

## ESTIMACIÓN DE LA EROSIÓN PRODUCIDA POR LAS CORRIENTES DE DERRUBIOS MEDIANTE TÉCNICAS DE MINERÍA DE DATOS CON DATOS DE CAMPO Y TOPOGRÁFICOS

Clàudia ABANCÓ<sup>1\*</sup>, Marcel HÜRLIMANN<sup>1</sup>

<sup>1</sup> Dpto. Ing. del Terreno, Cartográfica y Geofísica.  
Universitat Politècnica de Catalunya.

### RESUMEN:

*La erosión del lecho es un proceso común de las corrientes de derrubios, y aumenta considerablemente el volumen de éstas. En este estudio se analizan datos geomorfológicos y topográficos obtenidos de 17 torrentes donde se han producido corrientes de derrubios. Sobre la base de datos se aplicaron dos técnicas de minería de datos: regresión lineal múltiple y árbol de decisión, con la intención de desarrollar un modelo para predecir las tasas de erosión. Los modelos obtenidos mediante las dos técnicas fueron validados e implementados en grupos de datos independientes con la finalidad de evaluar su efectividad. El modelo del árbol de decisión mostró más ventajas que la regresión lineal múltiple. Finalmente se propuso un modelo, basado en el árbol de decisión, para estimar la tasa de erosión, que incluye tres factores condicionantes: la disponibilidad de sedimento, el pendiente del lecho y la forma de la sección transversal.*

### 1. INTRODUCCIÓN

Las corrientes de derrubios son movimientos rápidos, formados por una mezcla de agua y material sólido, que ocurren mayoritariamente en torrentes de alta pendiente. Representan un elevado riesgo para la población y las infraestructuras en zonas montañosas. La erosión basal del lecho del canal por donde transcurre el flujo es una característica común de las corrientes de derrubios. Esta incorporación de material suelto no consolidado a lo largo del transcurso de la corriente de derrubios afecta parámetros dinámicos del flujo, que son claves para la evaluación de la peligrosidad, como el volumen o el máximo alcance (Hungry et al., 2005; Guthrie et al., 2009; Berger et al., 2011). Por este motivo, la predicción del volumen de material potencialmente erosionable del canal es una cuestión clave.

A pesar de que existe una notable variedad de métodos para predecir volúmenes de eventos

(p.ej. Rickenmann, 1995; D'Agostino, 1996; Marchi y D'Agostino, 2004) pocos de estos contemplan específicamente la erosión del lecho (Hungar et al., 1984). Es un fenómeno complejo, y todavía existen muchas incertidumbres sobre la forma de definir el proceso de erosión. Aunque existen varios estudios publicados sobre erosión en eventos individuales (p.ej. Scheidl et al., 2008), pocos son los casos en que se presenten resultados obtenidos de varios eventos de una región (Fannin y Rollerson, 1993; Hungar et al., 2005; Gertsch, 2009). Este estudio tiene dos objetivos principales: en primer lugar, analizar los datos de erosión ocurrida durante corrientes de derrubios recientes con la intención de determinar los factores condicionantes del mecanismo de erosión; en segundo lugar, desarrollar una metodología para predecir la erosión producida por las corrientes de derrubios a lo largo del recorrido de las corrientes de derrubios. Esta metodología debería ser fácilmente aplicable.

## **2. BASE DE DATOS Y ZONAS DE ESTUDIO**

Para este estudio se recolectaron datos de la erosión ocurrida en 17 corrientes de derrubios granulares. La adquisición de datos en campo fue complementada con datos adquiridos mediante un Sistema de Información Geográfica (SIG). Los torrentes seleccionados para este estudio se encuentran en los Pirineos y los Alpes (Figura 1), y se han dividido en tramos, en los cuales se ha producido una amplia variedad de tasas de erosión ( $m^3$  material erosionado/m lineal de torrente) (Hungar y Morgenstern, 1984). La base de datos incluye información sobre factores condicionantes y tasas de erosión. Los volúmenes finales de los eventos fueron estimados en base a nuestras propias observaciones y datos de campo o a partir de informes técnicos. Las campañas de campo fueron mayoritariamente llevadas a cabo algunos días o semanas después de los eventos. Sólo en algunos casos la recolección de datos tuvo lugar después de un período de tiempo más largo.

Las estimaciones de las tasas de erosión se basaron en la observación en campo de marcas erosivas acompañadas de la reconstrucción de la disposición del sedimento previamente a la ocurrencia del evento. Para cada evento, se recolectaron dos tipos de datos: 1) observaciones de campo i 2) datos morfométricos derivados de un modelo digital de elevaciones (DEM) con tamaño de píxel entre 1 y 5 metros.

