https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

The Cryosphere Discussions

Combining snow physics and machine learning to predict avalanche activity: does it help?

Léo Viallon-Galinier[1,2,3], Pascal Hagenmuller[1], and Nicolas Eckert[2]

[1]Univ. Grenoble Alpes, Université de Toulouse, Météo-France, CNRS, CNRM, Centre d'Études de la Neige, Grenoble, France
[2]Univ. Grenoble Alpes, IRSTEA, UR ETNA, Grenoble, France
[3]École des Ponts, Champs-sur-Marne, France

Correspondence: Léo Viallon-Galinier (leo.viallon@meteo.fr)

Abstract. Predicting avalanche activity from meteorological and snow cover simulations is critical in mountainous areas to support operational forecasting. Several numerical and statistical methods have tried to address this issue. However, it remains unclear how the combination of snow physics, mechanical analysis of snow profiles and observed avalanche data improves avalanche activity prediction. This study combines extensive snow cover and snow stability simulations with observed avalanche occurrences within a Random Forest approach to predict avalanche days at a spatial resolution corresponding to elevations and aspects of avalanche paths in a given mountain range. We develop a rigorous leave-one-out evaluation procedure including an independent test set, confusion matrices, and receiver operating characteristic curves. In a region of the French Alps (Haute-Maurienne) and over the period 1960-2018, we show the added value within the statistical model of considering advanced snow cover modelling and mechanical stability indices instead of using only simple meteorological and bulk information. Specifically, using mechanically-based stability indices and their time-derivatives in addition to simple snow and meteorological variables increases the recall from around 65% to 76%. However, due to the scarcity of avalanche events and the possible misclassification of non-avalanche days in the training data set, the precision remains low, around 3.4%, due to the scarcity of avalanche days. These scores illustrate the difficulty of predicting avalanche occurrence with a high spatio-temporal resolution, even with the current cutting-edge data and modelling tools. Yet, our study opens perspectives to improve modelling tools supporting operational avalanche forecasting.

Keywords: snow avalanche, machine learning, mechanical stability indices, snow cover modelling, cross-validation, avalanche forecasting

1 Introduction

Avalanches are a significant issue in mountain areas (Wilhelm et al., 2000; Stethem et al., 2003) where they threaten recreationists and human infrastructure. The long-term (Keylock et al., 1999; Eckert et al., 2010b) and short-term (Schweizer et al., 2020) assessments of avalanche hazards and the related risks are therefore an important challenge for local authorities (Bründl and Margreth, 2021). Most of the countries facing such hazards rely on operational services for avalanche hazard forecasting (LaChapelle, 1977; Morin et al., 2019) and hazard mapping (Eckert et al., 2018). In this work, we focus on the issue of

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

short-term prediction (estimation of the outcomes of unseen data) of daily avalanche activity from simulated meteorological and snow data. Indeed, inferring the relation between avalanche activity and given weather and snow conditions is one of the essential components of operational avalanche hazard forecasting (prediction in the future on the basis of predicted snow and weather conditions).

Prediction of avalanche activity is mainly based on the knowledge of the snowpack evolution and of the mechanical processes leading to avalanches (e.g. LaChapelle, 1977; Morin et al., 2019). Information on the snowpack evolution can be collected through field observations and measurements (e.g. Coléou and Morin, 2018), and numerical simulations (e.g. Bartelt and Lehning, 2002; Vionnet et al., 2012). These data typically include a detailed description of the snowpack stratigraphy with vertical profiles of snow properties (Fierz et al., 2009). Several methods allow for identifying avalanche-prone situations from these profiles. Detection of weak layers based on mechanical and expert rules, such as the so-called lemons technique (Jamieson and Schweizer, 2005), comprises one qualitative approach. Numerical computation of stability indices based on mechanical theories constitutes an automated method to quantify the snowpack stability (Roch, 1966; Föhn, 1987; Lehning et al., 2004; Schweizer et al., 2006; Viallon-Galinier et al., 2021). These approaches rely on the knowledge of mechanical processes involved in avalanche release (Schweizer, 2017; Viallon-Galinier et al., 2021). Numerical models, which are currently used as an aid to decision-making for avalanche forecasters, generally combine mechanical stability indices and expert rules to provide information on snowpack stability (Morin et al., 2019; Schweizer et al., 2006; Giraud et al., 2002; Viallon-Galinier et al., 2021).

Machine learning techniques can approach the complex link between simple snow cover variables and avalanche occurrence. This method allows for taking advantage of the knowledge of past avalanche activity to determine objective delimitation lines around avalanche-prone conditions. The machine learning tools used so far mainly comprise classification trees (Kronholm et al., 2006; Hendrikx et al., 2014), random forests (e.g. Sielenou et al., 2021) or nearest neighbours (e.g. Navarre et al., 1987; Gassner and Brabec, 2002) techniques. Other techniques have also been tested, such as support vector machine (e.g. Pozdnoukhov et al., 2011; Choubin et al., 2019; Sielenou et al., 2021) and more advanced techniques appeared in the last years such as convolutional neural networks (e.g. Singh and Ganju, 2008; Dekanova et al., 2018). Most existing studies use meteorological variables as input or simple bulk variables such as snow depth to feed the statistical model, namely variables that are only surrogates for the true drivers of the avalanche processes. By contrast, studies using mechanically-based variables closely related to the processes involved in avalanche formation (e.g. Viallon-Galinier et al., 2021) are seldom used in machine learning approaches. However, these variables could increase the interpretability of the algorithm results and bring complementary non-linear information readily oriented towards the avalanche problem. Hence, they may reduce the complexity of statistical tools to implement (simpler statistical relations and a smaller number of variables to consider) and improve the overall predictive power.

Existing statistical prediction approaches are difficult to compare. Different spatial extensions are considered from large mountain ranges (e.g. Kronholm et al., 2006; Sielenou et al., 2021) to avalanche paths (e.g. Choubin et al., 2019). Different granularity of avalanche activity are also considered from binary classes (e.g. Kronholm et al., 2006; Hendrikx et al., 2014) to ordinal multi-classes (e.g. Mosavi et al., 2020; Sielenou et al., 2021). The most important difficulty is that existing studies

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

do not share a common evaluation process which includes a relevant segmentation of the train and test sets and common performance metrics. The absence of a homogeneous methodology for machine learning approaches evaluation among the snow and avalanche community limits the comparison between studies.

On this basis, this paper aims to study to which extent the combination of snow physics, mechanical analysis of snow profiles and observed avalanche data through machine learning techniques is of interest for predicting avalanche activity. We use random forest techniques to relate meteorological, snow conditions and mechanically-based stability indices to observed avalanche occurrences. We also employ time-derivative of mechanical indices to account for short time persistence of avalanche-prone conditions in certain cases. We eventually present a rigorous leave-one-out evaluation procedure of broad interest for evaluating avalanche prediction efficiency that includes an independent test set, confusion matrices, receiver operating characteristic (ROC) curves and additional scores derived from the confusion matrix. The study area is located in Haute-Maurienne in the French Alps, in which extensive avalanche data and snow cover reanalyses over 58 years (1960-2018) are available.

2 Material and methods

2.1 Study area

We selected in this study an area belonging to the Haute-Maurienne massif in the Northern French Alps, consisting of the three district municipalities of Bessans, Bonneval-sur-Arc and Lanslevillard (Figure 1). This area is frequently studied for avalanche related issues (e.g. Ancey et al., 2004; Eckert et al., 2009; Favier et al., 2014; Kern et al., 2021; Zgheib et al., 2020) because it is prone to intense avalanche activity and local series of avalanche events (see below) are particularly reliable, especially over the last decades. The area is characterized by a relatively high elevation ranging from 1400 to 3700 m, and its avalanche activity does not seem to be largely reduced by adverse climate warming effects (Lavigne et al., 2015; Zgheib et al., 2022). Located in the eastern French Alps next to the Italian border, the area experiences extreme snowfall events known as "easterly return", which drive most of the avalanche activity (Eckert et al., 2010a; Roux et al., 2021). We considered data on the winters between 1960 and 2018. When referring to the winter season, we consider days between the 15th of October and the 15th of March.

