1	466	0.000
dvia	0.839	89.621	1	466	0.000
geom	0.859	76.352	1	466	0.000
dem	0.898	52.375	1	466	0.000
drío	0.954	22.299	1	466	0.000
cobe	0.980	9.640	1	466	0.002
inso	0.991	4.450	1	466	0.035
slop	0.991	4.409	1	466	0.036
mpar	0.993	3.165	1	466	0.076
prft	0.995	2.553	1	466	0.111
se2p	0.996	2.061	1	466	0.152
curv	0.996	2.047	1	466	0.153
flac	0.996	1.926	1	466	0.166
flle	0.997	1.360	1	466	0.244
prfl	0.997	1.181	1	466	0.278
asp	1.000	0.223	1	466	0.637

Fuente: Elaboración propia

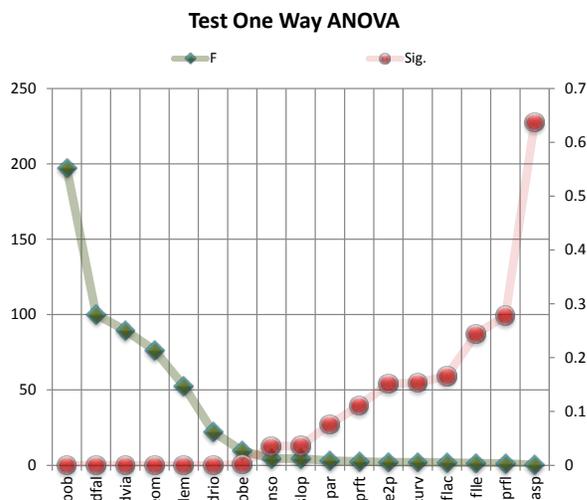


Figura 19. Test de igualdad de medias entre grupos (F de Fisher).
Fuente: Elaboración propia

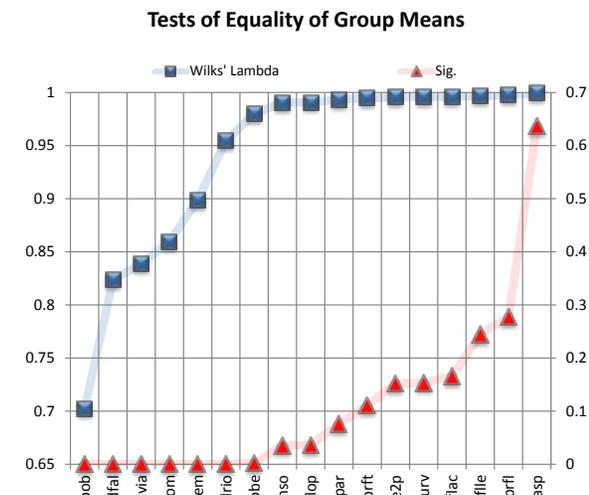


Figura 20. Test de igualdad de medias entre grupos (Lambda de Wilks).
Fuente: Elaboración propia

El análisis de varianza lleva a descartar las variables *asp*, *prfl*, *file*, *flac*, *curv*, *se2p*, *prft*, que aparecen en color azul en la Tabla 6, dado su bajo poder discriminante.

La correlación de los predictores o colinealidad se evalúa usando la matriz de covarianza entre grupos (Tabla 7). Una inspección de la correlación de las covarianzas muestra las correlaciones altas entre los predictores *dem*, *dfal* y *dpob*, y *slop* e *inso*.

Ante esta situación, es conveniente hacer un razonamiento: Considérese por ejemplo una variable de cobertura vegetal y otra de elevación. La vegetación puede variar sistemáticamente según la elevación. Un supuesto de independencia condicional de estas dos capas no es razonable. En ese caso debe valorarse cómo son afectados los patrones de predicción si el supuesto de independencia condicional es violado [6].

