

Predicción de aludes en el puerto de la Bonaigua

Análisis mediante árboles de clasificación (CART) de episodios de nevadas, actividad de aludes y patrones meteorológicos en el puerto de la Bonaigua (Val d'Aran)

Pablo Gorospe

Tutor: Ivan Moner

Cotutor: Joan Manuel Vilaplana

Máster de Recursos Minerales y Riesgos Geológicos

Especialidad: Riesgos Geológicos

Julio 2016



UNIVERSITAT DE
BARCELONA

UAB

Universitat Autònoma
de Barcelona

Tabla de contenidos

0	Abstract.....	1
1	Introducción	1
1.1	Zona de estudio	1
1.2	Antecedentes	2
1.3	Objetivos	2
2	Clima y dinámica de avalanchas en el puerto de la Bonaigua	3
2.1	Climatología.....	3
2.2	Terreno de avalanchas	4
3	Metodología	6
3.1	Datos meteorológicos	7
	• Fuentes de datos y su elección	7
	• Procesado de los datos. Variables primarias	7
	• Procesado de los datos. Variables derivadas	8
3.2	Datos de avalanchas	9
	• Registro de avalanchas	9
	• Definición día de avalanchas.....	10
3.3	Construcción de la base de datos.....	11
3.4	Análisis estadístico.....	13
	• Construcción del árbol máximo	13
	• Poda del árbol.....	13
	• Elección del árbol optimo	14
4	Resultados	14
4.1	Árboles del modelo general.....	14
4.2	Árboles sub-zonas	17
	• Sub-zona 1	17
	• Sub-zona 2.....	17
	• Sub-zona 3.....	19
5	Análisis.....	20
6	Conclusiones	21
7	Anexos.....	22
8	Bibliografía	22

Abstract

The highway C-28 is located in the Central Pyrenees and links the Aran valley with Catalonia along 20 km over the Bonaigua Pass. It constitutes a key access route for winter visitors. Most of the slopes affecting the road face to the south, with heights varying between 1600 and 2300 meters. The road runs through 58 avalanche paths of which 25 release frequently. Most of them are already capable to produce D2 or D3 avalanches, but in many places the road is carved in the hill side, what means that every single avalanche will certainly hit the road. We started from 12 years of meteorological and avalanche data collected by the local avalanche warning service of Aran Valley. Weather data were obtained from two automatic weather stations and a flowcapt, whereas avalanche activity was manually recorded in a GIS with a high level of detail. We selected 32 weather parameters including snow drift, time lagged, trend and categorical parameters and 137 avalanche days, in total. Using a classification tree method, we have developed a model to determine periods of significant avalanche activity in terms of the pre-defined avalanche day concept. The model is performed for the entire road in a combined analysis and also for three individual sections within the Pass. Results showed that time lagged conventional factors describing temperature and precipitation are more significant than the trend factors. Snow drift parameters segregated into prominent wind direction have not been found significant as solar radiation. A model for a single section with more homogeneous avalanche behavior allows us to highlight variables that were hidden in the general model applied for the overall highway.

1 Introducción

La carretera C-28 es una de las dos vías de comunicación del Valle de Aran con Cataluña y paso clave para la comunicación con Francia (Fig. 3a). Esta zona de los pirineos presenta abundantes precipitaciones en forma de nieve durante el invierno lo que provoca una intensa actividad de avalanchas. El trazado de la carretera a su paso por el Puerto de la Bonaigua sufre durante la temporada invernal cortes frecuentes al ser afectada por los aludes que se producen en las laderas contiguas lo que además puede causar daños a los vehículos en tránsito y sus ocupantes (Fig. 1). Estas incidencias repercuten de forma directa sobre el turismo invernal, principal actividad económica del valle.



Fig. 1: Auto caravana afectada por un alud en el sector Recta de Ruda. Ambos carriles de la carretera se han visto afectados. Foto de: Centre de Lauegi d'Aran, Conselh Generau d'Aran.

1.1 Zona de estudio

El trazado de la carretera C-28 a través del Puerto de la Bonaigua se extiende a lo largo de 20 Km (Fig. 3b) y atraviesa en su camino 58 trayectorias de avalanchas, de las cuales 25 presentan una actividad frecuente (Fig. 5). La mayoría de estas zonas de aludes son solo capaces de generar depósitos de tamaño D2 (pequeño) o D3 (mediano), dentro de una escala de cinco niveles (McClung and Shaerer, 1980), pero al encontrarse la carretera excavada en la ladera, cualquier avalancha que se inicie en la cabecera de las laderas afectara con toda seguridad al

trazado. Se han instalado elementos de protección permanente como barreras de viento y barreras fijas en algunas de las zonas de salida (Fig. 2) pero la mayoría de las existentes no disponen de estas estructuras de defensa. También se ha implantado una barrera para cortar el tráfico cuando existe peligro de avalanchas (Gavaldà and Benet, 2014).



Fig. 2: Ejemplo de defensas activas permanentes en la rama sur de la carretera bajo el sector de Comials. Foto de: Centre Lauegi d'Aran, Conselh Generau d'Aran

1.2 Antecedentes

Existen trabajos publicados recientemente de evaluación del riesgo de aludes en este tramo de la carretera C-28. Algunos cubren la totalidad del trazado del puerto como Xifre (2007) y otros como Gavaldà (2002) el sector de Comials (Fig. 5). En ambos estudios se proponen actuaciones para la reducción del riesgo de aludes en la carretera. Hasta el año 2003, la gestión del peligro de aludes para esta carretera estaba basada en el boletín de peligro de aludes de carácter regional que elabora el Institut Cartografic i Geologic de Catalunya (ICGC) para el Pirineo catalán. A partir del invierno 2003-2004 se puso en marcha el servicio de predicción local de aludes, Centre de Lauegi, perteneciente al Conselh Generau d'Aran. Este centro se encarga de gestionar el peligro de aludes en las carreteras de Bonaigua y Beret además de elaborar un boletín de peligro de aludes para montañeros. La predicción local se basa en los datos nivometeorológicos obtenidos a partir de cuatro estaciones automáticas y también de dos parcelas de observación gestionadas por los pisteros de la estación de esquí de Baqueira, así como de las observaciones de campo realizadas por los predictores del centro. Desde la temporada 2008-2009 se llevan a cabo desencadenamientos preventivos mediante el dispositivo DaisyBell® (Gavaldà and Moner, 2009). El centro de predicción también ha realizado una cartografía pormenorizada de las zonas de salida de los aludes que afectan a la carretera y de los eventos registrados. Como explican Gavaldà y Benet (2014) este nuevo modelo de gestión de la carretera basado en una predicción local ha demostrado ser mucho más eficiente y ha permitido conocer con mayor profundidad la dinámica de avalanchas de la carretera, la cual se ha dividido en siete sectores según su comportamiento (Fig. 5).

1.3 Objetivos

Fruto de la actividad del centro en las trece temporadas que lleva activo se ha generado un registro diario de datos nivometeorológicos y de cartografía de los eventos de aludes sucedidos en este tiempo. El objetivo principal de este trabajo es utilizar esta base de datos para determinar cuáles son las variables meteorológicas más significativas y potencialmente desarrollar un modelo predictivo que permita determinar si habrá actividad de avalanchas en este tramo de la carretera C-28 mediante el análisis con un árbol de clasificación. El conocimiento de la dinámica de los aludes en los diferentes sectores del trazado ha permitido diferenciar tres subzonas con un comportamiento avalanchoso más homogéneo. De esta manera, además del análisis para el conjunto del trazado, se ha llevado a cabo un estudio individualizado de las mismas para determinar si existen diferencias respecto al modelo general y se muestran relevantes otras variables.

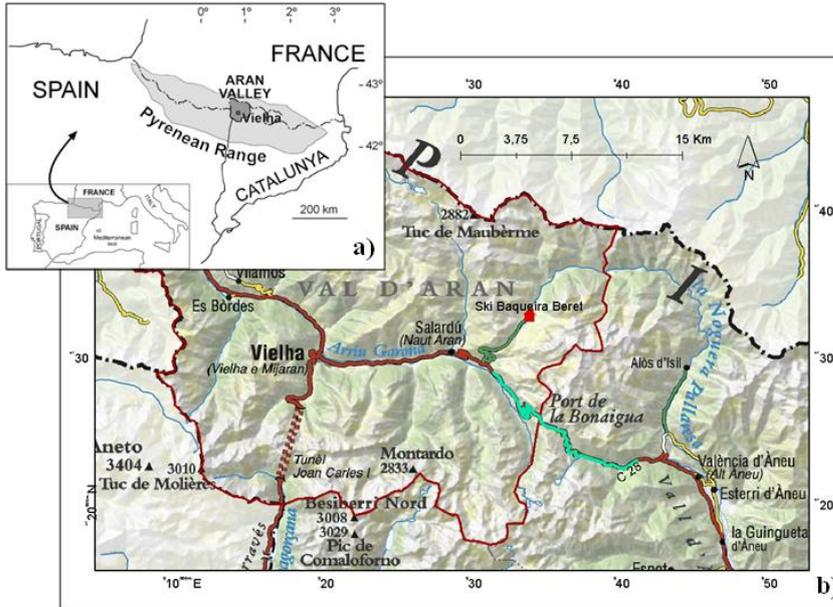


Fig. 3a: Situación del valle de Aran. Tomado de: (Gavaldà and Benet, 2014)

Fig. 3b: Situación de la carretera C-28. En azul el trazado objeto de estudio a su paso por el Puerto de la Bonaigua. © Base cartográfica propiedad del Institut Cartogràfic i Geològic de Catalunya, disponible en www.icgc.cat

2 Clima y dinámica de avalanchas en el puerto de la Bonaigua

2.1 Climatología

El Valle de Aran presenta un clima oceánico con fuerte influencia de los flujos húmedos del océano Atlántico que dejan abundantes precipitaciones a su paso por el valle. La media de precipitación líquida es de 1000 mm por año llegando incluso a 1500 mm en alturas superiores a 2000 m. Se caracterizan por una distribución homogénea a lo largo del año. Durante la temporada invernal las precipitaciones en forma de nieve por encima de los 2200 m suman un total de entre 600 y 700 cm anuales de promedio.

Los 20 Km de la carretera C-28 a su paso por el puerto de la Bonaigua salvan un desnivel de casi 600 m. El punto más alto del trazado a 2075 m de altura separa dos sectores climáticamente bien diferenciados. Como se recoge en el trabajo de Gavaldà and Benet (2014), la rama norte de la carretera (Fig. 3b), situada en territorio aranés, presenta una marcada influencia de los flujos atlánticos que afectan a la vertiente norte de los Pirineos. Por otro lado la rama sur de la carretera, en territorio del Pallars, también está dominada por las precipitaciones atlánticas pero son mucho menos abundantes. La topografía del Valle de Ruda, situado en las faldas del puerto de la Bonaigua canaliza los flujos procedentes del norte y del noroeste causando en la temporada invernal un transporte eólico intenso de nieve en la vertiente norte que genera placas de viento fundamentalmente en la vertiente sur del puerto.