2.2 Avalanche observations

Our proxy of avalanche activity relies on the *Enquête Permanente sur les Avalanches* (EPA). The EPA reports all avalanches in approximately 3,000 pre-defined paths over french mountain ranges and about 110 in the studied area are recorded (Bourova et al., 2016). The paths of selected area are shown in Figure 1. Each avalanche record indicates the time period during which the avalanche is likely to have released and some additional information such as the elevation and the aspect of the starting zone. EPA was initially designed to capture large avalanche events in exposed areas and extensively used for hazard mapping (Bourova et al., 2016). Hence, only avalanches whose run-out reaches a certain threshold (e.g. reaching the valley floor, see Figure 1) are recorded. The avalanche activity derived from the EPA is dependent on this specific sampling procedure.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

The Cryosphere
Discussions

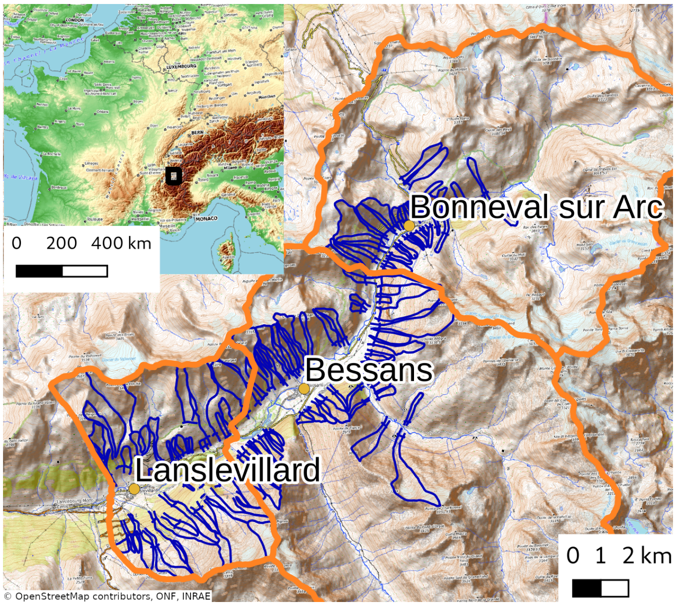

Figure 1. Studied area. General situation on the left and contour of avalanche paths surveyed each day (EPA) in blue for the three district municipalities of our studied area (delimited with orange lines). Only the avalanches that flow below a certain threshold (blue line at the bottom of each avalanche path) are reported.

Moreover, it relies on human-based observations and inevitably contains some uncertainties. However, EPA remains one of the longest avalanche activity records. The selected area is characterized by a dense observation network covering a large variety of avalanche paths. Besides, the steep topography of Haute-Maurienne reduces the effect of the threshold of observation as most of the avalanches reach the valley floor, providing a nearly exhaustive screenshot of natural avalanche activity. Further discussion on the EPA strengths and weaknesses is out of the scope of the paper and can be found in Jomelli et al. (2007) or Eckert et al. (2013).

One of the drawbacks of this data for the current study is the uncertainty of the date of some avalanche events, which can be large for remote sites or during low visibility periods. To associate meteorological and snow conditions to each observed avalanche, we removed from the dataset observations with an uncertainty of more than three days on the release date. When the uncertainty is larger than one day, the last day of the period was defined as the day of the avalanche event. Moreover, aspect and elevation of the starting zone were not reported in a few cases (representing less than 5% of the total number of events) because the starting point was not visible from the observation point or due to a lack of time. In these cases, the starting zone was defined by the average elevation and aspect of the typical release area defined for each avalanche path. We applied this definition of release day and zone to the 2779 avalanche events in the studied spatio-temporal domain.

We grouped these observations into eight aspect sectors (from North to North-West) and three elevation bands (centered at 1800, 2400 and 3000 m). This choice defines the spatial resolution of our model. All observations are represented in this geometry in Figure 2.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

2.3 Simulated snowpack

The SAFRAN-SURFEX/ISBA-Crocus model (Durand et al., 1999; Lafaysse et al., 2013) were used to simulate the snow and meteorological conditions on the Haute-Maurienne massif. SAFRAN provides meteorological information adapting numerical weather prediction on a gridded domain to the area of interest and assimilates observed meteorological data (Durand et al., 2009). We used the reanalysis publicly available (Vernay et al., 2020). Within this modelling scheme, meteorological conditions are assumed to depend only on elevation and aspect. The SURFEX/ISBA-Crocus model is a one-dimensional snowpack model representing snowpack evolution with a multi-layered scheme based on physical evolution laws (Brun et al., 1989; Vionnet et al., 2012). It uses as an input the meteorological data from SAFRAN model, and it is coupled to the soil scheme ISBA-DIF (Decharme et al., 2011) to represent energy and mass exchange at the bottom of the snowpack. Accordingly to the spatial resolution of the avalanche observations, snow conditions are computed for 8 aspects and 3 elevation levels (1800, 2400 and 3000 m). The temporal resolutions of the meteorological and snow conditions considered here were 1 h and 3 h, respectively.

These simulations retrieve meteorological and bulk snow conditions but also the full snowpack stratigraphy. Hence an additional step is required to take advantage of this information, which is here done through the computation of stability indices as presented right after.

2.4 Stability indices

Nine stability indices have been selected based on the applicability with our snow cover model: five for dry snow avalanches and four for wet snow avalanches. In addition, we also computed time-derivatives of these indices.

2.4.1 Dry snow indices

For dry snow, three indices are related to failure initiation, namely natural strength-stress ratio (S_n), skier strength-stress ratio (S_a) and external strength-stress ratio (S_r). These indices compare shear strength to shear stress where the stress originates from the weight of the overlying layers (S_n and S_a) and/or of an external load (skier, for instance) at the top of the snowpack (S_a and S_r) (Viallon-Galinier et al., 2021). Moreover, we selected two formulations of critical crack length for representing crack propagation (Viallon-Galinier et al., 2021): the original formulation by (Heierli et al., 2008; van Herwijnen et al., 2016) and the alternative approach by Gaume et al. (2017). Both approaches require a slab modulus, determined from density according to Scapozza (2004), and fracture energy estimated from strength Viallon-Galinier et al. (2021). Details on these indices are available in Viallon-Galinier et al. (2021).

These indices were computed for each layer. For each time step, five weak layers were then identified based on the values of each index. We defined a weak layer as a layer characterized by a local minimum of the considered stability index (excluding the top and the bottom layers). For each of the five weak layers, we computed the values of the five stability indices. This procedure leads to 25 variables (5 stability indices on 5 weak layers).

2.4.2 Wet snow indices

To characterize the conditions prone to wet snow avalanches, we used the mean liquid water content in the whole snowpack (Mitterer et al., 2013, 2016) and the thicknesses of humid snow layers. For the latter index, we considered a snow layer as humid as soon as its liquid water content exceeds either 0, 1 and 3% in volume. These three indices are denoted I_{h0}, I_{h1} and I_{h3}.

2.4.3 Time-derivatives

Stability indices values at a given time may not be sufficient to represent the avalanche activity. Time evolution of snow properties are supposed to be represented by snow cover models. However, Considering snowpack properties only at a given date and forgetting its past evolution does not indicate whether the snowpack is becoming more prone to avalanches or is in a stabilisation phase. For instance, low values of stability index may indicate an avalanche-prone situation but if these values are preceded by even lower ones, if an avalanche had to occur it will likely have occurred when the stability was minimal or even before but not after. However, few stability indices include the time dimension in the literature. To our knowledge, only Conway and Wilbour (1999) and MEPRA natural risk (Giraud et al., 2002; Viallon-Galinier et al., 2021) include explicit time dependence. Here, we decided to use time derivatives of the previously defined stability indices. We defined the time-derivative of stability index f on a given weak layer as $(f(t)-f(t-\mathrm{d}t))/\mathrm{d}t$ with several time intervals $\mathrm{d}t$ (6, 24 48, 72, 120 and 240 h). The derivatives represent 150 variables for dry snow indices and 24 for wet snow indices. Time derivatives on snow depth were also used as a straightforward indicator of stability for dry snow conditions (accumulation of new snow) and wet snow conditions (settling and melting).

2.5 Learning procedure

Random forests are used to relate snow and meteorological conditions to avalanche activity in the presented spatial resolution.