## **3. DEFINICIÓN DE LOS FACTORES CONDICIONANTES**

Los 17 torrentes se dividieron en tramos de acuerdo con las similitudes geomorfológicas y las tasas de erosión. La base de datos final se compone de 110 tramos de torrentes donde ocurrieron corrientes de derrubios. La longitud de los tramos varía entre decenas y más de un millar de metros.

En cada tramo, se han determinado algunos parámetros geomorfológicos, aquí denominados factores condicionantes. Finalmente se seleccionaron 4 factores condicionantes para el desarrollo de la metodología: 1) pendiente media del tramo ( $P_i$ ), 2) disponibilidad de sedimento ( $DS_i$ ), 3) forma de la sección transversal ( $ST_i$ ), y 4) área contributiva ( $AC_i$ ). Dichos factores pueden dividirse en dos grupos dependiendo en el método de obtención: 1) tres factores de campo ( $P_i$ ,  $DS_i$ ,  $ST_i$ ) i 2) un factor topográfico ( $AC_i$ ). Los factores se pueden dividir también en función de si son numéricos ( $P_i$ ,  $AC_i$ ); o categóricos ( $DS_i$ ,  $ST_i$ ). La

selección de los factores se basó en: 1) la simplicidad de su recolección, 2) las observaciones en campo que prueban su influencia en el proceso de la erosión, y 3) experiencias obtenidas desde estudios previos (p.ej. Hung et al., 2005).

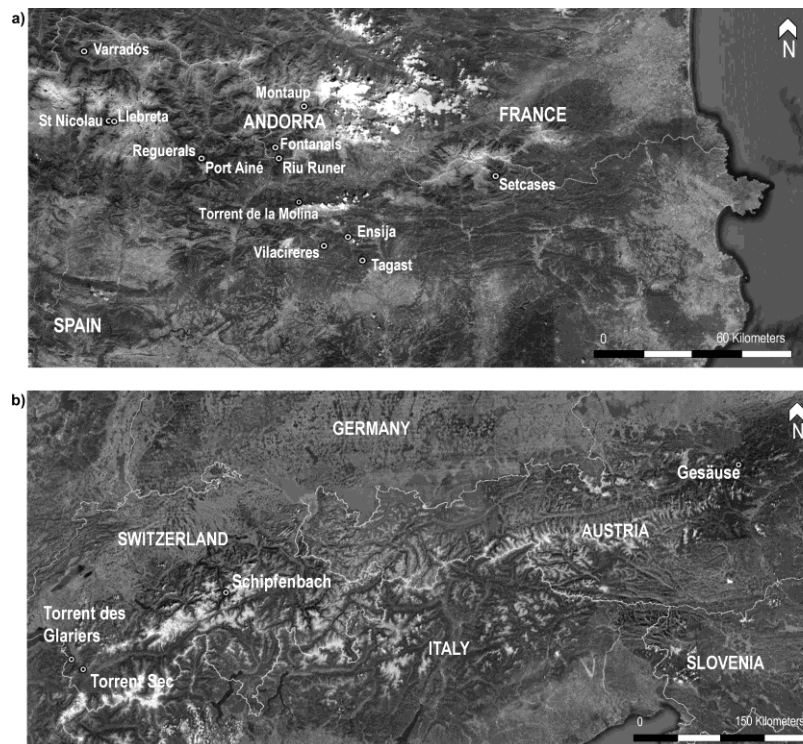


Figura 1: Situación de los torrentes estudiados en los Pirineos (a) i en los Alpes (b).

### 3.1. Factores de campo

#### a) Pendiente media del tramo ( $P_i$ )

La pendiente media de un tramo ( $P_i$ ) es un factor numérico medido en grados. En este trabajo la pendiente de los tramos fue medida en el campo con un clinómetro estándar: una única medida para los tramos cortos (pocas decenas de metros) y varias para tramos largos, de las cuáles se calculó el valor medio.

#### b) Disponibilidad de sedimento ( $DS_i$ )

La disponibilidad de sedimento en un tramo  $i$  ( $DS_i$ ) es un factor categórico. De acuerdo con las observaciones de campo, es uno de los factores más influyentes en la erosión producida durante una corriente de derrubios. En algunos torrentes, los tramos pueden mostrar disponibilidad ilimitada de sedimento. En tales casos la erosión está totalmente condicionada a los otros factores, mientras que en otros tramos, la falta de sedimento limita la erosión. En este trabajo se distinguen cinco clases de disponibilidad de sedimento en función del porcentaje de sección transversal cubierto por un cierto espesor de coluvión (Tabla 1).