La diferencia de cotas que salva esta carretera, la disposición frente a los flujos atlánticos de los relieves circundantes y la orientación preferente hacia el sur de la mayoría de las laderas que afectan al trazado determina que esta zona este sometida a distintas situaciones de avalanchas. Este término, acuñado por el servicio de aludes suizo (SLF), describe siete situaciones tipo en las que se puede encontrar el manto nival. Este método se ha revelado mucho más útil en las operaciones de predicción que otros métodos de clasificación basados en criterios morfológicos como el atlas de clasificación de la UNESCO (International Commission of Snow and Ice, 1981) además su uso se ha extendido en los últimos años en los boletines de peligro de aludes por su fácil comprensión para los usuarios. Las siete situaciones de avalanchas se reflejan en la tabla 1 con sus características. Como se adelantaba en la introducción el objetivo de este trabajo no es desarrollar un modelo para un tipo concreto de avalancha ni tampoco diferenciar entre nieve húmeda o seca. Sin embargo es conveniente recordar estas situaciones de avalanchas y su significado ya que ha sido el criterio que se ha utilizado para diferenciar las tres subzonas dentro del trazado de la carretera de acuerdo a su dinámica (Fig. 5).

Situaciones de aludes	Descripción	Condiciones
Nieve reciente	Después de nevadas significativas la nieve tarda un tiempo en estabilizarse. Esta nieve nueva puede desencadenarse en forma de alud de placa o de salida puntual.	-Fuerte intensidad de precipitación -Vientos fuertes (>40 km/h) -Temperaturas bajas (por debajo de -5 a -10°C) -Temperatura del aire en ascenso durante la nevada
Nieve venteada	El viento es el arquitecto de los aludes de placa, al ser el que forma las placas de viento. Se forma en lugares protegidos del viento como sotaventos.	- Vientos entre 30 y 80 km / h. - Nieve transportable: nevada reciente, o no cohesivas en superficie
Nieve vieja	A veces dentro del manto de nieve se desarrollan y conservan capas débiles, susceptibles de fracturarse por una sobrecarga y causar un alud de placa.	-Zonas con manto relativamente delgado - Muy a menudo en laderas sombrías
Nieve húmeda a lo largo del día	El agua proveniente de la fusión debilita el manto de nieve y puede causar avalanchas de nieve húmeda.	-Cielo cubierto, sobre todo de noche y madrugada -Temperatura alta / fuerte radiación solar -Rehielo nocturno débil
Nieve húmeda por lluvia	La lluvia debilita y sobrecarga el manto de nieve y puede causar avalanchas de nieve húmeda. La niebla con temperaturas positivas tiene un efecto similar.	-Lluvia sobre manto seco -Lluvia sobre nieve reciente, caída en los últimos 6 días -Lluvia abundante e intensa
Deslizamientos basales o glide	Las avalanchas de deslizamiento son un subtipo de aludes de fondo, y ocurren cuando todo el espesor del manto se desliza rápidamente pendiente abajo. Se producen tanto en nieve seca como húmeda, pero siempre en terreno muy liso, con poca rugosidad.	-Los de nieve húmeda tienen que ver con temperaturas altas e insolación, o con lluvia y/o niebla -Nevada abundante sobre suelo caliente
Favorable	Ningún problema que destacar. Previsiblemente sin aludes.	-A menudo esto se produce cuando después de un período de fusión las temperaturas vuelven a bajar y la fusión superficial de la nieve no es suficiente para provocar aludes.

Tabla 1: Descripción de las situaciones de aludes existentes y condiciones meteorológicas asociadas.
Fuente: (Centre de Lauegi, Conselh Generau d’Aran, 2016)

2.2 Terreno de avalanchas

La rama norte de la carretera C-28 presenta actividad avalanchosa desde el pueblo de Baqueira (1500 m) hasta el puerto de la Bonaigua (2072 m). Se trata de 9 Km de trazado en dirección NO-SE a través del valle de Ruda. La vegetación que tapiza esta zona está constituida por arbustos en las zonas más bajas y según se gana altura pasa a bosque de pinos (*Pinus uncinata*) que comienza a 1750 m y continúa por encima de los 2000 m. Las zonas de salida de los aludes se emplazan entre los 1700 y los 2300 m y se caracterizan por la presencia de praderas alpinas y árboles dispersos. Las laderas en esta zona tienen pendientes medias de 30° alcanzando algunas los 45° de inclinación. En esta vertiente la carretera atraviesa un total de 20 trayectorias de avalanchas que presentan un periodo de retorno inferior a un año (Fig. 5). La rama sur de la carretera se extiende desde el puerto de la Bonaigua (2072 m) y baja hasta Casa Sastrada (1400 m). El recorrido tiene una extensión de 11 Km y sigue el trazado del valle del río Bonaigua en dirección NW-SE. En las partes más elevadas de la ladera la vegetación consiste básicamente en praderas alpinas y solo en la zona más meridional aparece bosque caducifolio. Las laderas presentan pendientes desde 30° hasta más de 45° pero la mayor parte tienen una inclinación de entre 35° y 45°. La carretera en este tramo atraviesa un total de 28 zonas de avalanchas, de las

cuales algunas presentan periodos de retorno anuales y otras, mucho más elevados (Fig. 5). La experiencia acumulada por el centro de predicción local durante los años que lleva en activo les ha permitido elaborar una cartografía más detallada de las zonas de salida que la que se recoge en el mapa de zonas de aludes de ICGC (Fig. 4). Como se ve en la Fig. 4 algunas de las zonas de arranque cartografiadas se han dividido teniendo en cuenta el comportamiento de las mismas. De la misma forma, dentro de los 20 Km de puerto se han diferenciado siete sectores con una dinámica de avalanchas distinta. Cada sector comprende varias zonas de avalanchas con un comportamiento similar codificado en gran medida por la topografía. En el mapa de la Fig. 5 se pueden distinguir estos sectores, los nombres que se les han dado y su distribución en el puerto de la Bonaigua.

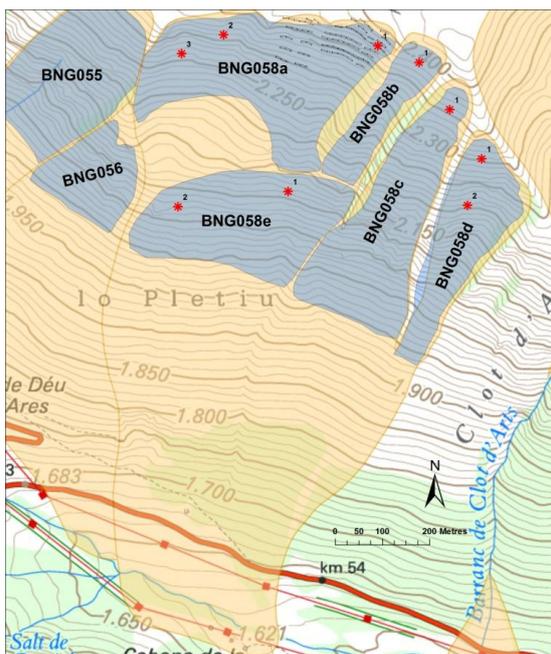


Fig. 4: Zonas de salida de aludes en detalle (polígonos azules) sobreimpuestas a la cartografía del ICGC (polígonos naranjas). Sobre cada polígono azul aparece el número de identificación del mismo basado en el código establecido en la cartografía de zonas de aludes del ICGC. Los asteriscos representan tiros de Daisybell. Sector Comials. Elaboración: Centre de Lauegi, Conselh generau d'Aran

3 Metodología

Como se indicaba en el apartado 1.3 del trabajo, el principal objetivo es desarrollar un modelo predictivo estadístico a partir de las variables meteorológicas recopiladas en el período que lleva activo el centro de predicción local. La herramienta estadística elegida para tal fin es un árbol de clasificación desarrollado en la obra “Classification and regression trees” (CART) de Breiman (1984). Esta herramienta destaca frente a otros métodos de análisis por su manejo sencillo y por la forma gráfica en la que se representan los resultados. Esto simplifica el procesado y proporciona un análisis de los resultados más intuitivo que otros métodos (Trujillo et al., 2008). Por todo ello se facilita su comprensión a usuarios con una experiencia limitada en este campo. La obtención de modelos estadísticos de predicción ha constituido un objetivo muy buscado por los nivólogos para su aplicación en operaciones de gestión de avalanchas y también en otras ciencias aplicadas como la medicina. Los estudios comparativos entre los árboles de clasificación y otros métodos como la regresión logística (RL) son frecuentes. Las tasas de mala clasificación (TMC) de CART suelen ser mayores que los de la regresión logística dando mejores resultados la regresión logística si se trata de muestras con variables con poca correlación y poblaciones pequeñas. Sin embargo si se trabaja con poblaciones mayores, entre 50 y 100 observaciones, y variables con alta correlación, los índices de TMC disminuyen en los modelos de CART hasta acercarse a los resultados obtenidos mediante la RL (Serna, 2009). Otros trabajos estadísticos comparativos entre métodos estadísticos como Díaz (2012) recomiendan el uso del CART frente a la regresión en el caso de desconocer la forma funcional

del modelo ya que se obtienen mejores resultados. Este caso, como señala el propio autor, es frecuente en investigadores con una experiencia limitada en análisis estadístico.

Muchos son los trabajos publicados en las últimas décadas con el objetivo de desarrollar un modelo estadístico que facilite la toma de decisiones en operaciones de gestión del medio nival. Las herramientas de análisis estadístico que han elegido los autores son variadas: análisis discriminante multivariante (Bois et al., 1975), vecino más próximo (Buser, 1983)(Purves et al., 2003), análisis canónico discriminante y análisis de varianza (ANOVA) (Floyer and McClung, 2003), árboles de clasificación y regresión (CART) (Davis et al., 1992) (Davis and Elder, 1994) (Davis et al., 1999) (Jones and Jamieson, 2001) (Rosenthal et al., 2002). La herramienta CART ha sido utilizada más recientemente con buenos resultados en importantes carreteras de Nueva Zelanda y Alaska para generar modelos predictivos de aludes para el conjunto del trazado (Hendrikx et al., 2005) (Murphy and Hendrikx, 2014). En el caso de este trabajo, ya que queremos implementar un modelo de predicción diario, es necesario generar una base de datos en la que se reflejen las variables meteorológicas y los eventos de avalanchas del registro. Para generar esta base de datos ha sido necesario llevar a cabo un tratamiento de los datos de partida, tanto meteorológicos como de avalanchas. Este proceso se describe a continuación.