2.5.1 Avalanche activity

Avalanche activity is based on EPA records in the selected area. Days were classified into two categories: avalanche and non-avalanche days, depending on whether one avalanche, at least, has been observed in the area (aspect sector, elevation band). The model resolution defined by classes of elevation and aspect is more demanding than more common approaches applied to whole mountain ranges (e.g. Hendrikx et al., 2014; Kronholm et al., 2006; Sielenou et al., 2021) but provides more detailed information, closer to the spatial resolution used in avalanche operational forecasting (Morin et al., 2019). The number of avalanche days observed by elevation and aspect range is shown in Figure 2.

2.5.2 General overview of input variables

For each elevation and aspect selected, input variables used are summarized in Table 1.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

Table 1. Variables used to predict avalanche activity using machine learning

Category		Name	time intervals	Number of variables
Meteo	Snowfall	Snowfall accumulation (mm)	24 and 72h	2
	Rainfall	Rainfall accumulation (mm)	24 and 72h	2
	Temperature	Min, max, mean values (K)	24 and 72h	6
	Wind	Max and mean wind speed (km/h)	24 and 72h	4
		Projected mean direction on N-S axis and E-W axis	24 and 72h	4
Bulk	Snow depth	Snow depth (m)	—	1
	Depth of new snow (m)	Depth of snow fallen since (see intervals)	24, 72, 120h	3
Stability	Dry snow	Stability indices (S_n, S_a, S_k, a_c, a_g) for the 5 identified weak layers and depths of each weak layer	—	25
		Depth of the corresponding weak layers	—	5
	Wet snow	Maximum mean liquid water content	24h	1
		Maximum height of wet snow wit thresholds of 0, 1, 3% of liquid water to consider layer as wet (m)	24h	3
	Snow depth variation (m)	Snow depth variation	24, 72, 120h	3
Derivatives	Dry snow indices	All dry snow indices	6, 24, 48, 72, 120, 240h	150
	Wet snow indices	All wet snow indices	6, 24, 48, 72, 120, 240h	24

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

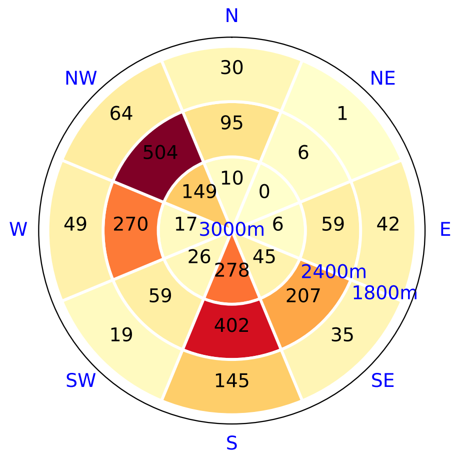

Figure 2. Number of avalanche days recorded on our study area over the full time period at the presented spatial resolution, i.e. per elevation band (1800, 2400 and 3000 m) and aspects (8 from N to NW).

These variables gather information from meteorological SAFRAN model (Meteo), SURFEX-ISBA/Crocus (Bulk), stability indices (Stability) computed on the base of modelled snowpack and derivatives of these variables (Derivatives) as described in Section 2.4. Hereafter, if no special mention is added, all these variables (All) are used but for studying variable importance, subsets of this list are also used.

2.5.3 Machine learning algorithm

To relate snow and meteorological conditions to avalanche activity as defined above, we used Random Forest (RF) techniques (Breiman, 2001; Hastie et al., 2009). Random Forest is an ensemble method used for classification. Each decision tree in the ensemble is built from a random subset of the data. This technique allows going beyond the limitations of single decision trees but without dramatically increasing the algorithm complexity and with similar introspection capabilities. Once trained, each tree of the Random Forest predicts a class for the input data. Aggregating all trees allow to define a probability for each class as the portion of trees predicting the given class.

Random Forest classifiers require two hyper-parameters: the number of trees and the tree depth. Here, we let the trees fully grow until there is only one element in each leaf as usually done (Hastie et al., 2009). An optimisation on our full dataset showed that 3000 trees were sufficient (more trees do not improve the results), so this value was selected for all this study.

We here have two classes, namely avalanche and non-avalanche days, that are highly unbalanced (mean of 1.1% of avalanche days in the winter season depending on elevation and aspect, see Figure 2). Machine learning techniques, if not handled with care, do not perform well on unbalanced data (e.g. Hastie et al., 2009; Sielenou et al., 2021). They are designed to optimize the overall classification accuracy or a similar score. Their results thus tend to be biased towards the majority class (Chawla et al., 2004; Sielenou et al., 2021), here the non-avalanche days. The most common techniques to limit this effect are oversampling

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

of the minority classes, undersampling of the majority classes or dedicated learning algorithms. We here used a combination of these techniques. We only considered days of the winter season and characterized by a simulated snow depth larger than 1 cm. This first selection led to the undersampling of the majority class. However, this first step was not sufficient to fully balance the dataset. We therefore used an adaptation of RF classifier to deal with unbalanced data (Chen et al., 2004): each tree of the forest is trained on a subset of the data randomly drawn; the probability law for drawing is adapted so that the probability of drawing non-avalanche or avalanche day are identical.

2.6 Evaluation methods

2.6.1 Evaluation process

We evaluated the model performance with a leave one year out approach (LOYO). The snowpack completely melted in summer, and new snowfall in autumn occurs on bare ground. Therefore, there is no memory between winter seasons and they are interchangeable. This is not the case between successive days during the winter season, whose snowpack characteristics are highly correlated. A simple leave one out (i.e. leave one day out) would yield better scores but would be less relevant. For each of the 58 seasons between 1960 and 2018, a test set is composed of one winter season and a learning set of the remaining 57 seasons. This leads to 58 sets of trained random forests, each one being evaluated on one year. On a single winter season, there are not enough avalanche events to be statistically relevant. Therefore, the confusion matrix of 58 test years were aggregated to compute scores with all information available. This leave one year out approach is used for all evaluations presented.

We also quantified the statistical uncertainty related to the sample size. As we used 58 years of evaluation data computed separately, we were able to define an uncertainty by bootstrapping evaluation years used to compute the considered score. In practice, 1000 independent draws of 58 years (with replacement) were randomly produced and the scores were computed on each draw. The 20th and 80th percentiles were used to quantify the uncertainty of the produced scores.

2.6.2 Scores

The Random Forest model produces the probability of being an avalanche day, defined as a day with at least one avalanche event, given the snow and meteorological conditions. We selected a threshold (t) on this probabilityy to discriminate avalanche days and non-avalanche days. It is possible to construct a confusion matrix (as presented in Table 2) based on this threshold. We derived three scores from the confusion matrix. The true positive rate or recall is the ratio between correctly predicted avalanche days divided by the number of observed avalanche days. It quantifies how many avalanche events have been correctly predicted. The false positive rate is the ratio between the number of false positive (non-avalanche days that are identified as avalanche days) and the total number of non-avalanche days. It corresponds to the probability that a false alarm will be raised. These two complementary indicators are interesting but do not fully characterize the performance of a binary classifier in case the two classes are unbalanced (which is the case here). We used a third score to represent how many predicted avalanche days are really observed as such. This score is called precision and is defined as the ratio between truly predicted avalanche days and the number of predicted avalanche days. The definition of these scores is summarized in Table 3.

Table 2. Confusion matrix: observed and predicted avalanche days ("Avalanche") and non-avalanche ones ("Non-avalanche").

		Predicted	
		Avalanche	Non-avalanche
Observed	Avalanche	AA	AN
	Non-avalanche	NA	NN

Table 3. Scores derived from the confusion matrix.

Name	expression
True positive rate or recall	$\frac{AA}{AA+AN}$
False positive rate	$\frac{NA}{NA+NN}$
Precision	$\frac{AA}{AA+NA}$

2.6.3 Scores presentation

These scores can be computed for any threshold t on the avalanche day probability. The impact of this threshold on the overall scores can be represented with two graphs: the ROC (Receiver Operating Characteristic) curve and the precision-recall graph. The ROC curve shows the true positive rate as a function of the false positive rate for all possible thresholds between 0 and 1. When the threshold is equal to 0, all days are considered avalanche days (true positive rate is 1, false positive rate is close to 1). When the threshold is 1, all days are considered non-avalanche days (true positive rate is 0 and false positive rate is close to zero). A perfect classifier would have a threshold value for which the true positive rate is 1, and the false positive rate is 0. A random classifier would have values of the true positive rate close to values of the false positive rate (first bisector). A common measure derived from this curve is the area between first bisector and ROC curve (the area under curve or AUC) (Bradley, 1997). The AUC quantifies how good the model is compared to a random classifier. We used the AUC value to compare different classifier configurations. In addition, recall is also plotted as a function of precision to capture the model capacity to identify avalanche days (precision) while limiting the number of false positives (recall). In this graph, the optimal point would be (1,1) i.e. a 100% precision and a 100% recall.