Tabla 7.
Correlación de covarianzas.

	dvia	dem	dfal	dpob	drio	geom	inso	slop	mpar	cobe
dvia	1,00	0,40	0,32	0,38	0,03	-0,30	-0,09	0,14	0,03	-0,13
dem	0,40	1,00	0,80	0,75	0,13	-0,51	-0,08	0,12	-0,01	-0,01
dfal	0,32	0,80	1,00	0,82	0,22	-0,64	-0,08	0,07	-0,02	-0,05
dpob	0,38	0,75	0,82	1,00	0,15	-0,58	-0,05	0,04	-0,10	-0,06
drio	0,03	0,13	0,22	0,15	1,00	-0,14	0,01	-0,02	0,04	0,01
geom	-0,30	-0,51	-0,64	-0,58	-0,14	1,00	0,11	-0,08	0,22	0,20
inso	-0,09	-0,08	-0,08	-0,05	0,01	0,11	1,00	-0,84	0,09	0,13
slop	0,14	0,12	0,07	0,04	-0,02	-0,08	-0,84	1,00	-0,04	-0,08
mpar	0,03	-0,01	-0,02	-0,10	0,04	0,22	0,09	-0,04	1,00	0,38
cobe	-0,13	-0,01	-0,05	-0,06	0,01	0,20	0,13	-0,08	0,38	1,00

Fuente: Elaboración propia

Dado el origen diferente de las variables *dem*, *dpob* y *dfal*, el hecho que ninguna de ellas se deriva de otra, y su alta influencia en el proceso discriminante de celdas estables e inestables (ver tablas y gráficas anteriores), se decide continuar el análisis con todas ellas, para posteriormente elaborar modelos alternativos descartando algunas y comparar los resultados. En el caso de *slop* e *inso*, se opta por probar con cada una también.

El test Box's M es utilizado para probar la homogeneidad de las matrices de covarianza, chequea el supuesto de igualdad de covarianzas entre grupos. Los log determinantes (Tabla 8) son una medida de la variabilidad de los grupos. Valores altos corresponden a grupos más variables. Diferencias grandes en los valores indican que los grupos tienen diferentes matrices de covarianza.

Tabla 8.
Chequeo de homogeneidad de covarianzas.

Desl	Log determinantes	
	Rank	Log determinant
0	10	87,7
1	10	76,48
Pooled within-groups	10	84,72

Fuente: Elaboración propia

Cuando el valor Box's M (Tabla 9) es significativa (*Sig.*=0), se debe chequear el modelo discriminante mediante matrices separadas para verificar si se obtienen valores de clasificación diferentes. En este caso, los modelos con matrices separadas produjeron resultados de clasificación muy similares o en algunos casos inferiores a los obtenidos con un solo grupo.

Tabla 9.
Test Box M.

Box's M		1227,832
F	Approx.	21,824
	df1	55
	df2	701264,161
	Sig.	0,000

Fuente: Elaboración propia

Tabla 10.
Lambda de Wilks para el modelo general. Fuente: Elaboración propia

Lambda de Wilks			
Lambda	Chi cuadrado	df	Sig.
0,636	208,370	10	0,000

Fuente: Elaboración propia

Tabla 11.
Valores propios y correlación para la función discriminante.

Eigenvalues			
Eigenvalue	% de varianza	Acumulado	Correlación canónica
0,571	100	100	0,603

Fuente: Elaboración propia

Tabla 12.
Matriz de estructura del modelo.

Matriz de estructura	
dpob	0,861
dfal	0,613
dvia	0,580
geom	-0,535
dem	0,445
drio	0,289
cobe	-0,190
inso	-0,129
slop	0,129
mpar	-0,109

Fuente: Elaboración propia

El Lambda de Wilks (Tabla 10) es una medida de la manera en que cada función separa los casos en grupos. Es igual a la proporción de varianza total en los coeficientes discriminantes no explicada por diferencias entre los grupos. Pequeños valores de lambda indican gran poder discriminante de la función. El Chi-cuadrado asociado chequea la hipótesis de que las medias de las funciones son iguales entre grupos. Valores pequeños de Sig. Indican que la función discriminante tiene resultados mejores que el azar en la separación entre grupos.