3.1 Datos meteorológicos

- **Fuentes de datos y su elección**

El centro de Lauegi elabora su boletín de peligro de aludes atendiendo a tres fuentes de información nivometeorológica: estaciones automáticas, datos NIMET tomados por los pisteros de la estación de Baqueira y perfiles estratigráficos y observaciones en distintos puntos del territorio aranes realizados por los técnicos del Centro de Lauegi. Actualmente disponen de cuatro estaciones automáticas que les proporcionan distintos parámetros meteorológicos. Como se aprecia en el mapa de la Fig. 5 se encuentran dispuestas en el entorno del puerto de la Bonaigua. Dentro de estas cuatro estaciones se pueden distinguir dos tipos: estaciones automáticas y flowcapt. Las primeras toman medidas de velocidad y dirección del viento, humedad, radiación solar y espesor de nieve y además incorporan un pluviómetro que proporciona datos de precipitación líquida. El segundo tipo mide la velocidad y dirección del viento y parámetros de transporte de nieve por el viento. Las características de cada tipo se reflejan en la tabla 2. En el mapa de la Fig. 5 se puede observar que solo se dispone de un flowcapt situado en Comalada y que las otras tres estaciones son del primer tipo. Los datos que aportan ambos tipos de estaciones automáticas son de carácter horario y excluyendo averías o fallos se dispone de un registro continuo desde la temporada 2003-2004 hasta la actualidad, a excepción de la estación de Cap del Port que fue instalada más tarde. Los datos NIMET tomados por los pisteros de la estación de esquí de Baqueira incluyen información sobre la nieve en superficie (penetración del primer tramo de sonda, densidad, tipo de cristales de nieve y su tamaño) y también datos meteorológicos como temperatura del aire y meteoros. Las parcelas donde se toman estos datos se sitúan en el dominio esquiable de la estación, también cercanas al puerto de la Bonaigua. A diferencia de los datos tomados con las estaciones automáticas, de estos no se dispone de un registro continuo. Hay lapsos de tiempo importantes en los que no se han podido tomar estos datos debido a grandes nevadas que impiden salir a los pisteros.

- **Procesado de los datos. Variables primarias**

Para llevar a cabo un análisis estadístico diario es necesario partir de una fuente de datos que cumpla este requisito. Como se ha explicado en el apartado anterior, de las fuentes disponibles solo las estaciones automáticas de Comalada y Bonaigua y el Flowcapt disponen de un registro continuo desde el comienzo de las operaciones en la temporada 2003-2004. La elección de las fuentes se ha basado en el conocimiento acumulado sobre el comportamiento de las estaciones y su disposición respecto al tramo de carretera estudiado. La estación de Bonaigua está emplazada en un pequeño valle cerrado en dirección NO-SE lo que implica que la dirección del viento que

mide está muy influenciadas por la orografía. Debido a esto los parámetros de velocidad y dirección del viento así como los de transporte de nieve se han tomado del Flowcapt. En un primer momento, se decidió tomar el resto de parámetros de la estación automática de Comalada. El emplazamiento de la misma era mucho más representativo de las condiciones en las zonas de salida del puerto de la Bonaigua. Finalmente esta opción resulto no ser viable debido a la falta de consistencia de los datos, que incluían errores y lapsos sin datos. Tras descartar esta vía se decidió tomar los parámetros restantes de la estación de Bonaigua.

DATOS	BONAIGUA 2266 m	COMALADA 2153 m	CAP DEL PORT 2250 m	FLOWCAPT 2100 m
Temp. Mínima	✓	✓	✓	✓
Temp. Media	✓	✓	✓	✓
Temp. Máxima	✓	✓	✓	✓
Humedad rel.	✓	✓	✓	
Radiación solar	✓	✓	✓	
Dirección viento	✓	✓	✓	
Vel. med. viento	✓	✓	✓	✓
Vel. min. viento	✓	✓	✓	✓
Vel. máx. viento	✓	✓	✓	✓
Drift 1m				✓
Drift 2 m				✓
Precip. líquida	✓	✓	✓	
Grueso n. total	✓	✓	✓	
Grueso n. 24 h	✓	✓	✓	

Tabla 2: Tipos de datos tomados por las estaciones automáticas.

- Flowcapt: Esta estación proporcionaba datos de velocidad máxima, velocidad media, dirección media del viento y transporte de la nieve por el viento a 1 metro sobre el suelo y de entre 1 m y 2 m. Como los datos que aporta son horarios fue necesario calcular los valores de estos mismos parámetros para intervalos diarios.
- Bonaigua: La titularidad de esta estación es del Servei Meteoròlogic de Catalunya (SMC). Al igual que el instrumento anterior, este también proporciona un registro horario. Sin embargo, en este caso, se solicitaron los datos ya tratados al ICGC y al SMC donde los datos horarios ya habían sido calculados. De esta fuente se utilizaron los datos de temperaturas medias, temperaturas mínimas, temperaturas máximas, oscilación térmica diaria, precipitación líquida equivalente, radiación solar, grueso total de nieve y nieve reciente diaria.

• **Procesado de los datos. Variables derivadas**

Una vez obtenidos los datos de las variables primarias para los años 2003-2015 se procedió al cálculo de las variables derivadas. Estas nuevas variables incluyen tanto datos continuos como categóricos. En el anexo 1(disponible en [enlace](#)) aparecen todas las variables que se han utilizado en el análisis, primarias o derivadas. Se incluyen la forma de cálculo, el nombre corto de cada variable y su procedencia. El grueso de las variables derivadas corresponde a sumatorios y medias que se han calculado para las temperaturas, humedad relativa, radiación solar, precipitación líquida y espesor de nieve. Estos cálculos se efectúan para intervalos de 24, 48 y 72 horas. El cálculo de los valores de las variables para intervalos previos es muy frecuente en la bibliografía. En trabajos como Hendriks et al. (2012) los intervalos elegidos tiene un

máximo de tres días, sin embargo, algunos autores como Peitzsch et al. (2012) los extienden hasta los seis días para el análisis de avalanchas de nieve húmeda. En el caso de esta operación la elección de periodos más cortos presenta mayor utilidad ya que el principal objetivo es implementar un modelo general en el que se incluyen todas las modalidades de avalanchas. Al tratarse de un modelo general se han introducido diferentes tipos de variables que puedan resultar significativas tanto para avalanchas húmedas como para secas. En el trabajo Baggi and Schweizer (2009) construyen variables categóricas dicotómicas denominadas “días de”. Este tipo de variables se ha aplicado en este trabajo para definir los días de lluvia, nieve u observación de avalanchas los días previos. Otro conjunto de variables muy utilizado en el modelizado de avalanchas de nieve húmeda son las variables de tendencia. Este tipo de parámetros reflejan la diferencia positiva o negativa entre días consecutivos. Esto se ha aplicado a humedad, temperatura y radiación tomando de ejemplo el trabajo Floyer and McClung (2003) y Baggi and Schweizer (2009).

Las avalanchas provocadas por placas de viento están relacionadas con otro conjunto de parámetros denominados “drift” que expresan la cantidad de nieve transportada por el viento. Como se indicó en la descripción de la estación flowcapt, esta aporta datos sobre esta variable, no obstante se han calculado nuevas variables “drift” a partir de los datos de velocidad del viento, precipitación y nieve nueva. Debido a la alta incidencia de accidentes asociados a la formación de estas placas de viento la bibliografía centrada en el estudio de estos parámetros es extensa y son muchas las fórmulas y métodos de cálculo de este tipo de variables. A modo exploratorio hemos incluido los métodos de varios autores para determinar cual tiene mayor significación y a su vez compararlos con los datos aportados por la estación Flowcapt. El método de cálculo concreto para cada uno de estos parámetros se indica de forma precisa en el anexo 1 (disponible en [enlace](#)). De manera general estas variables se obtienen a partir del producto de un parámetro de velocidad del viento y otro de precipitación. En Hendrikx et al. (2012) se usa la velocidad media del viento multiplicada por la precipitación líquida equivalente. En otros trabajos como Davis et al. (1999) en lugar de la precipitación líquida equivalente se utiliza el espesor de nieve nueva. En la formación de placas de viento la velocidad del viento es un factor crítico, por esto muchos autores calculan un conjunto distinto de variables en las que la velocidad es elevada a un exponente con el objetivo de dar peso a esta frente a la precipitación. El valor del exponente es diferente según los trabajos, Hendrikx et al. (2012) utiliza un exponente al cubo y Davis et al. (1999) uno a la cuarta. Además de la velocidad del viento, la dirección es clave en la formación de placas de viento asociadas a los sotaventos del relieve. Acotar los parámetros “drift” según distintas direcciones puede revelar información sobre los flujos preferentes que causan las placas de viento. En el trabajo Obled and Good, (1980), los autores segregaron las direcciones diarias del viento en ocho clases preferentes para un periodo acumulado de tres días consecutivos. Esta práctica ha sido aplicada posteriormente por Davis et al. (1999) pero en este caso el periodo acumulado es solo de un día. Este último método es el elegido para este trabajo. De forma resumida las variables “drift” que se han incluido en este análisis son cinco. De las dos variables calculadas a partir de la nieve reciente, en una se ha dado peso a la velocidad del viento elevándola al cubo. De la misma forma se ha procedido con las dos variables calculadas a partir de la precipitación líquida equivalente. El conjunto de estas variables “drift” a excepción de las segregadas por dirección han sido calculadas para periodos acumulados de 24, 48 y 72 horas al igual que en los trabajos Davis et al. (1999) y Hendrikx et al. (2012).

3.2 Datos de avalanchas

- **Registro de avalanchas**

Los aludes ocurridos en el puerto de la Bonaigua desde el invierno de 2003 hasta la actualidad están recogidos en una cartografía GIS de alto detalle (Fig. 6). Cada evento lleva asignado un código de identificación, código de la zona de avalancha donde se ha producido, fecha de caída, notas sobre el suceso, fotografías y fecha de toma de las mismas.

- **Definición día de avalanchas**

El análisis de predicción en las operaciones de control de avalanchas suele realizarse en términos dicotómicos, clasificando los días en días de avalancha (AVA) y días de no avalanchas (NO AVA). Otros trabajos incluyen más de dos categorías (Murphy and Hendriks, 2014) con el objetivo de afinar más la situación de aludes. La definición de día de avalanchas es uno de los pasos más importantes en este tipo de análisis. Es necesario conocer la dinámica avalanchosa de la zona de estudio para establecer un criterio certero que seleccione aquellos eventos potencialmente peligrosos para el trazado y que excluya los de entidad inferior para evitar la contaminación de la muestra. Estos criterios son diferentes según la zona de estudio, el objetivo del trabajo y las preferencias de los autores. Algunos autores han utilizado el número de eventos de avalancha diarios (Davis and Elder, 1994), la suma del tamaño de las avalanchas (Davis et al., 1999), umbrales de peligro para cada trayectoria de avalanchas basados en el criterio del experto (Floyer and McClung, 2003), porcentaje de trayectoria recorrida por la avalancha (Hendriks et al., 2012) o según los efectos potenciales sobre la carretera (Hendriks et al., 2005) entre otros.