Clase	Imagen	% rocoso en la sección transversal	% coluvión en la sección transversal	espesor de coluvión (m)
Completa limitación (CL)		> ~90	< ~10	< ~0.5
Alta limitación (AL)		> ~75	< ~25	< ~1
Media limitación (ML)		~50	~50	~1
Baja limitación (BL)		< ~25	> ~75	> 1
No limitación (NL)		< ~10	> ~90	> 1

Tabla 1: Clases de disponibilidad de sedimento

c) Forma de la sección transversal ( $ST_i$ )

La forma de la sección transversal es un factor categórico que se consideró con el ánimo de ilustrar el grado de incisión del tramo. Para determinar la forma de la sección transversal se adaptó el concepto propuesto por Gabet y Bookter (2008). Se determinó que la anchura del canal ( $W$ ) a una cierta altura de la vaguada ( $H$ ) era una medida significativa para describir la forma de la sección. La altura se debe determinar con criterio experto, mediante una aproximación del calado máximo que se produciría en caso que ocurriera una corriente de derrubios. Con la intención de simplificar el trabajo de campo, se describieron tres clases de forma: ancho, moderadamente incidido e incidido, en lugar de utilizar la anchura exacta. Los límites de las clases se establecieron en base a la observación durante las campañas de campo (Tabla 2).

Clase	Ancho	Moderadamente incidido	Incidido
Definición	$W \leq 2H$	$2H < W \leq 5H$	$W > 5H$

Tabla 2: Clases de forma de la sección transversal.

### 3.2. Factores topográficos

#### a) Área contributiva ( $AC_i$ )

El área contributiva ( $AC_i$ ) es un factor numérico que puede ser fácilmente determinado en un SIG. Se calcula como la suma de píxeles que drenan agua u otras sustancias al punto más bajo de un tramo. El valor de la suma de los píxeles es obtenido mediante la herramienta de Flow Accumulation que se encuentra en ArcGis® (ESRI, 2005), utilizando un modelo digital de elevaciones.

El área contributiva se aplicó en este trabajo con la intención de ilustrar las diferencias de caudales y transporte de sedimento que pueden existir entre tramos. Por ejemplo, una corriente de derrubios puede iniciarse con una concentración similar a un flujo hiperconcentrado (en un punto de pequeña área contributiva) pero desarrollarse hasta formar una corriente de derrubios madura a las partes bajas del torrente, cerca del abanico, donde las áreas contributivas son mayores.

## 4. MINERÍA DE DATOS

La minería de datos es el proceso de descubrimiento de patrones, a menudo escondidos, en una base de datos (Fayyad et al., 1996). En este trabajo el objetivo era desarrollar una metodología para predecir una aproximación de las tasas de erosión que se pueden producir en un cierto torrente a lo largo de una corriente de derrubios.

En primer lugar, se analizaron las relaciones simples entre variables: tanto entre factores condicionantes como entre estos y la variable objetivo ( $e_i$ ). El análisis se llevó a cabo mediante histogramas y correlaciones lineales de variables entre los 110 tramos de la base de datos. En segundo lugar, se aplicaron técnicas de minería de datos. Se utilizaron dos tipos de análisis basados en dos técnicas de minería de datos distintas: la regresión lineal múltiple y el árbol de decisión. Las técnicas se aplicaron sobre una parte de la base de datos: los datos de entrenamiento (formado por 93 tramos de la base de datos). A continuación se procedió a la validación de los modelos obtenidos mediante parámetros estadísticos, sobre un grupo de datos independiente de los datos de entrenamiento (datos de validación, formado por 10 tramos representativos del total de la base de datos). Finalmente se realizó una comprobación de la eficiencia de los modelos en un nuevo grupo de datos (datos de comprobación, formado por 7 tramos de un mismo torrente), comparando el total de volumen obtenido en un evento con el volumen observado en campo y el volumen obtenido mediante dos fórmulas empíricas.

### 4.1. Regresión lineal múltiple

La regresión lineal múltiple es una técnica de aprendizaje que tiene como objetivo encontrar la mejor relación lineal entre la tasa de erosión de un tramo ( $e_i$ ) y los factores condicionantes, mediante la expresión siguiente (1):

$$e_i = a_1 P_i + a_2 DS_i + a_3 ST_i + a_4 AC_i + b \quad (1)$$

Los coeficientes  $a_1$ ,  $a_2$ ,  $a_3$ ,  $a_4$  y  $b$  se determinan mediante la minimización de la suma del cuadrado de los errores originados por la aproximación de la erosión en cada uno de los

tramos de los datos de entrenamiento. Se otorgó un valor numérico a las variables categóricas (ST y DS). Por ejemplo: una sección transversal ancha se representó con el valor 1, una sección transversal medianamente incidida con el valor 2 y una sección transversal incidida se representó con un valor 3. Lo mismo se hizo con las clases de DS, con números del 1 (limitación de sedimento completa) al 5 (sedimento no limitado).