La tabla de valores propios (Tabla 11) proporciona información sobre la eficacia de la función discriminante. Cuando existen dos grupos, la correlación canónica es la medida más útil en la tabla, y es equivalente a la correlación de Pearson entre los coeficientes discriminantes y los grupos.

Hay varias tablas que evalúan la contribución de cada variable al modelo, incluyendo los tests de igualdad de medias grupales, los coeficientes de la función discriminante y la matriz de estructura.

La matriz de estructura (Tabla 12) muestra la correlación de cada predictor con la función discriminante y constituye una prueba de la importancia que cada variable tiene dentro

del modelo. El orden en esta matriz es igual al sugerido por el test de igualdad de medias grupales y es diferente del que aparece en la tabla de coeficientes estandarizados que constituye la ecuación del modelo lineal.

Los coeficientes estandarizados (Tabla 13) permiten comparar variables medidas en diferentes escalas. Coeficientes con valores absolutos altos corresponden a variables con alto poder discriminante. Puede observarse que dentro de las variables más influyentes individualmente se encuentran *dpob*, *dem*, *dvia*, *slope* y *drio*.

El modelo lineal obtenido con la función discriminante se ensambla con los coeficientes canónicos de la Tabla 14.

La Tabla 15 resume los resultados del modelo inicial en cuanto a capacidad de clasificación de casos. Puede leerse que el modelo acierta en la clasificación del 65% de las celdas estables, clasifica correctamente el 95.7% de las celdas inestables, y en total, tiene un acierto del 80.3% de casos.

Tabla 13.
Coeficientes canónicos estandarizados. Fuente: Elaboración propia

Coeficientes	
dvia	0,350
dem	-0,560
dfal	-0,068
dpob	0,140
drio	0,188
geom	-0,052
inso	0,025
slop	0,120
mpar	0,022
cobe	-0,079

Fuente: Elaboración propia

Tabla 14.
Coeficientes de la función discriminante. Fuente: Elaboración propia

Coeficientes	
dvia	0,001
dem	-0,001
dfal	0,000
dpob	0,000
drio	0,001
geom	-0,009
inso	0,003
slop	0,011
mpar	0,006
cobe	-0,006
constante	0,029

Fuente: Elaboración propia

Tabla 15.
Resultados de clasificación del modelo inicial.

Función discriminante	Resultados de la clasificación			Total
	0	1		
Count	0	152	82	234
	1	10	224	234
%	0	65,0	35,0	100
	1	4,3	95,7	100

Fuente: Elaboración propia

La ecuación lineal obtenida se ha aplicado a toda el área bajo estudio y sus valores se han normalizado entre 0 y 1. Este procedimiento hace que la asignación posterior de categorías de susceptibilidad sea más sencilla. El mapa resultado de esta

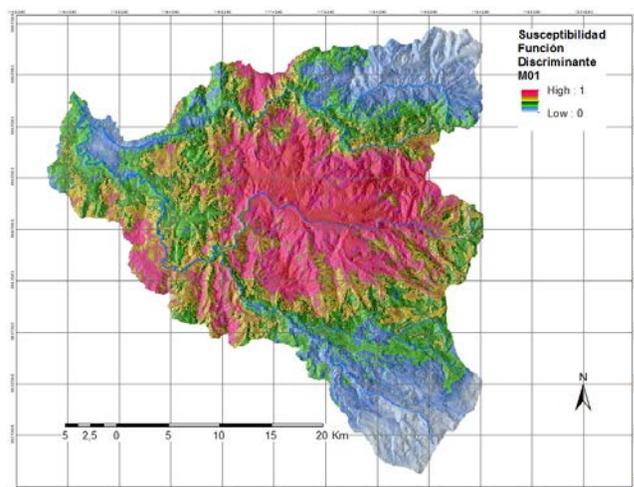


Figura 21. Susceptibilidad mediante una función discriminante.
Fuente: Elaboración propia