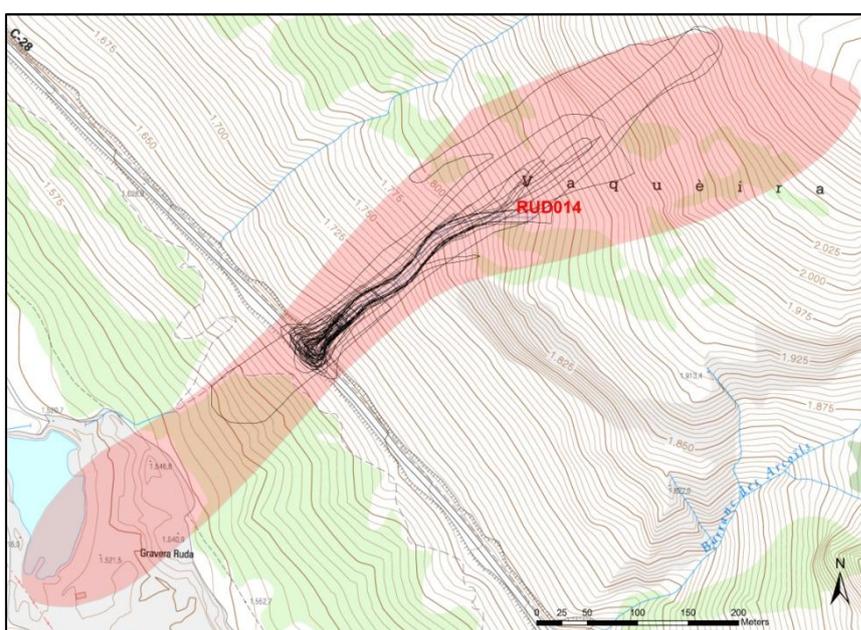


Fig 6: Cartografía GIS de detalle de los eventos de avalanchas en una de las zonas de aludes del puerto de la Bonaigua. El polígono rosa representa la cartografía de zonas de aludes del ICGC. Los polígonos contorneados en negro representan cada uno de los eventos registrados en esta zona desde 2003. Elaboración: Centre de Lauegi, Conselh generau d'Aran.

Para este estudio la selección se ha centrado en la afectación a la carretera. A partir de la cartografía SIG se han seleccionado aquellos eventos que cortaban al trazado. Al aplicar solo este criterio, quedaban excluidos gran cantidad de eventos que aunque no habían alcanzado la carretera si tenían una entidad suficiente como para causar problemas. Por ese motivo se decidió seleccionar además todos aquellos eventos cuyo depósito se situara a menos de 20 m del trazado. Esta distancia se decidió después de probar selecciones con otras medidas mayores y menores. Resulto ser esta distancia la que mejor seleccionaba los eventos con entidad suficiente para su inclusión en la muestra y la más conforme con el criterio de los expertos del centro. Hay que recordar que la gran mayoría de aludes que afectan a la carretera anualmente son de tamaño reducido (D2 o D3) pero al estar el trazado situado a mitad de ladera cualquier evento por pequeño que parezca puede causar el cierre de la carretera. Como no se trata de trayectorias de grandes dimensiones, la aplicación de los criterios expuestos anteriormente resultaría poco adecuada, ya que están diseñados para grandes zonas de aludes y quedarían excluidos muchos aludes, que aunque de poca entidad, constituyen un problema relevante para la carretera C-28 a su paso por el puerto. De esta manera cada día en el que se hubiera producido al menos un evento de alud seleccionado según los criterios elegidos se ha clasificado como día de avalanchas. Después de este proceso resultaron un total de 177 días de avalanchas. Siguiendo el trabajo Floyer and McClung (2003), los sectores en los que se dividía el puerto se han

clasificado en tres subzonas (Fig. 5). En cada una se han incluido los sectores con dinámicas avalanchosas similares. En la tabla 3 se especifica que sectores se incluyen y las características de su funcionamiento. Con este análisis pormenorizado se pretende, a modo de exploración, determinar si hay variables para las subzonas ocultas en el modelo general; a pesar de que al reducir la población de las muestras se puede disminuir la solidez de los resultados.

SUBZONAS	SECTORES	CARACTERÍSTICAS	POBLACIÓN
1	Glideland + corbes	Deslizamientos basales	165 días
2	Entre barrancs + Recta Ruda	Avalanchas de nieve húmeda y nieve reciente	200 días
3	Cap del Port	Especial incidencia de placas de viento	78 días

Tabla 3: Subzonas, sectores que las componen, comportamiento avalanchoso frecuente y población de estudio. Se puede ver su situación en el mapa de la Fig. 5.

3.3 Construcción de la base de datos

Una vez obtenidas las variables de toda la serie desde el invierno de 2003 hasta 2015 y también los días considerados de avalancha, se cruzan ambos datos. A partir de aquí se define la serie temporal que definirá el modelo. Se ha decidido utilizar los datos de las temporadas invernales 2004-2005 hasta 2014-2015 comprendiendo cada temporada los meses de noviembre a abril, ambos incluidos. El resto de meses de cada año se eliminan ya que la innivación no es significativa o inexistente. De esta base de datos donde se asocian días con variables es necesario eliminar algunos atendiendo a diferentes criterios.

En las observaciones de los fenómenos avalanchosos en el puerto y su posterior registro es frecuente que debido a los cortes en la carretera o a inestabilidad atmosférica no sea posible realizar observaciones diariamente. En periodos de varios días, cuando ocurra esto, mediante las observaciones posteriores de los depósitos de avalanchas en la carretera es imposible distinguir si se trata de uno o varios eventos y su distribución a lo largo de esos días. Por ello no se pueden clasificar con total seguridad estos días, en días de avalancha y días de no avalancha. Para evitar incluir errores en esta base de datos se han eliminado los días que responden al criterio explicado en la tabla 4. A partir de un registro de los cortes de carretera C-28 en el puerto de la Bonaigua se ha decidido excluir aquellos días en los que la carretera permaneciera cerrada y no haya registro de aludes. La imposibilidad del acceso no permite determinar con seguridad si son días de no avalanchas, por lo tanto se sacan del modelo.

avalanchas	carretera	día
Si	Abierta	Bueno
No	Abierta	Bueno
Si	Cerrada	Bueno
No	Cerrada	Malo

Tabla 4: Días eliminados (malos) o conservados (buenos) al cruzar los días de avalancha con los cortes de la carretera.

Una revisión uno por uno de los días de la serie temporal ha permitido descartar aquellos días que carecen de datos o incluyen errores en un número significativo de las variables.

En los días clasificados como de avalanchas, la discusión sobre la inclusión en el modelo de los días de desencadenamiento artificial es frecuente en la bibliografía. En el artículo de Hendrikx et al. (2005) los autores son partidarios de incluir los días con desencadenamiento artificial. En la zona de estudio de este trabajo las operaciones de control se realizan de forma regular tras los episodios de nevadas y en ocasiones durante las mismas (Peitzsch et al., 2012). En el caso concreto del puerto de la Bonaigua hay que recordar que los desencadenamientos artificiales se llevan a cabo mediante DaisyBell® transportada por un helicóptero. Esto implica que las operaciones de control no son tan regulares y además solo se realizan en condiciones de buen

tiempo que es cuando puede volar el helicóptero. Según Floyer and McClung (2003) este tipo de días debe eliminarse de la serie. El resultado ha sido un total de 22 días en los que se habían llevado a cabo desencadenamientos artificiales y que han sido eliminados de la muestra.

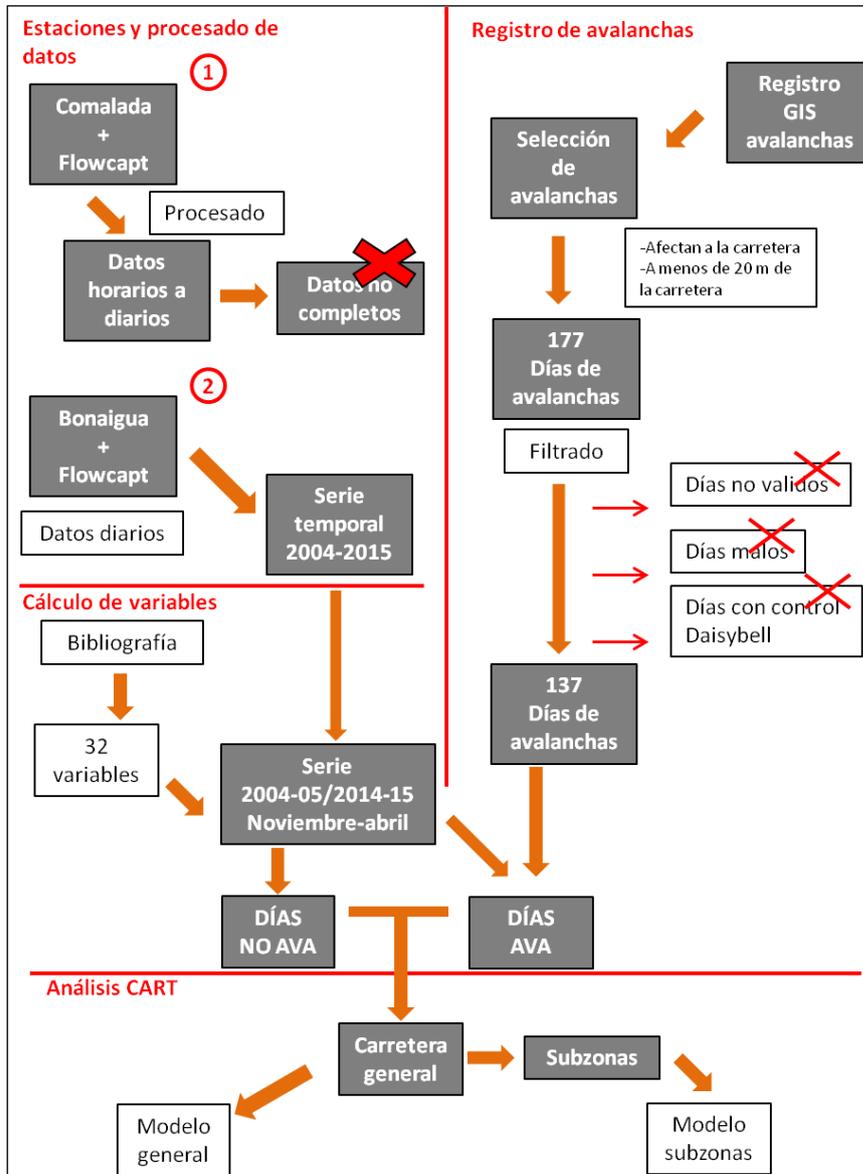


Fig. 7: Diagrama del procesamiento de los datos hasta el análisis con CART.

Después de todo el filtrado de los días de la serie se han obtenido un total de 137 días de avalanchas. El último paso para terminar la serie de datos que será analizada para desarrollar el modelo general, es seleccionar aleatoriamente de los días clasificados como de no avalanchas un número igual al anterior. Al introducir en el software de análisis una muestra con un número igual de días de ambos tipos las dos clases tendrán el mismo peso. En este proceso de selección hay que cerciorarse de que los días hayan sido elegidos de forma uniforme a lo largo de las temporadas y de la serie completa para evitar dar peso a unas temporadas sobre otras. Los días seleccionados se han revisado para evitar que se incluyeran días que presentan actividad avalanchosa aunque estos no hayan sido incluidos en la selección por no coincidir con el criterio propuesto. En el diagrama de la Fig. 7 se representa de forma sintética el proceso de tratamiento de los datos y de elaboración de la muestra para el modelado de la carretera.