Previamente a la aplicación de la técnica de regresión múltiple, se comprobó si todos los factores condicionantes presentaban una distribución normal. En caso de que no la presentasen, se realizó una conversión de los valores mediante la aplicación de una función  $\log(x)$ , con la intención de obtener una nueva variable con distribución normal (Gartner et al., 2008).

#### 4.2. Árbol de decisión

Los árboles de decisión son una familia de técnicas de minería de datos que permiten utilizar directamente valores categóricos. Se basan en dos tipos de elementos esenciales: los *nodos* y las *hojas*. En cada *nodo* se evalúa un factor condicionante y se compara con un valor (en caso de factor numérico) o se escoge una cierta categoría (en caso de factores categóricos). Así se avanza por el árbol hasta el punto en que se llega a una *hoja*, donde se obtiene un valor de la variable objetivo ( $e_i$ ).

En este trabajo se ha utilizado el programa WEKA (Hall et al., 2009) para obtener el árbol de decisión. WEKA es un programa que incorpora 77 algoritmos de técnicas de aprendizaje, de los cuales 16 son árboles de decisión. Después de una fase de evaluación previa con varios algoritmos de árboles, se escogió el algoritmo J48 (Breiman et al., 1984). Se utilizó un método de optimización llamado “validación cruzada en 10 particiones”.

El proceso de aprendizaje del árbol de decisión puede presentar problemas en caso que la variable objetivo no esté distribuida equitativamente, ya que esto puede provocar que se obvien partes de la base de datos debido al peso de las otras. Por esta razón se utilizan matrices de coste como estrategia para crear árboles de decisión en bases de datos desequilibradas (Witten et al., 2011). Se implementan en forma de metaclasificadores durante el proceso de aprendizaje y otorgan un coste mayor a la clasificación errónea de los grupos minoritarios de la base de datos y uno menor a los grupos más abundantes.

#### 4.3. Validación y comprobación de los modelos

Los modelos obtenidos se evaluaron por duplicado: primero los dos modelos se compararon con los datos de validación y segundo se implementaron los dos modelos en un torrente donde el volumen final era un dato conocido (datos de comprobación). En esta parte se estableció una comparación mediante los modelos y dos fórmulas empíricas.

La primera parte (validación) consiste en evaluar el éxito de los modelos. Normalmente la evaluación del éxito de los modelos predictivos se evalúa mediante análisis ROC. No obstante, el análisis ROC resulta complicado y poco intuitivo en casos de que la variable objetivo tenga más de dos clases. Su uso en este tipo de casos es poco común, por este motivo se optó por utilizar el parámetro F (*F-measure* en inglés), que combina el término “*precision*” (verdaderos positivos sobre totales positivos) y “*recall*” (verdaderos positivos sobre la suma

de verdaderos positivos y falsos negativos). Se calculó el valor de F para cada clase, así como también su valor medio ponderado para cada modelo, tanto para los datos de entrenamiento como para los de validación. El parámetro F se expresa de la forma siguiente (2):

$$F = \frac{2 \cdot \text{precision} \cdot \text{recall}}{(\text{precision} + \text{recall})} \quad (2)$$

La segunda parte (comprobación) consiste en aplicar ambos modelos a un torrente de la base de datos que no se utilizó ni para el entrenamiento ni para la validación. Se escogió el torrente de Schipfenbach (Hürlimann et al., 2003) (Figura 1), formado por 7 tramos, debido a su variabilidad en cuanto a los factores condicionantes. Los resultados obtenidos mediante los dos modelos se compararon con dos ecuaciones empíricas simples para el cálculo del volumen: (3) y (4) (Rickenmann 1995) y (5) (D'Agostino 1996).

$$\theta_{cono} \leq 15\% \Rightarrow V = (6.4\theta_{cono} - 23)\text{alcance}(m) \quad (3)$$

$$15\% < \theta_{cono} < 40\% \Rightarrow V = (110 - 2.5\theta_{cono})\text{alcance}(m) \quad (4)$$

$$V = 45A_{cuenca}^{0.9}\theta_{torrente}^{1.5}(\text{IndiceGeológico}) \quad (5)$$

## 5. RESULTADOS

### 5.1. Distribución espacial de los factores condicionantes

La base de datos contiene un total de 110 tramos, de los cuales 84 situados en los Pirineos y 19 en los Alpes. En la Figura 2 se muestran los histogramas y las curvas de frecuencia acumulada de los factores condicionantes y de la variable objetivo. En la figura se distinguen los datos de validación y entrenamiento de los de comprobación (torrente Schipfenbach).