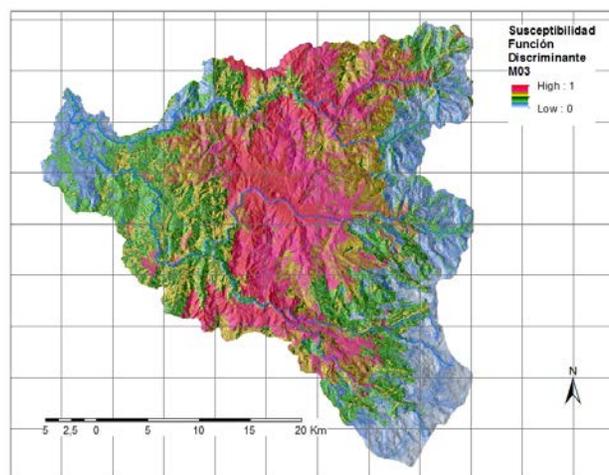


Figura 22. Modelo función discriminante MOD3.
Fuente: Elaboración propia

clasificación se presenta a continuación (Fig.21). Se trata de un mapa ráster con valores entre cero y uno que refleja el grado relativo de susceptibilidad de cada celda del terreno.

e) Diferentes modelos discriminantes probados

Se han elaborado siete modelos diferentes cambiando las variables incluidas en cada caso y se ha probado el desempeño en la clasificación de los datos. La configuración de cada uno de los modelos probados se presenta en la Tabla 16. El modelo MOD1 ya se presentó junto a la descripción del procedimiento en la Fig.21.

Los tres mejores modelos obtenidos fueron MOD1, MOD3 (Fig.22) Y MOD5, exhiben parámetros de rendimiento similares como se observa en la Tabla 17. El mejor modelo en capacidad de clasificación y sensibilidad es MOD3.

f) Rendimiento de los modelos obtenidos

En cuanto a sus capacidades de clasificación, en la Fig.23 se presentan las curvas ROC obtenidas para cada modelo. La curva es un dibujo de la probabilidad de tener un verdadero

positivo (un valor correctamente predicho) versus la probabilidad de un falso positivo (un valor incorrectamente predicho). Un modelo ideal muestra un valor del área bajo la curva AUC cercano a 1, mientras un valor cercano a 0,5 evidencia un inadecuado modelo.

Los valores del área bajo la curva para el caso de celdas inestables (deslizamientos), se muestran en la Tabla 17 bajo el nombre AUC_D; de allí se desprende que el modelo con mejor capacidad de clasificación es DIS03, con un valor AUC de 0.883, seguido muy de cerca por los otros dos modelos. En cuanto a clasificación de celdas estables, los valores AUC_E están cercanos a 0.5 en los tres casos, siendo deseables valores tendientes a cero. Sin embargo, para efectos de elección de un modelo de susceptibilidad a deslizamientos, es preferible aquel que acierte en clasificar correctamente celdas inestables que estables.

Tabla 16. Diferentes modelos de función discriminante probados.

VAR	MOD 1	MOD 2	MOD 3	MOD 4	MOD 5	MOD 6	MOD 7
dpob	X	X			X	X	X
dfal	X		X	X		X	X
dvia	X	X	X	X	X	X	X
geo	X		X	X			
m							
dem	X	X	X	X	X	X	X
drio	X	X	X	X	X	X	X
cobe	X	X	X	X	X	X	X
inso	X	X	X			X	
slop	X			X	X		X
mpar	X	X	X	X	X	X	X
prft							
se2p							
curv							
flac							
file							
prfl							
aspe							

Fuente: Elaboración propia

Tabla 17. Parámetros de rendimiento modelos discriminantes.