2.6.4 Importance of variables

The importance of variables was estimated through the separative power of each variable in the trees by computing the normalized mean decrease of impurity (also called Gini importance) on nodes where the given variable is used to separate the data in two groups (Breiman, 2001). A variable importance of zero means that the variable could be removed without reducing model performance and a high value denotes a high separative capacity (between avalanche and non-avalanche days) of the variable. If two variables contain similar information, each variable will be picked randomly in the tree construction and these variables

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

Table 4. Confusion matrix for the test dataset: observed and predicted avalanche days ("Avalanche") and non-avalanche ones ("Non-avalanche") summed over elevation and aspect ranges. A threshold value of 0.01 is used, i.e. predicted probabilities over 0.01 are considered to identify avalanche days. Corresponding recall is 75.3%, false positive rate is 23.6% and precision is 3.4%.

		Predicted	
		Avalanche	Non-avalanche
Observed	Avalanche	1 895	623
	Non-avalanche	55 005	178 357

will consequently share out the importance of the common information (Breiman, 2001). Hence, we also use for a more robust discrimination of the importance of variable through independent learning with different subsets of the variables (see subsets in Table 1).

3 Results

3.1 Model performance

The ROC curve of the model trained with all input variables at our spatial resolution and evaluated independently on each winter season since 1958 is shown in Figure 3a. Fortunately, the model is far better than a random classifier (ROC curve above the first diagonal) but it also remains far from an optimal classifier (no points closed to (0, 1)). The uncertainty around the ROC curve is very low, which indicates that a sufficient amount of data is available to constrain the model and that the evaluation is not sensitive to the choice of the winter season. The optimal threshold, defined as the threshold which leads to the ROC point closest to (0, 1), is here 0.01. In other words, a day is considered an avalanche day when the model probability is larger than 0.01. For this threshold, we provide, the corresponding confusion matrix in Table 4, classifying days between observed and predicted avalanche and non-avalanche days in all elevation and aspect bands. The corresponding scores are 75.3% for the true positive rate or recall, 23.6% for the false positive rate and 3.4% for precision. This means that more than three-quarter of the observed avalanche days are correctly identified but only 3.4% of the days predicted as avalanche one are really observed as such. An alternate point of view is to consider precision and recall rather than true and false positive rates (Figure 3b). The maximal precision that can be reached with our model is around 30% but with a very low value of recall (below 5%). With higher values of recall, the precision ranges between 2 and 10%.

3.2 Variable importance

As described in Section 2.6.4, the predictive power of the input variables can be estimated in two ways.

First, we computed the feature importance of all variables and aggregated (summed) them by groups (defined in Table 1) (Figure 4). The most important variables are related to snow depth (Figure 4) and, in particular, new snow amounts or variations for a time step of 72 and 120 h (with respective importance of 4 and 4.3% for new snow amount and 1.9 and 2.6% for snow

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

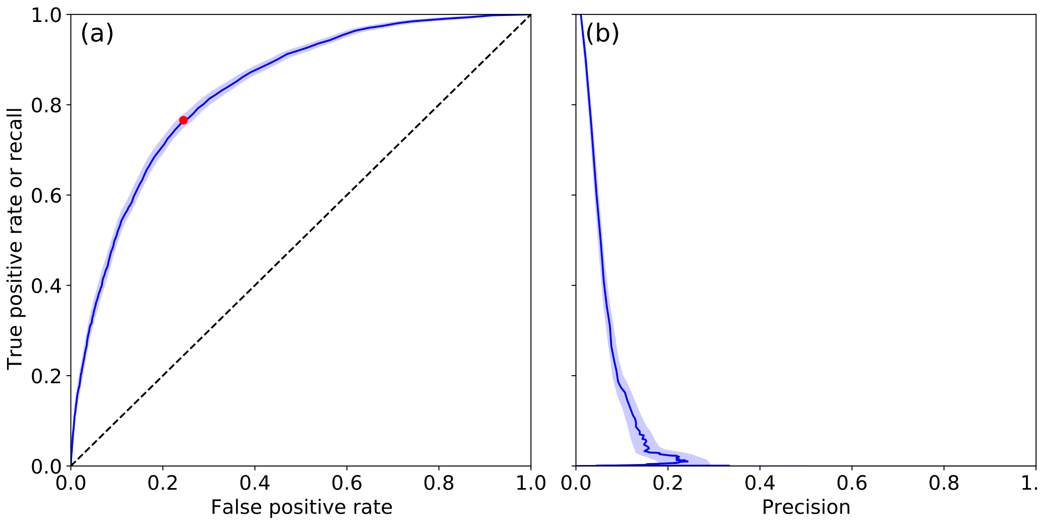

Figure 3. (a) ROC curve for the test dataset. The optimal point (threshold value of 0.01) is represented by a red dot. (b) Precision and recall (Table 3) curve. Shading represents the uncertainty by bootstrap on test years (see methods section).

depth variation). Variables related to dry snow stability appear to have similar importance but with much more variables in the corresponding group: 30 dry stability indices whereas only 7 variables for snow depth. Derivatives of dry snow indices decrease in importance with time step, whereas for wet snow indices, the importance is more pronounced for a time step of 72 h. Temperature and wind are also important, even described with a low number of involved variables. By contrast, snowfall and rainfall (on 24 h) are variables with low importance. The variability between years is limited, giving confidence in the robustness of these results.

Second, we studied the importance of variable groups by removing the data before learning and observing changes in evaluation scores. We selected 6 subsets of the presented variables (see Table 1): the meteorological variables only (Meteo), bulk variables only (Bulk), both together (Meteo+Bulk), stability variables without derivatives (Stability), stability variables and derivatives (Stability+Derivatives) and all variables (All). The ROC curves for all these subsets are presented in Figure 5. The associated scores for the optimal threshold are reported in Table 5. The ROC curve of the model trained only on meteorological variables is very close to the first bisector (Area Under the Curve AUC=0.09, Figure 5). In other words, this model is almost not much better than a random classifier. Using the bulk variables (snow depth and derivatives) allows for a first improvement in scores with an AUC of 0.19. As expected, adding meteorological variables to the bulk set (Meteo+Bulk) does not change the model scores significantly. Using only the 37 stability variables increases the AUC to 0.27 and combining with the associated 174 time derivatives increases the AUC to 0.32, which is close to the value (AUC=0.33) obtained by using all variables. However, the uncertainty linked to inter-annual variability is larger than the difference between the two latter approaches. This means that using all stability indices and their derivatives contains all relevant information available (in the context of other variables tested in this study) for discriminating avalanche and non-avalanche days. The other scores (false positive rate FPR,

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

The Cryosphere
Discussions
Open Access

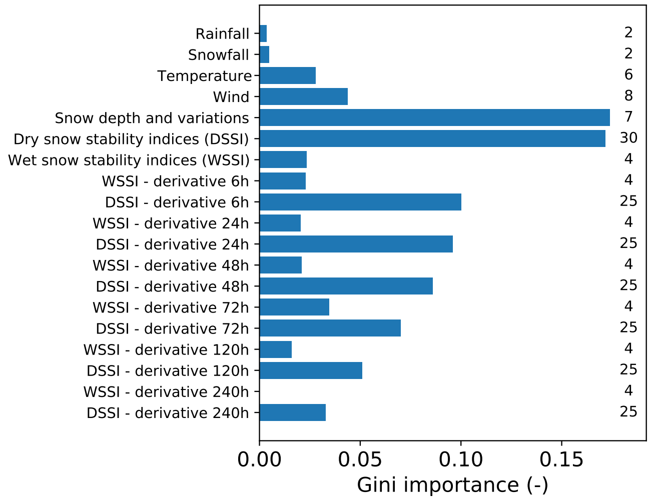

Figure 4. Feature importance (Gini importance) as computed from trained random forest with train dataset, aggregated (summed) by groups of variables. On the right, are reported the number of variables in each group.