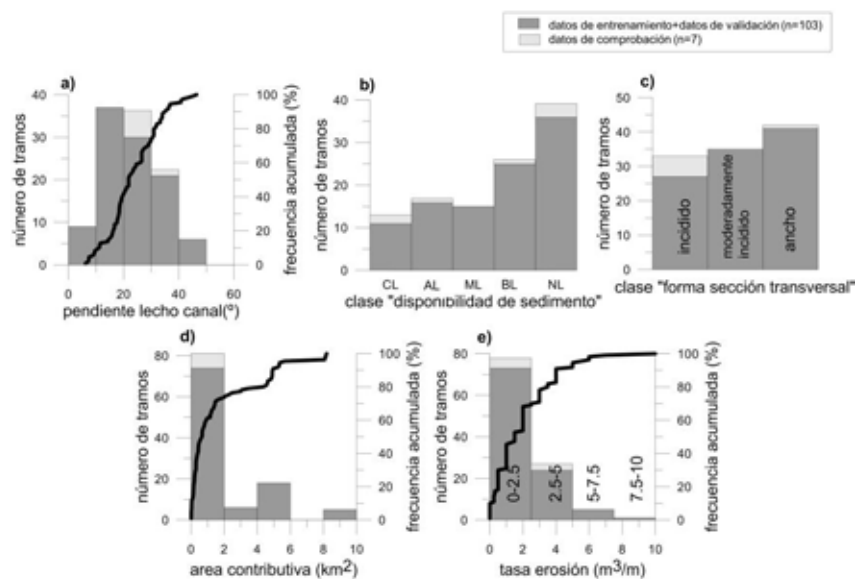


Figura 2: Histogramas de los 4 factores condicionantes (a, b, c, d) y la tasa de erosión (ei) para los 110 tramos incluidos en la base de datos. La frecuencia acumulada se muestra para los factores condicionantes numéricos y la tasa de erosión (ei). Ver Tabla 1 para abreviaciones de las clases de SA clases.

El histograma de la pendiente del lecho del canal muestra un pico en la clase que comprende los valores entre  $10^\circ$  y  $20^\circ$ . Los tramos de pendiente superior a  $40^\circ$  e inferior a  $10^\circ$  representan menos de un 20% de la base de datos. Estos resultados están acorde con otras publicaciones (Marchi y D'Agostino, 2004; Hungr et al., 2005). La mayoría de los tramos no presentan limitaciones de sedimento disponible y los tramos de forma de la sección incidida son los más comunes en la base de datos. Las áreas contributivas toman valores de hasta  $8.2 \text{ km}^2$ , pero un 60% de los tramos tienen áreas contributivas inferiores a  $1 \text{ km}^2$ .

La tasa de erosión muestra un gráfico de frecuencia acumulada que incrementa fuertemente para los valores de erosiones bajas y suavemente para las altas tasas de erosión. A pesar de que en algunos tramos se describieron erosiones altas, los valores típicos de este estudio son menores que en otros estudios publicados, como el de Hungr et al. (1984). Ésta diferencia de escala se puede atribuir a que en el presente estudio no se consideraron las roturas laterales del torrente, sumado al hecho de que el trabajo de Hungr se sitúa en un contexto geológico diferente.

A parte de los histogramas, se elaboraron gráficos de relaciones lineales por pares de factores, o en grupos de 3. La mayoría de gráficos mostraron grandes dispersiones, cosa que reafirma la complejidad del proceso y refuerza la necesidad de utilizar técnicas más complejas para realizar predicciones de volumen de erosión más adecuadas.

## 5.2. Regresión lineal múltiple

Previamente a la obtención de la fórmula que describe la tasa de erosión en función de los factores condicionantes se realizó una transformación de los factores que no presentaban una distribución normal (y eran factores numéricos). Dichos factores fueron el área contributiva y la propia tasa de erosión, tal y como muestran los histogramas (Figura 2). Se comprobó que, evidentemente, el modelo de regresión lineal obtenido en el segundo caso mostraba mejores resultados ( $R^2$  y error estándar) que en caso de no normalizar las distribuciones de los factores. Finalmente, la expresión resultante de la regresión lineal múltiple fue la siguiente:

$$\log(e_i + 1) = 0.007P_i + 0.102DS_i - 0.066ST_i + 0.053\log(AC + 1) - 0.312 \quad (6)$$

## 5.3. Árbol de decisión

El árbol resultante de aplicar el algoritmo J48 a los datos de entrenamiento se muestra en la Figura 3. La tasa de erosión se dividió en 4 clases de  $2.5 \text{ m}^3/\text{m}$  (las mismas clases que las indicadas en el histograma de la Figura 2e). El intervalo de  $2.5 \text{ m}^3/\text{m}$  se considera una resolución razonable para la resolución de la metodología a la vez que la simplifica.

Durante la creación del árbol se utilizó una matriz de costes, con el fin de forzar la aparición de las 4 clases. Se atribuyó mayor coste a la mala clasificación de los tramos con altas tasas de erosión mientras que los costes de la mala clasificación eran menores para los tramos con menor tasa de erosión.