Parámetro	MOD1	MOD3	MOD5
Área bajo curva (D)	0,882	0,883	0,881
Área bajo curva (E)	0,487	0,487	0,487
fn	5	3	5
tp	229	231	229
tn	116	116	118
fp	118	118	116
Eficiencia/Precisión	0,737	0,741	0,741
Sensibilidad	0,979	0,987	0,979
Especificidad	0,504	0,504	0,496
Gilbert Score	0,651	0,656	0,654
Peirce Skill Score	-0,304	-0,171	-0,296
Heidke Skill Score	0,474	0,483	0,483
Odds Ratio Skill Score	0,957	0,974	0,958

Fuente: Elaboración propia

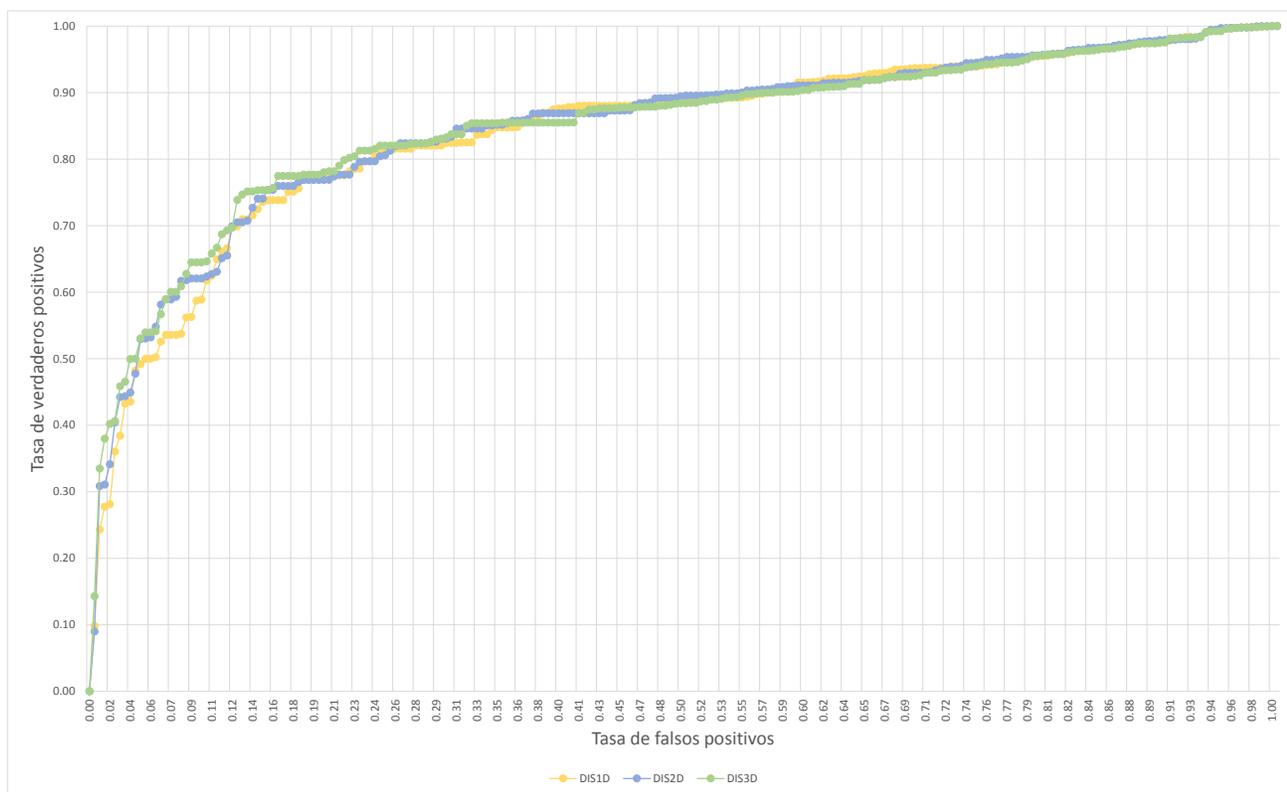


Figura 23. Curvas de rendimiento modelos discriminantes.
Fuente: Elaboración propia