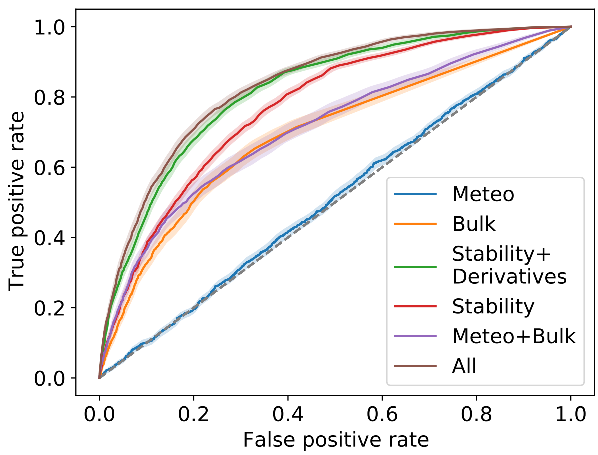

Figure 5. ROC curves of the model trained with different sets of variables. Shading represents the uncertainty by bootstrap on test years (see methods section). Labels of subsets of variables correspond to those of Table 1. Scores associated to the optimal points (nearest to (0, 1)) are reported in Table 5

recall, precision, see Table 3) present similar trends between groups compared to AUC except for the stability group that have higher FPR with lower recall compared to stability and derivatives or all variables highlighting that AUC or the selection of optimal threshold is a compromise between these two scores.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

Table 5. Predictive performance of the model trained with different sets of variables. The scores include area under ROC curve (AUC), false positive rate (FPR), recall and precision and the associated optimal threshold (associated to the point of the ROC curve nearest to the optimal one). Subsets of variables correspond to those of Table 1

Subset	AUC	FPR (%)	recall (%)	precision (%)	threshold
Meteo	0.009	49.9	51.7	1.1	0.025
Bulk	0.195	32.9	65.3	2.1	0.001
Meteo and Bulk	0.211	27.4	60.0	2.3	0.008
Stability	0.273	30.9	71.8	2.4	0.01
Stability and derivatives	0.319	26.6	76.3	3.0	0.01
All	0.334	23.6	75.3	3.4	0.01

4 Discussion

4.1 Machine learning for predicting avalanche activity

The model performance in the studied area decomposed into 8 aspects and 3 elevation bands, is summarized with the confusion matrix shown in Table 4. Values of recall (75.3%), false positive rate (23.6%) or precision (3.4%) may seem quite low compared to current literature. Hendrikx et al. (2005) or Kronholm et al. (2006) obtained accuracy for separation around 85% with regression trees and meteorological variables or simple bulk variables (snow depth or simple melting model). The accuracy of our model is 76.5% but this metric may not be the most informative when classes are highly unbalanced, as in our problem because it mainly gathers information on non-avalanche days. Sielenou et al. (2021) reported scores above 95% for accuracy but did not exploit other metrics. Hendrikx et al. (2014) reported a recall (focusing on observed avalanche days) of 76 to 79%, close to our value of 75.3%. Some studies, such as Pérez-Guillén et al. (e.g. 2021); Mayer et al. (e.g. 2021), did similar work using different targets (manually predicted avalanche hazard or measured stability) with accuracy also in comparable ranges (72 to 88%). Precision is highly influenced by the balance between avalanche and non-avalanche classes that are highly unbalanced, but representative of avalanche activity in the considered area. Moreover, low values of precision (around 3% for our model) are not uncommon for such difficult problems in related but different contexts (e.g. Rubin et al., 2012).

However, it remains difficult to compare scores to other studies due to differences in evaluation methods and reported scores. All studies used different methods for defining a train and a test set. In this study, we used a robust and conservative method, consisting in isolating winter seasons for evaluation. Indeed, with the snow melting between season, we get rid of the snowpack memory and provide a robust separation between train and test sets, leading to strong results with our evaluation method. Moreover, we discard all the days where the snow depth is less than 10 cm in the release zone, and days outside the winter period where avalanche release is very unlikely. Therefore, our evaluation is quite strict because it does not include obvious non-avalanche days. It is therefore more strict than Sielenou et al. (2021) for instance who used the random forest out of bag method with oversampling of the minority class. It resembles the methodology of Hendrikx et al. (2014) method who selected two independent years for evaluation. Our method may be used for future benchmarks to compare competing methods on a robust and homogeneous basis. In addition, the scores reported are not homogeneous between studies either. Some focus on global accuracy (e.g. Kronholm et al., 2006; Pérez-Guillén et al., 2021), other on accuracy per class (e.g. Sielenou et al., 2021) and few propose other metrics such as recall, precision or F1 score (harmonic mean of precision and recall) (e.g. Hendrikx et al., 2014). The choice of the score depends on the goal of each study and must be adapted to it. However, limiting to a few values for summarizing the model performances limits the information available. These differences in the evaluation processes, both separation between test and train sets and computed scores limits the possibility of model comparison.

Our model predicts the probability that at least one avalanche occurs on a given day and a spatial unit corresponding to one elevation band (centred at 1800, 2400 and 3000 m) and one aspect (among 8 aspects). This spatial resolution has been chosen to be closer to the information currently used for avalanche forecasting in many countries (Morin et al., 2019). This prediction goal is more demanding than prediction at larger scales as generally used in literature. It inevitably leads to lower performances for similar models but provides more precise information about the spatial distribution of the avalanche hazard (Statham et al., 2018). Indeed most studies considered avalanche activity at the scale of mountain ranges, of some thousands of km^2 (e.g. Kronholm et al., 2006; Hendrikx et al., 2014; Sielenou et al., 2021; Pérez-Guillén et al., 2021). These approaches have the advantage of using machine learning to also aggregate information at larger scales but provide a less geographically precise indicator of avalanche activity. More local approaches have the advantage of providing a relation between snow and meteorological conditions and observed or expected avalanche activity.

4.2 Added value of physical modelling of snow cover and stability for predicting avalanche activity

We tested different input variables to train our model: meteorological variables, bulk variables (mainly snow depth), stability indices and derivatives of bulk variables and stability indices. We evaluated the added value of the different groups of variables with two different methods (described in Section 2.6.4). Snow depth appeared as an essential variable in both methods: its Gini importance is high (Figure 4) and adding it to the input variables improves a lot the model performance (Figure 5). This result is consistent with current literature identifying snow depth as the first statistical predictor for avalanche activity (e.g. Schweizer et al., 2003; Castebrunet et al., 2012; Sielenou et al., 2021). Meteorological information is not sufficient by itself (Figure 5). Indeed, most of the developed models used, at least, some basic output of snow cover models or observed snow evolution such as snow depth (e.g. Hendrikx et al., 2014). Some studies used more advanced diagnostics from snow cover models (e.g. Gassner and Brabec, 2002; Pérez-Guillén et al., 2021; Mayer et al., 2021) or computed expert aggregated variables similarly to what snow cover models do from temperature and precipitations (e.g. Kronholm et al., 2006). Snow modelling with physical models for taking into account snowpack history thus appears of high interest for automatic avalanche activity prediction.

The novelty of our model is to add a wide range of stability indices to reduce the complex information of snow cover models with the help of knowledge of physical processes. The stability indices appear to have a significant importance (Figure 4). The introduction of stability indices could help identify avalanche-prone situations with statistical models. This group of variables also gather much information as it nearly replaces the information in other variables as the results are quite similar when using only stability indices and its derivatives and all variables (Figure 5). This result indicates that stability indices are a relevant way to summarize the information of meteorological and snow cover models for avalanche-prone conditions prediction, which is a new way of validating the interest of such stability indices.

Computing feature importance can drive the selection of relevant input variables but correlation between variables can affect the computed importance. Re-training the full model with a subset of input variables provides a robust estimation of their effective added value. In particular, the analysis of feature importance allows for selecting the right time steps for derivative computations in the wide range of possibilities included (last column of Table 1). The most important derivatives are the short-time ones (6 to 72 h) for dry snow and 72 h for wet snow (Figure 4). This result is consistent with the knowledge of involved processes (van Herwijnen et al., 2018) whereas it was never demonstrated from a statistical approach. The spontaneous release in dry snow occurs during or immediately after snowfall whereas wet snow problems are more linked to the progressive wetting of the snowpack with solar radiations (time scales of one to several days) (e.g. Reuter et al., 2022). Variable importance allows for selecting the most relevant variables which may be kept for further work, especially on stability variables and derivatives, which prove to be of interest (Figure 5).