3.4 Análisis estadístico

Para elaborar los árboles de clasificación se ha elegido el software estadístico CART de Salford Systems® basado en el algoritmo desarrollado en la publicación de Breiman (1984). CART se trata de un árbol de decisión. Se estructura en ramas que representan conjuntos de decisiones que se generan sucesivamente para continuar con la partición y formar de esta forma grupos homogéneos respecto a la variable a discriminar, en nuestro caso AVA/NO AVA. Se trata de un método no-paramétrico de segmentación binaria en el que las particiones se realizan de forma recursiva hasta que se alcanza un criterio de parada. A partir de los datos históricos se construye el árbol de decisión que se utilizará para clasificar nuevos datos. La estructura del árbol consiste en un nodo inicial llamado raíz o *madre* y que se divide en dos grupos *hijos* que a su vez se les aplica otro procedimiento de partición por separado. Las divisiones se seleccionan de forma que la impureza de la muestra del modo resultante sea inferior a la del nodo madre. El objetivo consiste en partir la muestra en grupos homogéneos y mantener un tamaño de árbol razonablemente pequeño. Este proceso de análisis consiste en tres pasos fundamentales: Construcción de un árbol máximo o *sobredesarrollado*, poda del árbol y selección del árbol óptimo mediante un procedimiento de validación cruzada.

- **Construcción del árbol máximo**

Este árbol inicial se trata de un modelo que describe el conjunto de entrenamiento, el conjunto de datos original y que normalmente está *sobredesarrollado*. En este primer paso el crecimiento del árbol continúa hasta que se produce cualquiera de las siguientes posibilidades: Solo hay una observación en cada nodo hijo, es imposible establecer el criterio de máxima pureza ya que todas las funciones tienen la misma probabilidad asignada en los nodos terminales o se ha establecido un límite de desarrollo previo al análisis. Hay una gran cantidad de niveles y nodos que no producen una mejor clasificación y puede ser demasiado complejo. El árbol que se ha generado clasifica bien los registros utilizados para su desarrollo pero no se garantiza su capacidad de generalización cuando se enfrenta a nuevos registros. Para determinar la mejor variable en cada partición se utiliza una función de impureza. De esta forma el programa selecciona aquellas variables que nos aportan una división en la que los nodos hijos presentan una pureza máxima. La variable que consigue una mayor pureza se utilizara en primer lugar. Existen distintas formas de medir la impureza que nos permiten analizar las respuestas del árbol. En este caso se ha elegido el índice Gini, uno de los más extensamente utilizados.

Este índice tiende a separar la categoría más grande en un grupo aparte, al contrario que otros indicadores. La representación matemática de este índice es:

$$G = 1 - p_1^2 - p_2^2$$

Donde p_1^2 son el porcentaje de población que pertenece al grupo 1, por ejemplo AVA y p_2^2 el porcentaje de muestra que pertenece al grupo 2, NO AVA. El peor escenario que se puede presentar en una partición es que el nodo hijo presente un valor de Gini de 0,5 lo que significaría que la proporción de ambas clases es de 50/50. Si una partición ha sido óptima el valor se acerca a 0 lo que determina que una de las clases predomina sobre la otra.

- **Poda del árbol**

El proceso de poda del árbol *sobredesarrollado* consiste en la eliminación sucesiva de ramas y nodos hasta la obtención de un tamaño adecuado y con capacidad de generalizar su uso en otros conjuntos de datos. Para determinar cuáles de estos nodos se deben eliminar se utiliza una función coste-complejidad. Esta función valora la complejidad en número de nodos y velocidad de procesamiento frente a la precisión que aporta en la clasificación. Los árboles que presenten menor valor de esta función serán los buscados.

- **Elección del árbol óptimo**

Tras el proceso de poda del conjunto de árboles resultantes es necesario seleccionar el mejor. El criterio que se impone en este caso es elegir aquel árbol que presente un menor porcentaje de error en la clasificación de la muestra utilizada para su implementación. Es necesario a su vez llevar a cabo una comprobación con un conjunto de muestras no utilizado para el desarrollo del modelo. En este caso se ha utilizado el método de validación cruzada que no precisa de un conjunto de prueba independiente. Este sistema de revalidación del modelo divide en K partes la muestra de datos antes del desarrollo del modelo (normalmente K=10). De esta manera cada una de estas K partes se reserva en análisis consecutivos como conjunto de prueba. El error de predicción obtenido para estos K modelos es acumulado para obtener el error medio absoluto del conjunto de prueba.

4 Resultados

4.1 Árboles del modelo general

Para el conjunto de la carretera se han obtenido dos árboles. El primero de ellos se realizó introduciendo en el software todas las variables calculadas. Este primer procesado *grosso modo* nos permite observar que variables son relevantes y cuales son descartables. De esta forma se puede afinar el modelo añadiendo solo aquellas más representativas para alcanzar un árbol más preciso. En el Anexo 2 (disponible en [enlace](#)) se muestran los resultados de este primer análisis y la importancia de las variables.

En esta primera fase se obtuvo un árbol de gran tamaño con un elevado número de nodos que, como se ha especificado en la sección 3.4, no es recomendable para su puesta en práctica. El porcentaje de aciertos en los test de comprobación mediante validación cruzada atribuyen una media de aciertos a este modelo de 75,70 %. En el caso de los días de avalanchas es capaz de predecir correctamente un 80,29 %. El porcentaje de aciertos desciende a un 71,11 % en el caso de los días de no avalanchas. En la Fig. 8 se pueden observar las primeras particiones de este árbol. La primera variable que utiliza el modelo es *sumsnow_72*, como se describe en el Anexo 1 (disponible en [enlace](#)) se trata del sumatorio de nieve nueva en las últimas 72 horas. El valor de corte es 18,10 cm, esto nos indica que en periodos de 72 horas en los cuales se supere este espesor de nieve nueva serán clasificados como días de avalanchas. Aquellos días que no superen estos valores quedarían incluidos en el nodo de la izquierda, donde se procede a una segunda partición. En esta ocasión la variable utilizada es *total_snow* que corresponde al espesor total de nieve acumulada. El valor de partición 155,35 cm clasifica todos aquellos días con un espesor superior como días de avalanchas y aquellos con valor inferior como días de no avalanchas.

Desde un punto de vista nivológico, la elevada importancia estadística del sumatorio de nieve para 72 horas se corresponde con la dinámica de aludes del puerto. Un periodo acumulado de nevadas en el que se alcanza un valor crítico de nieve nueva se relaciona directamente con una mayor incidencia de avalanchas. Por otro lado, a pesar de que el espesor total de nieve en las zonas de salida de aludes es de gran importancia para la predicción, el parámetro *total_snow* no corresponde exactamente a este significado ya que los valores no se han tomado en las propias cabeceras de las zonas de aludes. Teniendo en cuenta que este árbol tiene como objetivo gestionar los cierres de la carretera, no es posible utilizarlo ya que clasifica como día de avalanchas cualquier día con un valor superior a 155 cm. Este espesor total se alcanza anualmente en la estación automática de la Bonaigua cuando la temporada se encuentra avanzada. En este caso, lo único que nos está indicando es en qué momento de la temporada nival estamos. Es decir, la actividad de aludes será más intensa cuanto más avanzada este la temporada y mayor espesor acumulado de nieve haya en la montaña. A esta misma conclusión llegaron los autores Floyer and McClung (2003). En el modelo que desarrollaron el espesor total de nieve presentaba una gran relevancia estadística. Sin embargo afirmaron que esta variable podría sustituirse por otra que expresara los días transcurridos desde el inicio de temporada.

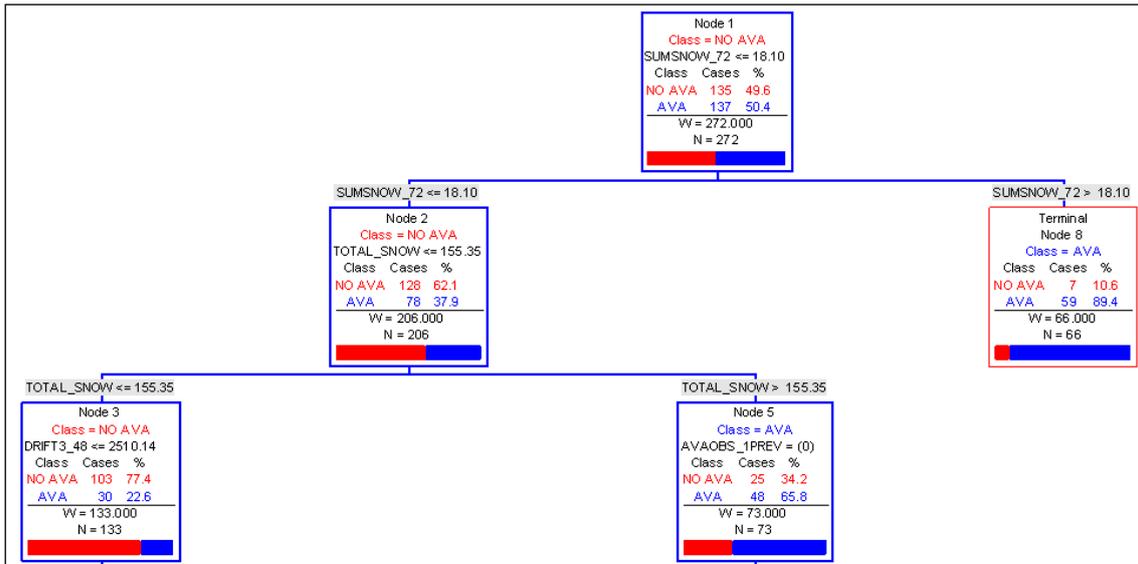


Fig. 8: Primeras particiones del árbol obtenido a partir del análisis conjunto de todas las variables. Cada cuadro representa un nodo y dentro del mismo están reflejadas sus características, porcentaje de días de avalancha (AVA) frente a no avalancha (NO AVA), la variable de partición y su valor de corte y la población (N).