4. Conclusiones

Cuando se habla de planificación del territorio, se requiere una valoración o dimensionamiento como paso inicial para fines de gestión. Aunque en la literatura abundan enfoques cualitativos para valoración, y tradicionalmente se han obtenido modelos desde este enfoque, se prefiere el uso de análisis cuantitativos, pues permiten eliminar la subjetividad de los estudios, y hacen posible la comparación directa del nivel de susceptibilidad y amenaza entre áreas. De otro lado, es la única manera posible de hacer estimaciones posteriores de amenaza, en términos de probabilidad de ocurrencia, y de riesgo, en términos de niveles de pérdidas.

Un aspecto diferenciador de la metodología expuesta es la incorporación de factores de contexto, que reflejan la influencia antrópica dentro del territorio, aspecto fundamental cuando se aborda el tema de riesgo dado su carácter dinámico de construcción social.

La metodología planteada pretende interrelacionar variables para obtener zonificaciones y en ningún caso construir modelos que obedezcan a principios físicos, los cuales requieren otro enfoque. Se ha presentado una de las métricas posibles para obtener modelos de susceptibilidad en términos cuantitativos, lo cual constituye el primer paso en un análisis de riesgo por deslizamientos. Posteriormente deberá obtenerse una estimación de los umbrales detonantes y con ello un cálculo de amenaza, en términos de probabilidades de excedencia del umbral, para después analizar el modelo de exposición y vulnerabilidad y llegar al

cálculo de riesgo en términos cuantitativos de pérdidas.

Aunque existen otras metodologías más sofisticadas para obtener modelos de susceptibilidad, los resultados obtenidos demuestran que siguiendo una serie de pasos lógicos, es posible llegar a zonificaciones aceptables para escalas de análisis de tipo regional, con precisiones del orden de 75% y sensibilidades altas.

Bibliografía

- [1] Baeza, C. and J., Corominas. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surface Processes and Landforms* 26(12), pp. 1251-1263. DOI: 10.1002/esp.263. 2001.
- [2] Bianchini, S., Del Soldato, M., Solari, L., Nolesini, T., Pratesi, F. and Moretti, S., Badland susceptibility assessment in Volterra municipality (Tuscany, Italy) by means of GIS and statistical analysis. *Environmental Earth Sciences*. 75(889), pp. 1-14, 2016. DOI: 10.1007/s12665-016-5586-5
- [3] Carrara, A., Multivariate models for landslide hazard evaluation. *Mathematical Geology*, 15(3), pp. 403-426, 1983. DOI: 10.1007/BF01031290
- [4] Corominas, J., et al., Recommendations for the quantitative analysis of landslide risk. *Bulletin of Engineering Geology and the Environment* 73(2), pp. 209-263, 2014. DOI: 10.1007/s10064-013-0538-8
- [5] Chacón, J., Irigaray, C. y Fenández, T., Metodología para la cartografía regional de movimientos de ladera y riesgos asociados mediante un sistema de información geográfica. III Simposio Nacional Sobre taludes y laderas inestables. La Coruña. 1, pp. 121-133, 1992.
- [6] Chung, C., Using likelihood ratio functions for modeling the conditional probability of occurrence of future landslides for risk assessment. *Computers Geosciences*, 32(8), pp. 1052-1068, 2006. DOI: 10.1016/j.cageo.2006.02.003