La primera hoja es la de la disponibilidad de sedimento. Las ramas con mayor limitación (AL y CL) llegan directamente a la hoja final, obteniendo la clase de menor tasa de erosión. Las



ramas con más sedimento disponible (ML, BL, NL) necesitan uno o dos niveles adicionales para llegar a la hoja con el valor de la tasa de erosión.

La forma de la sección transversal es el segundo factor que aparece en dos de las clases de la disponibilidad de sedimento, mientras que la pendiente aparece en la tercera. En el siguiente nivel se encuentra el área contributiva o la pendiente. En general, la clase dominante en el árbol de decisión es la tasa de erosión menor: " $0 \leq e_i \leq 2.5 \text{ m}^3/\text{m}$ ", que aparece en la mitad de las hojas del árbol. La tasa de erosión máxima sólo aparece en tramos de erosión ilimitada y alta pendiente.

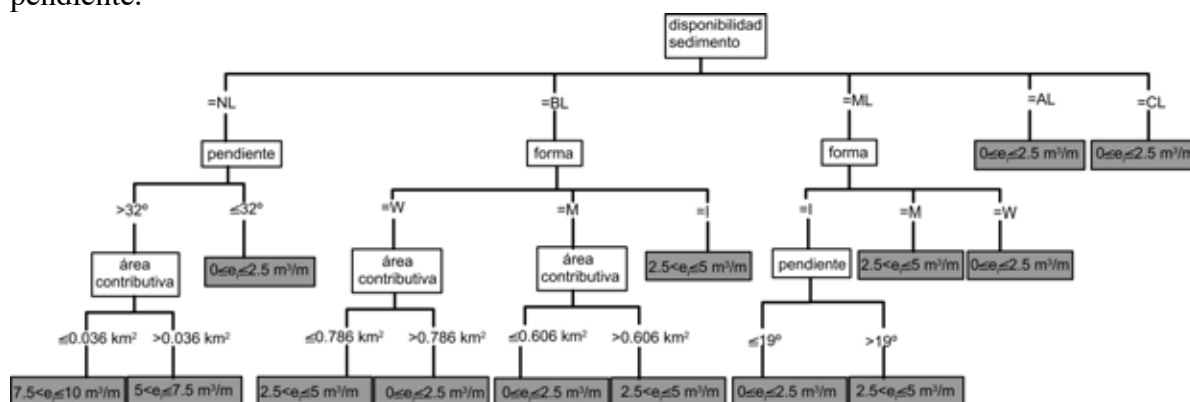


Figura 3: Árbol de decisión para la estimación de la tasa de erosión estimada. Se ha elaborado aplicando el algoritmo J48 sobre los datos de entrenamiento.

#### 5.4. Validación y comprobación de los modelos

Se calcularon los valores de  $F$  para cada clase de tasa de erosión para los datos de entrenamiento y de validación. Los valores se muestran en la Tabla 3, donde se puede observar que ambos modelos muestran buenos resultados para la clase de erosión más baja (menor de  $2.5 \text{ m}^3/\text{m}$ ) en los dos grupos de datos. Este es un resultado lógico ya que se trata de la clase de erosión más frecuente en la base de datos. En contraposición, la regresión lineal múltiple (MLR) muestra malos resultados, especialmente para las clases de erosión mayores. En cuanto a los valores de la media ponderada del parámetro  $F$ , los mejores resultados se obtienen con el árbol de decisión (AD).

Los resultados de la comprobación muestran que los volúmenes calculados con las fórmulas empíricas sobreestiman de manera importante el volumen observado en campo (Figura 4). Este resultado puede atribuirse al hecho de que los modelos empíricos muestran una gran dispersión, o bien al hecho de que fueron desarrollados en una zona específica y su aplicación fuera de esa zona puede generar mucha inexactitud. La regresión lineal múltiple muestra los mejores resultados, mientras que el árbol de decisión define un rango de volúmenes en los cuales el valor observado está incluido.

Modelo	Clase	Datos de entrenamiento		Datos de validación	
		n (total 93)	F	n (total 10)	F
Regresión lineal múltiple	$0 \leq e_i \leq 2.5 \text{ m}^3/\text{m}$	67	0.940	6	0.714
	$2.5 < e_i \leq 5 \text{ m}^3/\text{m}$	17	0.117	3	0
	$5 < e_i \leq 7.5 \text{ m}^3/\text{m}$	8	0	1	0
	$7.5 < e_i \leq 10 \text{ m}^3/\text{m}$	1	0	0	0
	Media ponderada	0.632		0.429	
Árbol de decisión	$0 \leq e_i \leq 2.5 \text{ m}^3/\text{m}$	67	0.917	6	0.857
	$2.5 < e_i \leq 5 \text{ m}^3/\text{m}$	17	0.649	3	0.8
	$5 < e_i \leq 7.5 \text{ m}^3/\text{m}$	8	0.4	1	0
	$7.5 < e_i \leq 10 \text{ m}^3/\text{m}$	1	0.5	0	0
	Media ponderada	0.819		0.754	

Tabla 3: Valores del parámetro F para cada clase de los dos modelos para predecir la erosión en el datos de entrenamiento y los de validación.