Basándonos en los resultados obtenidos en este primer análisis, se excluyeron aquellas variables no significantes o con una importancia baja, además del parámetro *total_snow*. Con las restantes se probaron distintas combinaciones, incluyendo y excluyendo las variables en análisis sucesivos hasta obtener un modelo con mayor precisión. El resultado es el árbol de la Fig. 9 (más información en Anexo 2, disponible en [enlace](#)). El porcentaje de aciertos que ha obtenido en los test de validación cruzada es de un 75,74% para el conjunto de los datos, un 80,00% para los días de no avalanchas y un 71,53% para los días de avalanchas. Tanto para este último árbol como para el anterior, se ha utilizado un valor de K=8 para el test de validación cruzada debido a que el número de días usado para el análisis es 272. Dividiéndolo en 8 partes generamos grupos más homogéneos, lo que es fundamental cuando se trabaja con bases de datos de tamaño reducido. La precisión que presenta este árbol es relativamente elevada teniendo en cuenta que se sustenta en una serie temporal no muy extensa. Hay que destacar el menor índice de mala clasificación para los días de no avalanchas (20,00%) frente a los días de avalanchas (28,47%). Estos resultados son relevantes de cara a la aplicación del modelo en la gestión del puerto. Una mayor precisión en la clasificación de los días no avalanchosos es de gran importancia, ya que los falsos negativos pueden generar una mayor problemática que los falsos positivos. Al igual que ocurría en el modelo anterior, la variable responsable de la primera partición es el *sumsnow_72* y además utiliza el mismo valor de corte (18,10 cm). Aquellos días que hayan ido a parar al nodo de la izquierda, con una precipitación en 72 horas inferior a 18,10 cm, se trata de días en los que se han producido nevadas débiles. La siguiente partición la lleva a cabo la variable *drift_24*. Los días con un valor de este parámetro inferior a 25,05 serán aquellos en los que la precipitación de nieve ha sido escasa o nula. Por el contrario los que tengan un valor de *drift_24* mayor a 25,05 presentarán nevadas escasas (inferiores a los 18,10 cm del primer nodo), pero de mayor entidad que el conjunto anterior. Aunque el objetivo de esta variable es representar la cantidad de nieve transportada y por lo tanto ser un indicador de la formación de placa de viento, en este caso la interpretación que se le ha dado es que refleja la cantidad de nieve precipitada. La forma de cálculo es el producto de la velocidad media del viento por la precipitación líquida para 24 horas. Teniendo en cuenta esto y el valor tan bajo de corte (25,05), le atribuimos más peso en el resultado a la precipitación que al viento. El valor de corte propuesto podría alcanzarse por ejemplo con precipitación de 5 cm y una velocidad del viento de 5 m/s, valor demasiado bajo para que se produzca un transporte relevante de nieve. A partir

de aquí el árbol divide en dos ramas cada una de las cuales representa las situaciones de avalanchas que más se producen en el puerto de la Bonaigua. Por un lado, la rama de la izquierda que se desarrolla a partir del nodo 3 (Fig. 9) es característica de la dinámica de aludes de nieve húmeda y glides que se produce muy frecuentemente debido a la orientación preferente de las laderas al sur.

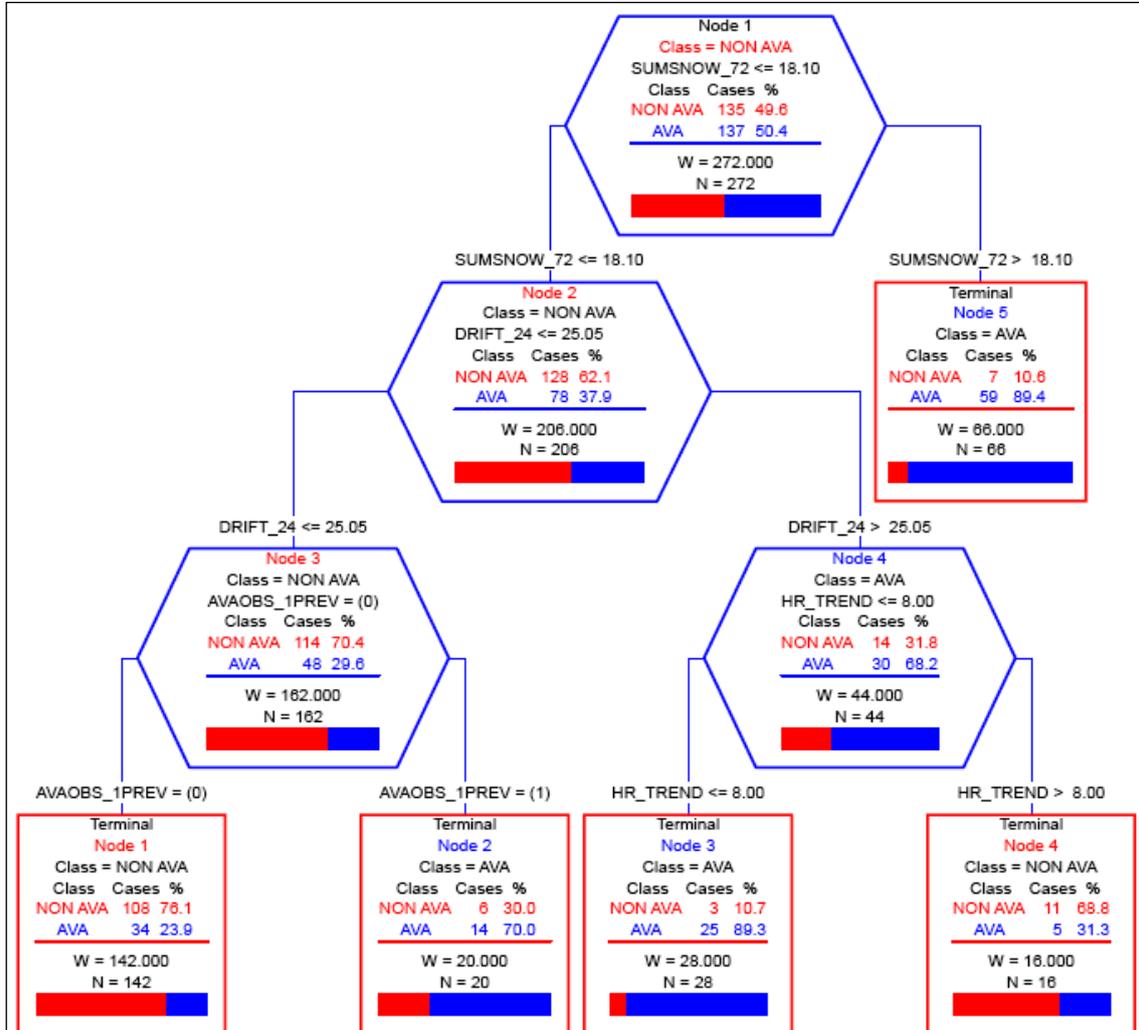


Fig. 9: Árbol de decisión final para el conjunto de la carretera.

Por otro, la rama de la derecha que parte del nodo 4 explicaría la situación de avalanchas de nieve seca. Al dividir el nodo 3 mediante la variable *avaobs_1prev*, que indica si el día anterior se registraron aludes, obtenemos los nodos terminales 1 y 2. En el 1 se incluyen todos aquellos días cuyo valor de la variable es cero, es decir el día anterior no se produjeron avalanchas. Este grupo clasificado como de no avalanchas indicaría que las zonas de salida no presentan nieve debido a la escasa o nula precipitación de los días anteriores, o bien a que las zonas de salida ya se han purgado completamente. Los casos del nodo terminal 2 en los que si se observan avalanchas el día anterior han sido clasificados como días de avalanchas. Los días incluidos en este grupo son aquellos en los que se producen avalanchas de nieve húmeda y deslizamientos basales debido a que las zonas de salida están cubiertas por manto nivoso. Es frecuente que en episodios de estabilidad atmosférica con ventanas de buen tiempo se produzcan en las laderas de Bonaigua aludes de tipo deslizamiento basal que dejan limpias las zonas de avalanchas. Este funcionamiento es el que representa esta rama. El nodo 4 incluye a aquellos días con nevadas débiles pero que tienen mayor entidad que el nodo 3. Para la partición de los casos incluidos en

este grupo se aplica la variable *Hr_trend* que representa la tendencia de la humedad relativa en un periodo de 24 horas. Esta variable separa los días que presentan una disminución de la humedad de los que sufren un aumento respecto al día previo. El significado de este cambio se atribuye a la llegada o a la marcha de una borrasca. Si la humedad aumenta el frente se aproxima, las débiles nevadas registradas son solo el preludio de otras más importantes por llegar, que pueden dar lugar a más episodios de avalanchas de nieve seca (nodo terminal 3). Si por el contrario la humedad baja, el frente, poco activo ya que no ha acumulado más de 18,10 cm (nodo 1) dará paso a un periodo de estabilidad y menos incidencia de avalanchas (nodo terminal 4). Una de las virtudes de este modelo es que representa las distintas fases de un frente, desde su llegada hasta su fin. Las tres ramas del árbol permiten seguir la evolución de la tormenta desde su inicio y responder a los cambios que se produzcan. La combinación de un primer parámetro de partición que responde a un periodo acumulado de tres días con el resto de variables calculadas para un día permite seguir la tendencia a lo largo del tiempo que duren las precipitaciones. A modo de ejemplo, si en las primeras 24 h de inestabilidad se registraran valores de *sumsnow_24* inferiores a 18 cm la predicción se centrara en las dos ramas del nodo 2. Si con el paso de las horas la precipitación de nieve acumulada supera este valor crítico se clasificara como día de avalanchas (nodo terminal 5).

4.2 Árboles sub-zonas

En el caso de estas subzonas (Fig. 5) los test de validación cruzada se han llevado a cabo utilizando un valor de $K=10$. La información sobre el proceso de análisis de cada árbol se encuentra en el Anexo 3 (disponible en [enlace](#))

- **Sub-zona 1**

Esta subzona comprende los sectores de Glideland y Corbes donde la actividad de deslizamientos basales es muy frecuente. En el modelo desarrollado para esta zona (Fig. 10) se han obtenido unos porcentajes de acierto de 87,34 % para los días de no avalancha y 73,23 % para los días de avalancha. Al igual que en los dos árboles anteriores la variable que efectúa la primera partición es el sumatorio de nieve nueva para un periodo de 72 horas. En este caso el valor de corte (14,75 cm) es menor que en los casos anteriores. De nuevo se vuelven a mostrar relevantes variables asociadas a humedad relativa y transporte de nieve. La interpretación que se le ha dado al parámetro *sumhrm_72* es la de indicador de la situación atmosférica al igual que la variable *hr_trend* en el segundo árbol de la sección 4.1. Si esta variable excede el valor de corte de 255,50 se tratara de días con inestabilidad pero en los que las nevadas son muy escasas, inferiores a los 14,75 cm del nodo 1. Días con humedades altas, por encima del 80 %, suelen correlacionarse con días nublados. Si por el contrario los valores del sumatorio permanecen por debajo de ese valor la situación atmosférica será anticiclónica. Como en el segundo árbol de la sección 4.1 la variable “drift” seleccionada nos proporciona más información sobre la cantidad de la precipitación que sobre el transporte eólico. Los días que sobrepasan el valor de corte son días de inestabilidad pero con precipitación escasa. Aunque poco abundantes estas pueden cargar las zonas de salida de avalanchas y causaran los deslizamientos basales una vez pasada la nevada. En los días con nula o muy escasa precipitación estas zonas de salida permanecerán limpias y no se producirán aludes. Como ya apuntábamos en la sección 4.1, el espesor total de nieve en las zonas de salida es una variable de suma importancia para la predicción de aludes. A pesar de que en el registro no disponemos de estos datos, hay variables que revelan esta información aunque su objetivo inicial fuera otro distinto.

- **Sub-zona 2**

En este grupo se incluyen los sectores de Recta de Ruda y Entre Barrancs, ambos caracterizados por los aludes de nieve reciente y nieve húmeda. Como se puede ver en la Fig. 11, la primera variable de partición es el sumatorio de nieve para 72 horas, al igual que en los árboles anteriores. Al contrario de lo que ocurría con las variables “drift” de los casos anteriores, esta representa el transporte de nieve estrictamente.