- [7] Jones, F.O., Embody, D.R. and Peterson, W.L., Landslide along the Columbia river valley, Northeastern Washington. Geological Survey Professional Paper, 367, pp. 1-94, 1961.
- [8] Lorentz, J.F., Calijuri, M.L., Marques, E.G. And Baptista, A.C., Multicriteria analysis applied to landslide susceptibility mapping. Natural Hazards. 83, pp. 41-52, 2016. DOI: 10.1007/s11069-016-2300-6
- [9] Myronidis, D., Papageorgiou, C. and Theophanous, S., Landslide susceptibility mapping based on landslide history and analytic hierarchy process (AHP). Natural Hazards. 81, pp.245-263, 2016. DOI: 10.1007/s11069-015-2075-1
- [10] Neuland, H., A prediction model of landslides. Catena, 3, pp. 215-230, 1976.
- [11] Persichillo, M.G., Bordoni, M. and Meisina, C., The role of land use changes in the distribution of shallow landslides. Science of the Total Environment. 574, pp. 924-937, 2017. DOI: 10.1016/j.scitotenv.2016.09.125
- [12] Santacana, N., Baeza, C., Corominas, J., de Paz, A. y Marturià, J., Análisis de la susceptibilidad del terreno a la formación de deslizamientos superficiales mediante el uso de un Sistema de Información Geográfica. Aplicación a La Pobla de Lillet (Pirineo Oriental). En: F.J. Ayala y J. Corominas (eds). Mapas de susceptibilidad a los movimientos de ladera con técnicas SIG. Fundamentos y aplicaciones en España. IGME. Madrid, pp. 55-82, 2003.
- [13] SGC. Sistema de Información de Movimientos en Masa SIMMA. Servicio Geológico Colombiano, 2016.
- [14] Soeters, R. and van Westen, C.J., Slope instability. Recognition, analysis and zonation. In: A.K. Turner and R.L. Schuster (Editors), Landslide: Investigations and mitigation. Special Report 247. Transportation Research Board. National Research Council. National Academy Press., Washington, D.C, pp. 129-177, 1996.
- [15] Suarez, J., Deslizamientos Volumen I: Análisis geotécnico. Publicaciones Universidad Industrial de Santander. Bucaramanga, Colombia. 2009, 582 P.
- [16] Van Westen, C.J., GIS in landslide hazard zonation: a review with examples from the Colombian Andes. In: Price, M.F. and Heywood, D.I. (eds.) Taylor and Francis, London, 1994, pp. 135-165.
- [17] Youssef, A., Pourghasemi, H., El-Haddad, B. and Dhahry, B., Landslide susceptibility maps using different probabilistic and bivariate statistical models and comparison of their performance at Wadi Itwad Basin, Asir Region, Saudi Arabia. Bulletin of Engineering Geology and the Environment, 75, pp. 63-78, 2016. DOI: 10.1007/s10064-015-0734-9

J.P. Londoño-Linares, Es Ing. Civil y MSc. en Medio Ambiente y Desarrollo de la Universidad Nacional de Colombia. Dr. por la Universidad Politécnica de Cataluña, España. Cuenta con 15 años de experiencia como consultor independiente en temas relacionados con medio ambiente, gestión del riesgo y desarrollo territorial, especialmente en cuanto a modelación de procesos físicos mediante métodos numéricos.
ORCID: 0000-0002-0830-2981



UNIVERSIDAD NACIONAL DE COLOMBIA

SEDE MEDELLÍN
FACULTAD DE MINAS

Área Curricular de Medio Ambiente

Oferta de Posgrados

Especialización en Aprovechamiento de
Recursos Hidráulicos
Especialización en Gestión Ambiental
Maestría en Ingeniería Recursos Hidráulicos
Maestría en Medio Ambiente y Desarrollo
Doctorado en Ingeniería - Recursos Hidráulicos
Doctorado Interinstitucional en Ciencias del Mar

Mayor información:

E-mail: acma_med@unal.edu.co
Teléfono: (57-4) 425 5105