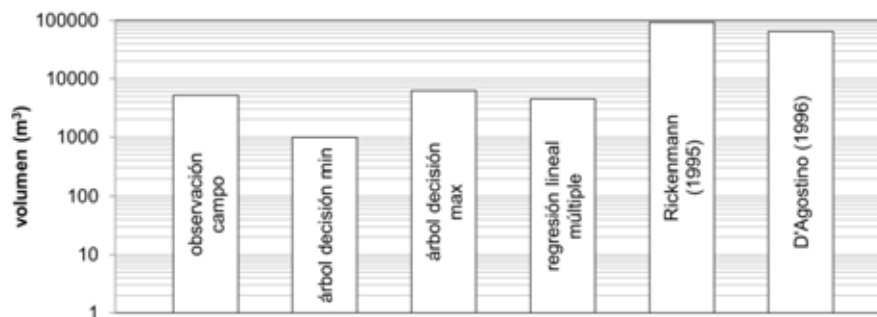


Figura 4: Resultados de la comprobación de los modelos.

## 6. MODELO GENERAL PROPUESTO PARA LA ESTIMACIÓN DE LA TASA DE EROSIÓN

Se seleccionó el modelo del árbol de decisión para proponer un nuevo modelo general para estimar la tasa de erosión. La comparación de la regresión lineal múltiple (RLM) y el árbol de decisión (AD) mostraron varias ventajas del AD sobre RLM: 1) es más simple, 2) puede ser más fácilmente adaptable a otras regiones, y 3) mostró mejores resultados que RLM en términos de parámetros estadísticos. El nuevo modelo general se basa en el árbol obtenido de nuestra base de datos combinado con el criterio experto (Figura 5). Es importante tener en mente que el árbol se obtuvo con una base de datos que incluye eventos con volúmenes desde miles hasta decenas de miles de metros cúbicos. Por este motivo, eventos extremos (como de centenares de miles o millones de metros cúbicos) no se han considerado. Entonces, estos resultados son representativos para eventos de corrientes de derrubios de media-alta frecuencia ocurridos en los Pirineos y los Alpes, pero no para eventos de baja frecuencia, asociados a volúmenes considerablemente mayores.

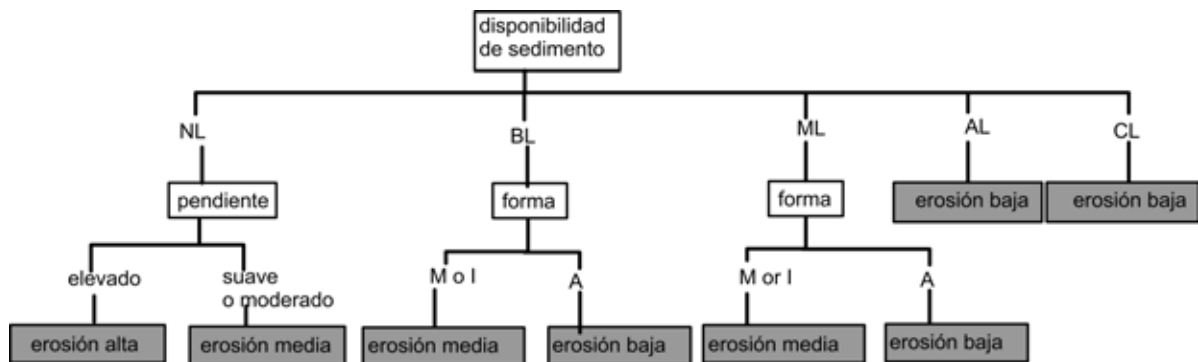


Figura 5: Árbol de decisión propuesto como modelo de aprendizaje para estimar la tasa de erosión en un tramo de torrente.

## 7. CONCLUSIONES

La erosión es un proceso complejo. El análisis preliminar de la información de campo y topográfica obtenida en 110 tramos de 17 torrentes donde se produjeron corrientes de derrubios muestra que no hay correlaciones claras entre factores, la cual cosa evidencia la complejidad del proceso de la erosión.

Las técnicas de minería de datos se aplicaron con la intención de describir un modelo capaz de predecir una estimación de la tasa de erosión en un cierto tramo. La minería de datos mostró que la disponibilidad de sedimento es el factor más importante seguido por el pendiente y la forma de la sección transversal. Los resultados fueron mejores con el algoritmo del árbol de decisión que con la regresión lineal múltiple, aunque al aplicar los dos modelos en el torrente de Schipfenbach RLM se ajustó más al resultado observado en campo que el AD.