4.3 Other pro's and con's of the approach

We here used the EPA as the ground truth of avalanche activity. This dataset is unique in its spatial and temporal extension but mainly focuses on large avalanches reaching valley floors or infrastructures. In consequence, the high-elevation avalanche activity and smaller avalanches are not reported which leads to a limited number of avalanche days in the dataset. For the spatio-temporal domain selected in this study (Haute-Maurienne, 1960-2018), the number of avalanche events reported in the EPA remains large enough (2518 avalanche situations) but in other regions, the scarcity of reported avalanche events might become a problem. Moreover, this study presents a binary classification as there is rarely more than one avalanche per day and spatial unit (aspect-elevation), which limits the definition of several classes of avalanche activity. In the future, these techniques may benefit from the use of other sources of data to complement EPA data and identify more avalanches, such as remote sensing (e.g. Karas et al., 2022), infrasound (Mayer et al., 2020) or seismic detection (van Herwijnen and Schweizer, 2011).

The interest of physics to predict avalanche activity was studied through a specific statistical tool which is random forest. This method is highly popular due to its very simple background (decision trees (Breiman et al., 1984)) which allows for in-depth analysis and interpretation to some extent and the capacity to represent non-linear phenomena (e.g. Sielenou et al., 2021; Pérez-Guillén et al., 2021; Mayer et al., 2021). Many other statistical methods are available but random forests have been shown to be as relevant as other ones (e.g. Sielenou et al., 2021). We introduced time-derivatives and cumulative values to represent the importance of history for snowpack-related processes. Methods in the range of recurrent neural networks are specifically designed to cope with processes having a memory of previous states (Hochreiter and Schmidhuber, 1997). This kind of statistical methods could be compared to our random forest approach to strengthen our results on the interest of combining snow physics and machine learning for predicting avalanche activity.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

Our results were obtained with a reanalysis of meteorological and snow conditions, that is to say, input data that have been retroactively corrected with all available observations. This may not be completely representative of operational forecasting (prediction in the future) situation in which models are corrected by observations of the past but run unconstrained for the forecast. This transposition to forecasting context would be the next step in terms of complexity for machine learning methods. However, the use of the reanalysis allows for a better evaluation of the capabilities of the machine learning model, with fewer input errors, which was the goal of this paper.

5 Conclusion and outlooks

This paper combines snow cover modelling, mechanical stability indices and observational data through machine learning for avalanche prediction. In particular, we considered numerous stability indices and their time-derivatives. To evaluate the random forest model, we define a robust method adapted to the specific behaviour of the snowpack (long-term memory). This evaluation was conducted on three district municipalities of the French Alps on 58 years of a comprehensive dataset of avalanche observations, with a high spatial resolution (8 aspects and 3 elevation ranges) and an extended set of variables describing both meteorological, mechanical stability variables and their time evolution.

The combination of snow physics through snowpack modelling, stability indices and their derivatives, with random forest proves to be valid. The snow depth remains the most important predictor but this study highlights the interest in using mechanical stability indices and their derivatives. This is the main finding of the paper as this was never demonstrated with such a large variety of indices and their derivatives as no stability inputs were used in most of the studies (Kronholm et al., 2006; Hendrikx et al., 2014; Sielenou et al., 2021, e.g.) or mainly simple stability indices (Mayer et al., 2021). It also more generally underlines the interest of physically-based snow cover models and stability indices for avalanche-prone conditions identification.

Obtained scores of recall (75.3%), false positive rate (23.6%) and recall (3.4%) are consistent with current literature with similar goals and methods. These scores illustrate this pregnant difficulty to predict avalanche occurrence with high spatio-temporal resolution, even with the cutting-edge data and modelling tools currently available. Moreover, we used a rather strict evaluation method leading to lower but more representative scores compared to other studies (e.g. Sielenou et al., 2021). This method may comprise a first step for future formal comparison between approaches. Yet, with its high spatio-temporal resolution and use of cutting-edge physical and mechanical models, our study opens perspective to improve modelling tools supporting operational avalanche forecasting.

We here focus on the avalanche activity reported by EPA. The method may be extended in the future to other target variables describing more precisely avalanche hazard such as release volume or typical situation (Schweizer et al., 2003; Statham et al., 2018; Reuter et al., 2021; Mayer et al., 2021). Similarly, we used meteorological reanalysis for snow modelling for the quality of the data but this may not be completely representative of forecast conditions and tests have to be conducted with re-forecasts rather than reanalysis.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

References

Ancey, C., Gervasoni, C., and Meunier, M.: Computing extreme avalanches, Cold Regions Science and Technology, 39, 161–180, https://doi.org/10.1016/j.coldregions.2004.04.004, 2004.

Bartelt, P. and Lehning, M.: A physical SNOWPACK model for the Swiss avalanche warning: Part I: numerical model, Cold Regions Science and Technology, 35, 123 – 145, https://doi.org/10.1016/S0165-232X(02)00074-5, 2002.

Bourova, E., Maldonado, E., Leroy, J., Alouani, R., Eckert, N., Bonnefoy-Demongeot, M., and Deschatres, M.: A new web-based system to improve the monitoring of snow avalanche hazard in France, Natural Hazards and Earth System Sciences, 16, 1205–1216, https://doi.org/10.5194/nhess-16-1205-2016, 2016.

Bradley, A. P.: The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognition, 30, 1145–1159, https://doi.org/10.1016/s0031-3203(96)00142-2, 1997.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/a:1010933404324, 2001.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J.: Classification and regression trees, Cole Advanced Books and Software, 1984.

Brun, E., Martin, E., Simon, V., Gendre, C., and Coleou, C.: An Energy and Mass Model of Snow Cover Suitable for Operational Avalanche Forecasting, Journal of Glaciology, 35, 333–342, https://doi.org/10.3189/S0022143000009254, 1989.

Bründl, M. and Margreth, S.: Integrative risk management: The example of snow avalanches, in: Snow and Ice-Related Hazards, Risks, and Disasters, pp. 259–296, Elsevier, https://doi.org/10.1016/b978-0-12-817129-5.00002-0, 2021.

Castebrunet, H., Eckert, N., and Giraud, G.: Snow and weather climatic control on snow avalanche occurrence fluctuations over 50 yr in the French Alps, Climate of the Past, 8, 855–875, https://doi.org/10.5194/cp-8-855-2012, 2012.

Chawla, N. V., Japkowicz, N., and Kotcz, A.: Editorial, ACM SIGKDD Explorations Newsletter, 6, 1–6, https://doi.org/10.1145/1007730.1007733, 2004.

Chen, C., Liaw, A., Breiman, L., et al.: Using random forest to learn imbalanced data, University of California, Berkeley, 110, 24, https://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf, 2004.

Choubin, B., Borji, M., Mosavi, A., Sajedi-Hosseini, F., Singh, V. P., and Shamshirband, S.: Snow avalanche hazard prediction using machine learning methods, Journal of Hydrology, 577, 123 929, https://doi.org/10.1016/j.jhydrol.2019.123929, 2019.

Coléou, C. and Morin, S.: Vingt-cinq ans de prévision du risque d'avalanche à Météo-France, La Météorologie, 100, 79–84, https://doi.org/10.4267/2042/65147, 2018.

Conway, H. and Wilbour, C.: Evolution of snow slope stability during storms, Cold Regions Science and Technology, 30, 67 – 77, https://doi.org/10.1016/S0165-232X(99)00009-9, 1999.

Decharme, B., Boone, A., Delire, C., and Noilhan, J.: Local evaluation of the Interaction between Soil Biosphere Atmosphere soil multilayer diffusion scheme using four pedotransfer functions, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD016002, 2011.

Dekanova, M., Duchon, F., Dekan, M., Kyzek, F., and Biskupic, M.: Avalanche forecasting using neural network, in: 2018 ELEKTRO, IEEE, https://doi.org/10.1109/elektro.2018.8398359, 2018.

Durand, Y., Giraud, G., Brun, E., Mérindol, L., and Martin, E.: A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting, Journal of Glaciology, 45, 469–484, https://doi.org/10.3189/S0022143000001337, 1999.

Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., and Mérindol, L.: Reanalysis of 44 Yr of Climate in the French Alps (1958–2002): Methodology, Model Validation, Climatology, and Trends for Air Temperature and Precipitation, Journal of Applied Meteorology and Climatology, 48, 429–449, https://doi.org/10.1175/2008JAMC1808.1, 2009.

Eckert, N., Parent, E., Faug, T., and Naaim, M.: Bayesian optimal design of an avalanche dam using a multivariate numerical avalanche model, Stochastic Environmental Research and Risk Assessment, 23, 1123–1141, https://doi.org/10.1007/s00477-008-0287-6, 2009.

Eckert, N., Coleou, C., Castebrunet, H., Deschatres, M., Giraud, G., and Gaume, J.: Cross-comparison of meteorological and avalanche data for characterising avalanche cycles: The example of December 2008 in the eastern part of the French Alps, Cold Regions Science and Technology, 64, 119–136, https://doi.org/10.1016/j.coldregions.2010.08.009, 2010a.

Eckert, N., Naaim, M., and Parent, E.: Long-term avalanche hazard assessment with a Bayesian depth-averaged propagation model, Journal of Glaciology, 56, 563–586, https://doi.org/10.3189/002214310793146331, 2010b.

Eckert, N., Keylock, C., Castebrunet, H., Lavigne, A., and Naaim, M.: Temporal trends in avalanche activity in the French Alps and subregions: from occurrences and runout altitudes to unsteady return periods., Journal of Glaciology, 59, 93 – 114, https://doi.org/10.3189/2013JoG12J091, 2013.

Eckert, N., Naaim, M., Giacona, F., Favier, P., Lavigne, A., Richard, D., Bourrier, F., and Parent, E.: Repenser les fondements du zonage règlementaire des risques en montagne « récurrents », La Houille Blanche, 104, 38–67, https://doi.org/10.1051/lhb/2018019, 2018.

Favier, P., Eckert, N., Bertrand, D., and Naaim, M.: Sensitivity of avalanche risk to vulnerability relations, Cold Regions Science and Technology, 108, 163–177, https://doi.org/10.1016/j.coldregions.2014.08.009, 2014.

Fierz, C., R.L, A., Y, D., P, E., Greene, E., Mcclung, D., Nishimura, K., Satyawali, P., and Sokratov, S.: The International Classification for Seasonal Snow on the Ground, IHP-VII Technical Documents in Hydrology, 83, https://unesdoc.unesco.org/ark:/48223/pf0000186462, 2009.

Föhn, P. M.: The stability index and various triggering mechanisms, IAHS publication, 162, 195–214, https://iahs.info/uploads/dms/iahs_162_0195.pdf, 1987.

Gassner, M. and Brabec, B.: Nearest neighbour models for local and regional avalanche forecasting, Natural Hazards and Earth System Science, 2, 247–253, https://doi.org/10.5194/nhess-2-247-2002, 2002.

Gaume, J., Van Herwijnen, A., CHAMBON, G., Wever, N., and Schweizer, J.: Snow fracture in relation to slab avalanche release: critical state for the onset of crack propagation, The Cryosphere, 11, 217–228, https://doi.org/10.5194/tc-11-217-2017, 2017.

Giraud, G., Navarre, J.-P., and Coléou, C.: Estimation du risque avalancheux dans le système expert MEPRA, Tech. rep., CNRM - Centre national de recherches météorologiques, https://hal.archives-ouvertes.fr/hal-02638748, unpublished work, 2002.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning, Springer New York, https://doi.org/10.1007/978-0-387-84858-7, 2009.

Heierli, J., Gumbsch, P., and Zaiser, M.: Anticrack Nucleation as Triggering Mechanism for Snow Slab Avalanches, Science, 321, 240–243, https://doi.org/10.1126/science.1153948, 2008.

Hendrikx, J., Owens, I., Carran, W., and Carran, A.: Avalanche activity in an extreme maritime climate: The application of classification trees for forecasting, Cold Regions Science and Technology, 43, 104–116, https://doi.org/10.1016/j.coldregions.2005.05.006, 2005.

Hendrikx, J., Murphy, M., and Onslow, T.: Classification trees as a tool for operational avalanche forecasting on the Seward Highway, Alaska, Cold Regions Science and Technology, 97, 113 – 120, https://doi.org/10.1016/j.coldregions.2013.08.009, 2014.

Hochreiter, S. and Schmidhuber, J.: Long short-term memory, Neural computation, 9, 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735, 1997.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

The Cryosphere Discussions

Jamieson, B. and Schweizer, J.: Using a checklist to assess manual snow profiles, Avalanche News, 72, 57–61, 2005.

Jomelli, V., Delval, C., Grancher, D., Escande, S., Brunstein, D., Hetu, B., Filion, L., and Pech, P.: Probabilistic analysis of recent snow avalanche activity and weather in the French Alps, Cold Regions Science and Technology, 47, 180–192, https://doi.org/10.1016/j.coldregions.2006.08.003, 2007.

Karas, A., Karbou, F., Giffard-Roisin, S., Durand, P., and Eckert, N.: Automatic Color Detection-Based Method Applied to Sentinel-1 SAR Images for Snow Avalanche Debris Monitoring, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–17, https://doi.org/10.1109/tgrs.2021.3131853, 2022.

Kern, H., Eckert, N., Jomelli, V., Grancher, D., Deschatres, M., and Arnaud-Fassetta, G.: Brief communication: Weak control of snow avalanche deposit volumes by avalanche path morphology, The Cryosphere, 15, 4845–4852, https://doi.org/10.5194/tc-15-4845-2021, 2021.

Keylock, C. J., McClung, D. M., and Magnússon, M. M.: Avalanche risk mapping by simulation, Journal of Glaciology, 45, 303–314, https://doi.org/10.3189/S0022143000001805, 1999.

Kronholm, K., Vikhamar-Schuler, D., Jaedicke, C., Isaksen, K., Sorteberg, A., and Kristensen, K.: Forecasting snow avalanche days from meteorological data using classification trees; Grasdalen, Western Norway, in: Proceedings of the International Snow Science Workshop, Telluride, Colorado, pp. 1–6, Citeseer, https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.855.5195&rep=rep1&type=pdf, 2006.

LaChapelle, E. R.: Snow Avalanches: A review of Current Research and Applications, Journal of Glaciology, 19, 313–324, https://doi.org/10.3189/s0022143000215633, 1977.

Lafaysse, M., Morin, S., Coléou, C., Vernay, M., Serça, D., Besson, F., Willemet, J.-M., Giraud, G., Durand, Y., and Météo-France, D.: Towards a new chain of models for avalanche hazard forecasting in French mountain ranges, including low altitude mountains, in: Proceedings of International Snow Science Workshop Grenoble–Chamonix Mont-Blanc, vol. 7, pp. 162–166, http://arc.lib.montana.edu/snow-science/objects/ISSW13_paper_O1-02.pdf, 2013.

Lavigne, A., Eckert, N., Bel, L., and Parent, E.: Adding expert contributions to the spatiotemporal modelling of avalanche activity under different climatic influences, Journal of the Royal Statistical Society: Series C (Applied Statistics), 64, 651–671, https://doi.org/10.1111/rssc.12095, 2015.

Lehning, M., Fierz, C., Brown, B., and Jamieson, B.: Modeling snow instability with the snow-cover model SNOWPACK, Annals of Glaciology, 38, 331–338, https://doi.org/10.3189/172756404781815220, 2004.

Mayer, S., van Herwijnen, A., Ulivieri, G., and Schweizer, J.: Evaluating the performance of an operational infrasound avalanche detection system at three locations in the Swiss Alps during two winter seasons, Cold Regions Science and Technology, 173, 102 962, https://doi.org/10.1016/j.coldregions.2019.102962, 2020.

Mayer, S., van Herwijnen, A., and Schweizer, J.: A random forest model to assess snow instability from simulated snow stratigraphy, in: EGU General Assembly Conference Abstracts, pp. EGU21–12 259, https://ui.adsabs.harvard.edu/abs/2021EGUGA..2312259M/abstract, 2021.