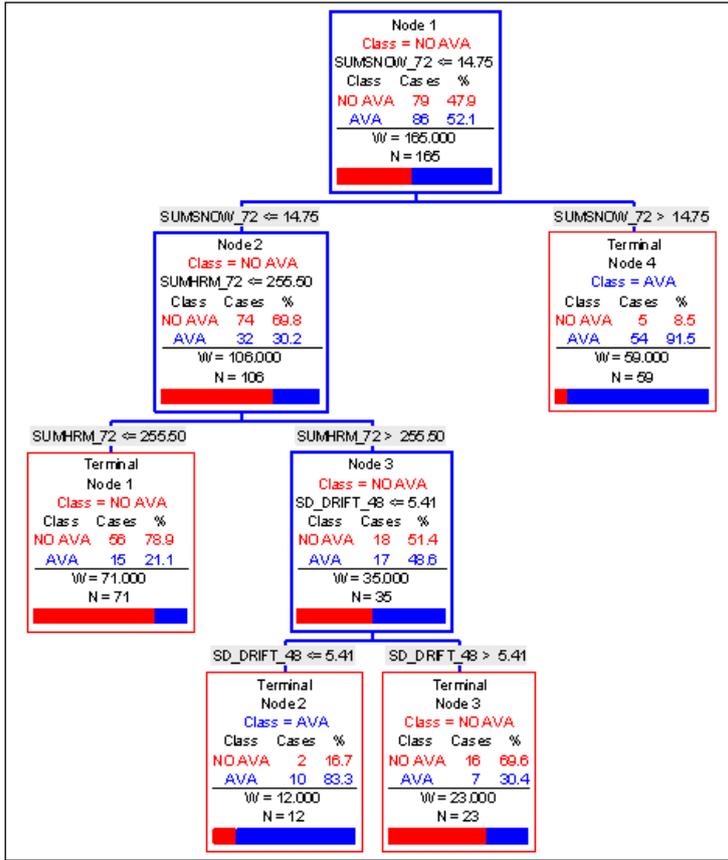
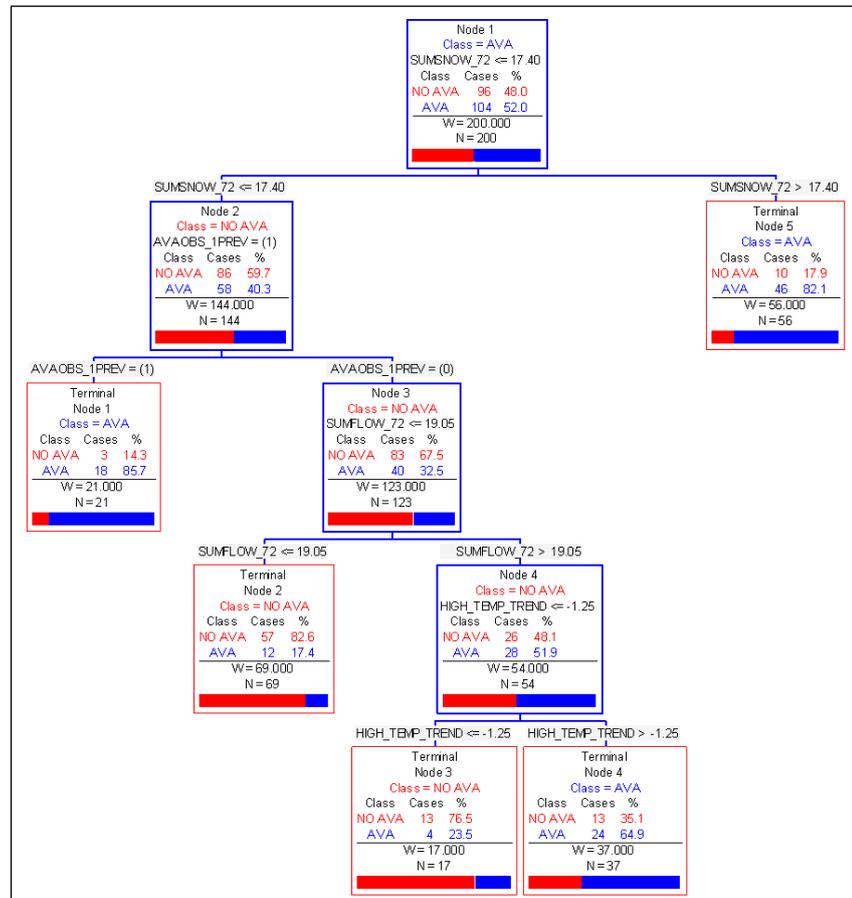


Fig. 10: Árbol de clasificación de la subzona 1. Incluye los sectores de Glideland y Corbes donde predominan los deslizamientos basales (glides).

Fig. 11: Árbol de clasificación de la subzona 2. Dentro están incluidos los sectores de Entre Barrancs y Recta de Ruda. Se caracterizan por la actividad de aludes de nieve húmeda y nieve reciente.



El valor de corte que utiliza (19.05) es bajo, sobre todo si tenemos en cuenta que se trata del sumatorio de transporte de nieve para un periodo de tres días. Esto se justifica porque los días que se incluyen en este nodo presentan nevadas muy escasas ($sumsnow_{72} < 17,40$ cm). Los días con un valor de transporte eólico inferior al valor de corte son clasificados como de no avalanchas y los que superen este umbral darán lugar a la acumulación de las zonas de salida y a la formación de placas de viento. Como el nodo 4 que recoge los días con mayor transporte eólico tiene una proporción muy similar de ambos casos (AVA y NO AVA) se procede a una nueva división en base a la variable *High_temp_trend*. Hasta este momento, en ninguno de los árboles presentados, el modelo había incluido un parámetro de temperatura. Se trata de un indicador de la tendencia de las temperaturas en un periodo de 24 horas. De esta manera si los valores de esta variable para un día se encuentran por encima de -1,25 significa que las temperaturas han disminuido y se clasificara como día de avalanchas. Tras un episodio de nevadas si las temperaturas permanecen bajas en los días sucesivos se impedirá la estabilización del manto reciente favoreciendo que la actividad de aludes se alargue durante un período más largo. Si por el contrario se produce un aumento de las temperaturas se estabilizará el manto y la incidencia de avalanchas bajará.

- **Sub-zona 3**

Esta subzona se diferencia de las otras dos en que la práctica totalidad de los aludes que se producen son de nieve seca. Es especialmente sensible a la formación de placas de viento por la acción del transporte eólico. Las variables que se esperaban relevantes para esta zona eran aquellas que indicaran cantidad de nieve precipitada, la velocidad del viento y parámetros “drift” debido al comportamiento tan diferenciado respecto a los otras sub-zonas. Sin embargo, como se puede ver en la Fig. 12, ambas variables hacen referencia a la temperatura. Los porcentajes de acierto conseguidos por este modelo se asemejan al del conjunto de los árboles expuestos anteriormente con una media de 74,36 %. La explicación que se puede dar a este resultado es que las temperaturas frías se asocian con episodios de nevadas y por tanto con actividad de aludes. Es evidente, como se adelantaba en la sección 3.2, que la merma de la población de estudio ha generado resultados poco confiables. Esta sub-zona estaba compuesta solo por 78 días de avalanchas como se indica en la Tabla 3.

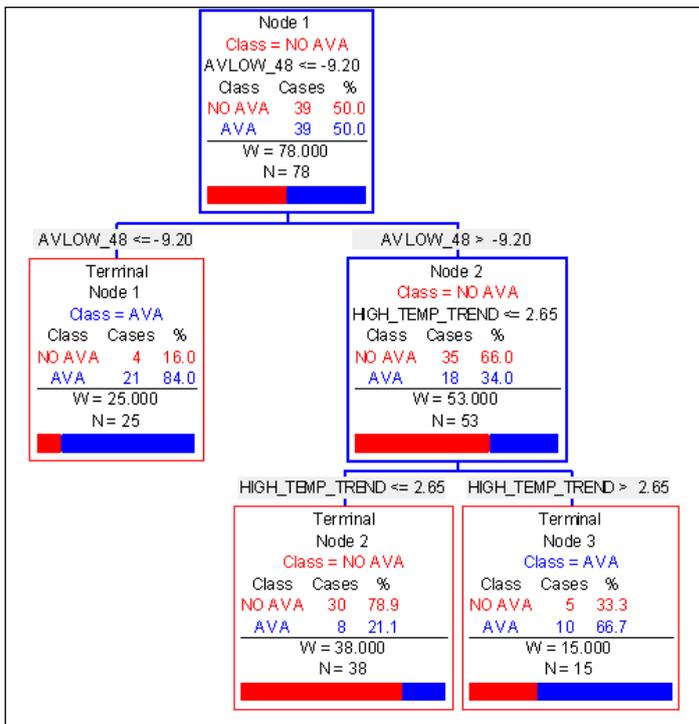


Fig. 12: Árbol de clasificación de la subzona 3 compuesta por el sector Cap del Port caracterizada por los aludes de nieve venteadas.

5 Análisis

La orientación preferente hacia el sur de las laderas que afectan a la carretera C-28 a su paso por el puerto de la Bonaigua y la latitud y altitud en la que se encuentra la zona de estudio, condicionan una dinámica en la que predominan los aludes de nieve húmeda. Como se mencionó en la sección 2.1, los aludes de nieve húmeda están relacionados con la radiación solar, la lluvia y la humedad pero fundamentalmente con la temperatura. Basándonos en esto, se esperaba que el modelo general destacara la relevancia de las variables de temperatura. Sin embargo la importancia de estas ha sido escasa. La explicación que se ha encontrado a esta falta de significancia es que en los días de subida de temperaturas es frecuente que las zonas de salida y las laderas se encuentren limpias de nieve y por lo tanto no se producen avalanchas. Ya se explicó en la sección anterior que debido a la dinámica de deslizamientos basales (tabla 1) es frecuente que las laderas se purguen tras el episodio de nevada con la llegada del buen tiempo y que permanezcan sin o con poca nieve hasta las nuevas nevadas. Esta situación podría explicar el poco peso de las variables de temperatura y los episodios de aludes en el análisis estadístico.

El sector de Cap del Port (subzona 3; Fig. 5) supone una excepción dentro de la dinámica avalanchosa del puerto. En este emplazamiento la mayoría de los episodios se deben a aludes de nieve seca y de nieve venteada. Uno de los objetivos de realizar el análisis por subzonas era analizar de forma más detallada su dinámica. Sin embargo la escasa población para el análisis no lo ha permitido. Las distintas variables “drift” que se calcularon estaban enfocadas en gran medida a su aplicación en esta subzona. De los distintos métodos de cálculo utilizados (Anexo 1, disponible en [enlace](#)) los más significativos para el modelo general han sido los *Drift_XX*, calculados a partir de la precipitación líquida, frente a los calculados a partir del espesor de nieve reciente. También se han impuesto sobre los valores de transporte aportados por el *flowcapt* que no presentaban ninguna relevancia estadística. A pesar de esto, en la interpretación de los árboles de la zona general y de la subzona 1, se ha visto que estos parámetros “drift” responden mejor como un indicador de precipitación que como uno de transporte, especialmente en las divisiones terminales del modelo. El último grupo de variables “drift” segregadas por direcciones del viento pretendía poner de manifiesto la relación entre la formación de placas en Cap del Port y los flujos atlánticos procedentes del N y NW. Estas variables han resultado ser totalmente irrelevantes.