Finalmente se ha propuesto un modelo general para estimar la tasa de erosión por tramos de torrente que permite obtener una estimación del volumen de material erosionado en un torrente debido a un evento de corriente de derrubios. Debe tenerse en cuenta que este modelo no tienen cuneta ni las roturas laterales que puedan ocurrir en el torrente debido al paso de la corriente de derrubios.

## AGRADECIMIENTOS

Esta investigación ha sido financiada por el Ministerio de Economía y Competividad, Secretaría de Estado de Investigación, Desarrollo e Innovación, proyecto con código CGL2011-23300 (DEBRIS-START). Los autores quieren agradecer al Institut Geològic de Catalunya, al Parc Nacional d'Aigüestortes i Estany de Sant Maurici, al Nationalpark Gesäuse, a François Xavier Marquis y al WSL todo su soporte para la adquisición de datos.

## REFERENCIAS

Berger, C., McArdeil, B.W., Schlunegger, F., 2011. Sediment transfer patterns at the Illgraben catchment, Switzerland: Implications for the time scales of debris flow activities. *Geomorphology*, 125, 421-432.

- Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. *Classification and regression trees*, London.
- D'Agostino, V., 1996. *Analisi quantitativa e qualitativa del trasporto solido torrentizio nei bacini montani del Trentino Orientale, 1a Sezione, Volume presentato in occasione del Convegno di Studio: I problemi dei grandi comprensori irrigui*, Novara, pp. 111-123.
- ESRI, 2005. *ArcGIS 9 : what is ArcGIS 9.1?* ESRI, Redlands, California, 123 pp.
- Fannin, R.J., Rollerson, T.P., 1993. Debris flows: some physical characteristics and behaviour. *Canadian Geotechnical Journal*, 30, 71-81.
- Fayyad, U., Piatetsky-Saphiro, G., Smyth, P., 1996. From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17 (3).
- Gabet, E.J., Bookter, A., 2008. A morphometric analysis of gullies scoured by post-fire progressively bulked debris flows in southwest Montana, USA. *Geomorphology*, 96, 298-309.
- Gartner, J.E., Cannon, S.H., Santi, P.M., DeWolfe, V.G., 2008. Empirical models to predict the volumes of debris flows generated by recently burned basins in the western U.S. *Geomorphology*, 96, 339-354.
- Gertsch, E., 2009. *Geschiebelieferung alpiner Wildbachsysteme bei Grossereignissen - Ereignisanalysen und Entwicklung eines Abschätzverfahrens* Universität Bern, Bern, 204 pp.
- Guthrie, R.H., Hockin, A., Colquhoun, L., Nagy, T., Evans, S.G., Ayles, C., 2009. An examination of controls on debris flow mobility: Evidence from coastal British Columbia. *Geomorphology*, 114 (4), 601-613.
- Hall, M.A., Frank, E., Holmes, G., Pfahringer, B., Reutmann, P., Witten, I.H., 2009. *The WEKA Data Mining Software: An Update* SIGKDD Explorations, 11 (1), 10-18.
- Hungr, O., Morgan, G.C., Kellerhals, R., 1984. Quantitative analysis of debris torrent hazards for design of remedial measures. *Canadian Geotechnical Journal*, 21, 663-677.
- Hungr, O., Morgenstern, N.R., 1984. High velocity ring shear tests on sand. *Géotechnique*, 34 (3), 415-421.
- Hungr, O., McDougall, S., Bovis, M., 2005. Entrainment of material by debris flow. In: Jakob, M., Hungr, O. (Eds.), *Debris-flow Hazards and Related Phenomena*. Springer, Berlin, pp. 135-158.
- Hürlimann, M., Rickenmann, D., Graf, C., 2003. Field and monitoring data of debris-flow events in the Swiss Alps. *Canadian Geotechnical Journal*, 40 (1), 161-175.
- Marchi, L., D'Agostino, V., 2004. Estimation of debris-flow magnitude in the Eastern Italian Alps. *Earth Surface Processes and Landforms*, 29 (2), 207-220.
- Rickenmann, D., 1995. Beurteilung von Murgängen. *Schweiz. Ingenieur und Architekt*, 113 (48), 1104-1108.
- Scheidl, C., Rickenmann, D., Chiari, M., 2008. The use of airborne LIDAR data for the analysis of debris flow events in Switzerland. *Nat. Hazards Earth Syst. Sci.*, 8, 1113-1127.
- Witten, I.H., Frank, E., Hall, M.A., 2011. *Data Mining. Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, 629 pp.