El conjunto de variables que se ha demostrado más significativo son las que expresan el espesor de nieve. En el primer modelo desarrollado para el conjunto del trazado la variable *total_snow* presentaba un muy buen nivel de clasificación, aunque no resultara operativa para los objetivos planteados. En todos los árboles a excepción de la subzona 3 la variable *sumsnow_72* es la responsable de la primera partición. A esto hay que añadir el significado de la cantidad de precipitación sobre las zonas de salida que expresan las variables “drift”. Es evidente la necesidad de una variable que exprese la cantidad de nieve en las zonas de salida y su especial relevancia en la predicción. No en vano el boletín de peligro de aludes diario no se realiza solo en base a los parámetros meteorológicos que aportan las estaciones si no que se incorporan datos de campo sobre el manto nivoso y su espesor.

De las variables categóricas que se han utilizado solo *avaobs_Iprev* ha resultado significativa tanto para el modelo general como para las subzonas. El resto no ha resultado determinante. Esta variable presenta esta relevancia porque confirma que existen las condiciones para que se produzcan aludes. Es decir si ayer se desencadenaron avalanchas es probable que hoy se sigan produciendo bajo las mismas condiciones. Este dato, la ocurrencia o no de avalanchas en días previos, es especialmente crucial para la predicción de deslizamientos basales.

El análisis pormenorizado por sub-zonas no ha conseguido generar modelos particulares distintos al general o revelar variables que difieran de las usadas para el conjunto de la carretera. En el caso de las subzonas 1 y 2, estas presentan un modelo de árbol muy similar al desarrollado para el conjunto de la carretera (Fig. 9). De forma sintética, se comienza la clasificación con una variable que expresa la cantidad de nieve acumulada en las últimas 72 horas. Los días que

superen una altura de nieve crítica son identificados como días de avalanchas y los que no se vuelven a subdividir utilizando, por un lado un parámetro de humedad que indique la evolución de la inestabilidad y por otro un parámetro de “drift” que matice la cantidad reducida de nieve que está cayendo. Esta similitud entre el modelo general y las subzonas 1 y 2 no es de extrañar teniendo en cuenta que los tres conjuntos de datos representan dinámicas de aludes similares asociadas a nieve húmeda. Para la subzona 3, la única que presenta una dinámica marcadamente diferente, la reducida población de estudio ha impedido obtener un modelo propio o variables distintas que expliquen su funcionamiento.

6 Conclusiones

El árbol implementado para el conjunto de la carretera obtiene un porcentaje de aciertos de 75,77 % para el conjunto de la muestra. Un resultado bastante elevado teniendo en cuenta que el modelo solo trabaja con datos de las estaciones automáticas. En comparación con otros trabajos que disponen de series de datos manuales más largas y que incluyen parámetros nivológicos, el modelo obtenido en este trabajo presenta resultados más limitados. Esto hace visible la necesidad de incluir en el modelo datos de estaciones manuales en el futuro. Para ello habrá que crear una serie con datos de nieve significativos y continuos.

En el modelo analizado, las distintas ramas de su estructura permiten seguir las distintas fases de un periodo de inestabilidad, desde su comienzo hasta su final. Cada una de estas ramas representa las diferentes situaciones de aludes que se producen en el puerto, las de nieve seca o las de nieve húmeda tras los episodios de precipitación.

El análisis por subzonas no ha conseguido su objetivo que era desvelar el funcionamiento pormenorizado de estas zonas. En el caso de las subzonas 1 y 2 la similitud de ambos modelos con el modelo general se justifica ya que sus dinámicas de avalanchas son idénticas.

En el caso de la subzona 3 la escasez del registro de análisis ha impedido obtener buenos resultados. La evidencia de que presenta una dinámica avalanchosa muy diferente a la del resto del puerto plantea la necesidad de realizar futuros análisis cuando se disponga de series temporales más largas.

La variable *Drift_XX* ha resultado ser la más significativa dentro del conjunto de distintas variables “drift” calculadas. Sin embargo, su función dentro de los modelos es aportar información sobre precipitación y, en cambio, no funcionan como parámetros de transporte de nieve. Las variables de transporte aportadas por el *flowcapt* no presentan ninguna relevancia estadística para el modelo general.

La exclusión de los parámetros de temperatura de todos los modelos ha sido una de las principales revelaciones del análisis. Sin embargo se justifica por la falta de nieve que a menudo presentan las zonas de aludes durante los episodios de subida de temperaturas.

Los datos de espesor de nieve acumulada se han revelado especialmente eficaces en el análisis estadístico. Los datos de espesor de nieve aportados por la estación automática de Bonaigua presentan grandes diferencias con la realidad de las zonas de salida de aludes del Puerto y ello dificulta y limita la correlación. Es evidente que disponer de una serie temporal completa de espesor en las zonas de salida supondría un incremento significativo de la eficacia del modelo.

7 Anexos

Anexos 1, 2 y 3 disponibles en el enlace:

https://drive.google.com/open?id=0B_TeX3256rwWVZNOEJXdGJrWEU

8 Bibliografía

- Baggi, S., and Schweizer, J., 2009, Characteristics of wet-snow avalanche activity: 20 years of observations from a high alpine valley (Dischma, Switzerland): *Natural Hazards*, v. 50, no. 1, p. 97–108, doi: 10.1007/s11069-008-9322-7.
- Bois, P., Obléd, C., and Good, W., 1975, Multivariate data analysis as a tool for day-to-day avalanche forecast, *in* *Proceedings of the International Symposium on Snow Mechanics*, I.A.H.S., Grindelwald, Switzerland, p. 391–403.
- Breiman, L. (Ed.), 1984, *Classification and regression trees*: Wadsworth International Group, Belmont, Calif.
- Buser, O., 1983, Avalanche forecast with the method of nearest neighbours: An interactive approach: *Cold Regions Science and Technology*, v. 8, no. 2, p. 155–163, doi: 10.1016/0165-232X(83)90006-X.
- Centre de Lauegi, Conselh Generau d’Aran, 2016, Situaciones de aludes: Centre de Lauegi d’Aran, URL: <http://lauegi.conselharan.org/les-situacions-dallaus/>. Accedido:11/06/16 19:54
- Davis, R.E., and Elder, K., 1994, Application of classification and regression trees: Selection of avalanche activity indices at Mammoth Mountain, *in* *Proceedings of International Snow Science Workshop*, Snowbird, Utah, p. 285–294.
- Davis, R.E., Elder, K., and Eddy Bouzaglou, 1992, Applications of classification tree methodology to the avalanche data management and forecasting, *in* *Proceedings of International Snow Science Workshop*, Breckenridge, Colorado, p. 126–133.
- Davis, R.E., Elder, K., Howlett, D., and Bouzaglou, E., 1999, Relating storm and weather factors to dry slab avalanche activity at Alta, Utah, and Mammoth Mountain, California, using classification and regression trees: *Cold Regions Science and Technology*, v. 30, no. 1–3, p. 79–89, doi: 10.1016/S0165-232X(99)00032-4.
- Díaz Sepulveda, J.F., 2012, Comparación entre árboles de regresión CART y regresión lineal [Master thesis]: Universidad Nacional de Colombia, 79 p.
- Floyer, J.A., and McClung, D.M., 2003, Numerical avalanche prediction: Bear Pass, British Columbia, Canada: *Cold Regions Science and Technology*, v. 37, no. 3, p. 333–342, doi: 10.1016/S0165-232X(03)00074-0.
- Gavaldà, P., 2002, Estudi del risc d’allaus en la zona de Comials (Port de la Bonaigua) Tesina Enginyeria Geològica: Universitat de Barcelona. Facultat de Geologia. UPC. Escola Tècnica Superior d’Enginyers de Camins Canals i Ports de Barcelona, 187 p.
- Gavaldà, J., and Benet, J., 2014, Ten years of avalanche forecasting on Bonaigua and Beret roads, Aran Valley, Spain, *in* *17th International Road Weather Conference*, Andorra.

- Gavaldà, J., and Moner, I., 2009, Handling of avalanche risk for the “Bonaigua” and “Beret” roads, Aran Valley, Spain, *in* Proceedings of International Snow Science Workshop, Davos, Switzerland, p. 577–582.
- Hendrikx, J., Murphy, M., and Onslow, T., 2012, The use of clasification trees as a tool for forecasting avalanche activity on the Seward highway, Alaska, *in* Proceedings of International Snow Science Workshop, Anchorage, Alaska, p.553-560.
- Hendrikx, J., Owens, I., Carran, W., and Carran, A., 2005, Avalanche activity in an extreme maritime climate: The application of classification trees for forecasting: Cold Regions Science and Technology, v. 43, no. 1–2, p. 104–116, doi: 10.1016/j.coldregions.2005.05.006.
- International Commission of Snow and Ice (Ed.), 1981, Avalanche atlas: illustrated international avalanche classification: Unesco, Paris.
- Jones, A.S., and Jamieson, B., 2001, Meteorological forecasting variables associated with skier-triggered dry slab avalanches: Cold Regions Science and Technology, v. 33, no. 2–3, p. 223–236, doi: 10.1016/S0165-232X(01)00040-4.
- McClung, D., and Shaerer, P., 1980, Snow avalanche size classification, *in* Proceedings of International Snow Science Workshop, Vancouver, BC, Canada, p. 12–30.
- Murphy, M., and Hendrikx, J., 2014, Operational use of a statistical forecasting tool on the Seward Highway, Alaska, *in* Proceedings of International Snow Science Workshop, Banff, Alberta, Canada, p. 470-474.
- Obled, C., and Good, W., 1980, Recent developments of avalanche forecasting by discriminant analysi techniques: a methodological eview and some aplications to Parsenn area (Davos, Switzerland): journal of Glaciology, v. 25, no. 92, p. 315–346.
- Peitzsch, E.H., Hendrikx, J., Fagre, D.B., and Reardon, B., 2012, Examining spring wet slab and glide avalanche occurrence along the Going-to-the-Sun Road corridor, Glacier National Park, Montana, USA: Cold Regions Science and Technology, v. 78, p. 73–81, doi: 10.1016/j.coldregions.2012.01.012.
- Purves, R., Morrison, K., Moss, G., and Wright, D.S., 2003, Nearest neighbours for avalanche forecasting in Scotland—development, verification and optimisation of a model: Cold Regions Science and Technology, v. 37, no. 3, p. 343–355, doi: 10.1016/S0165-232X(03)00075-2.
- Rosenthal, W., Kelly Elder, and Robert E. Davis, 2002, Operational decision tree avalanche forecasting, *in* Proceedings of International Snow Science Workshop, Penticton, British Columbia, p. 152–158.
- Serna Pineda, S.C., 2009, Comparación entre Árboles de regresión y clasificación y regresión logística [Master thesis]: Universidad Nacional de Colombia, 60 p.
- Trujillo, J., Antonio Sarria-Santamera, Aureli Esquerda, Mariona Badia, Matilde Palma, and Jaume March, 2008, Aproximación a la metodología basada en árboles de decisión (CART). Mortalidad hospitalaria del infarto agudo de miocardio: Gaceta Sanitaria, v. Gac Sanit vol.22 no.1.
- Xifre, D., 2007, Avaluació del risc d’allaus al Port de la Bonaigua [Tesina Enginyeria Geològica]: Universitat Politècnica de Catalunya.