Mitterer, C., Techel, F., Fierz, C., and Schweizer, J.: An operational supporting tool for assessing wet-snow avalanche danger, in: Proceedings ISSW, vol. 33, https://www.researchgate.net/profile/Frank_Techel/publication/261672459_An_operational_supporting_tool_for_assessing_wet-snow_avalanche_danger/links/00b49534f63d14b4ae000000/An-operational-supporting-tool-for-assessing-wet-snow-avalanche-danger.pdf, 2013.

Mitterer, C., Heilig, A., Schmid, L., van Herwijnen, A., Eisen, O., and Schweizer, J.: Comparison of measured and modelled snow cover liquid water content to improve local wet-snow avalanche prediction, in: International Snow Science Workshop Proceedings, https://arc.lib.montana.edu/snow-science/objects/ISSW16_O6.01.pdf, 2016.

Morin, S., Horton, S., Techel, F., Bavay, M., Coléou, C., Fierz, C., Gobiet, A., Hagenmuller, P., Lafaysse, M., Ližar, M., Mitterer, C., Monti, F., Müller, K., Olefs, M., Snook, J. S., van Herwijnen, A., and Vionnet, V.: Application of physical snowpack models in support of operational avalanche hazard forecasting: A status report on current implementations and prospects for the future, Cold Regions Science and Technology, p. 102910, https://doi.org/https://doi.org/10.1016/j.coldregions.2019.102910, 2019.

Mosavi, A., Shirzadi, A., Choubin, B., Taromideh, F., Hosseini, F. S., Borji, M., Shahabi, H., Salvati, A., and Dineva, A. A.: Towards an Ensemble Machine Learning Model of Random Subspace Based Functional Tree Classifier for Snow Avalanche Susceptibility Mapping, IEEE Access, 8, 145 968–145 983, https://doi.org/10.1109/access.2020.3014816, 2020.

Navarre, J., Guyomar'ch, G., and Giraud, G.: Un modèle statistique pour la prévision locale des avalanches, IAHS-Publ, 162, 571–580, Unmod{è}lestatistiquepourlapr{é}visionlocaledesavalanches, 1987.

Pérez-Guillén, C., Hendrick, M., Techel, F., van Herwijnen, A., Volpi, M., Tasko, O., Pérez-Cruz, F., Obozinski, G., and Schweizer, J.: Data-driven automatic predictions of avalanche danger in Switzerland, in: EGU General Assembly, Copernicus GmbH, https://doi.org/10.5194/egusphere-egu21-6154, 2021.

Pozdnoukhov, A., Matasci, G., Kanevski, M., and Purves, R. S.: Spatio-temporal avalanche forecasting with Support Vector Machines, Natural Hazards and Earth System Sciences, 11, 367–382, https://doi.org/10.5194/nhess-11-367-2011, 2011.

Reuter, B., Viallon-Galinier, L., Mayer, S., Hagenmuller, P., and Morin, S.: A tracking algorithm to identify slab and weak layer combinations for assessing snow instability and avalanche problem type, in: EGU General Assembly Conference Abstracts, Copernicus GmbH, https://doi.org/10.5194/egusphere-egu21-1578, 2021.

Reuter, B., Viallon-Galinier, L., Horton, S., van Herwijnen, A., Mayer, S., Hagenmuller, P., and Morin, S.: Characterizing snow instability with avalanche problem types derived from snow cover simulations, Cold Regions Science and Technology, 194, 103 462, https://doi.org/10.1016/j.coldregions.2021.103462, 2022.

Roch, A.: Les déclenchements d'avalanche, IAHS Publ., 69, 86–99, https://iahs.info/uploads/dms/069021.pdf, 1966.

Roux, E. L., Evin, G., Eckert, N., Blanchet, J., and Morin, S.: Elevation-dependent trends in extreme snowfall in the French Alps from 1959 to 2019, The Cryosphere, 15, 4335–4356, https://doi.org/10.5194/tc-15-4335-2021, 2021.

Rubin, M. J., Camp, T., Herwijnen, A. V., and Schweizer, J.: Automatically Detecting Avalanche Events in Passive Seismic Data, in: 2012 11th International Conference on Machine Learning and Applications, IEEE, https://doi.org/10.1109/icmla.2012.12, 2012.

Scapozza, C.: Entwicklung eines dichte-und temperaturabhängigen Stoffgesetzes zur Beschreibung des visko-elastischen Verhaltens von Schnee, Ph.D. thesis, ETH Zurich, https://doi.org/10.3929/ethz-a-004680249, 2004.

Schweizer, J.: On recent advances in avalanche research, Cold Regions Science and Technology, 144, 1 – 5, https://doi.org/10.1016/j.coldregions.2017.10.014, international Snow Science Workshop 2016 Breckenridge, 2017.

Schweizer, J., Bruce Jamieson, J., and Schneebeli, M.: Snow avalanche formation, Reviews of Geophysics, 41, https://doi.org/10.1029/2002RG000123, 2003.

Schweizer, J., Bellaire, S., Fierz, C., Lehning, M., and Pielmeier, C.: Evaluating and improving the stability predictions of the snow cover model SNOWPACK, Cold Regions Science and Technology, 46, 52 – 59, https://doi.org/10.1016/j.coldregions.2006.05.007, 2006.

Schweizer, J., Mitterer, C., Techel, F., Stoffel, A., and Reuter, B.: On the relation between avalanche occurrence and avalanche danger level, The Cryosphere, 14, 737–750, https://doi.org/10.5194/tc-14-737-2020, 2020.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

Sielenou, P. D., Viallon-Galinier, L., Hagenmuller, P., Naveau, P., Morin, S., Dumont, M., Verfaillie, D., and Eckert, N.: Combining random forests and class-balancing to discriminate between three classes of avalanche activity in the French Alps, Cold Regions Science and Technology, p. 103276, https://doi.org/10.1016/j.coldregions.2021.103276, 2021.

Singh, A. and Ganju, A.: Artificial Neural Networks for snow avalanche forecasting in Indian Himalaya, in: Proceedings of 12th International Conference of International Association for Computer Methods and Advances in Geomechanics, IACMAG, vol. 16, 2008.

Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., and Kelly, J.: A conceptual model of avalanche hazard, Natural hazards, 90, 663–691, https://doi.org/10.1007/s11069-017-3070-5, 2018.

Stethem, C., Jamieson, B., Schaerer, P., Liverman, D., Germain, D., and Walker, S.: Snow Avalanche Hazard in Canada – a Review, Natural Hazards, 28, 487–515, https://doi.org/10.1023/a:1022998512227, 2003.

van Herwijnen, A. and Schweizer, J.: Monitoring avalanche activity using a seismic sensor, Cold Regions Science and Technology, 69, 165–176, https://doi.org/10.1016/j.coldregions.2011.06.008, 2011.

van Herwijnen, A., Gaume, J., Bair, E. H., Reuter, B., Birkeland, K. W., and Schweizer, J.: Estimating the effective elastic modulus and specific fracture energy of snowpack layers from field experiments, Journal of Glaciology, 62, 997–1007, https://doi.org/10.1017/jog.2016.90, 2016.

van Herwijnen, A., Heck, M., Richter, B., Sovilla, B., and Techel, F.: When Do Avalanches Release: Investigating Time Scales in Avalanche Formation, in: Proceedings, International Snow Science Workshop, 2018.

Vernay, M., Lafaysse, M., Hagenmuller, P., Nheili, R., Verfaillie, D., and Morin, S.: The S2M meteorological and snow cover reanalysis in the French mountainous areas (1958 - present), version 2020.2, https://doi.org/10.25326/37, 2020.

Viallon-Galinier, L., Hagenmuller, P., and Lafaysse, M.: Forcing and evaluating detailed snow cover models with stratigraphy observations, Cold Regions Science and Technology, 180, 103 163, https://doi.org/10.1016/j.coldregions.2020.103163, 2021.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., and Willemet, J.-M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geoscientific Model Development, 5, 773–791, https://doi.org/10.5194/gmd-5-773-2012, 2012.

Wilhelm, C., Wiesinger, T., and Ammann, W.: The avalanche winter 1999 in Switzerland-an overview, Tech. rep., WSL/SLF Davos, 2000.

Zgheib, T., Giacona, F., Granet-Abisset, A.-M., Morin, S., and Eckert, N.: One and a half century of avalanche risk to settlements in the upper Maurienne valley inferred from land cover and socio-environmental changes., Global Environmental Change, 65, 102 149, https://doi.org/10.1016/j.gloenvcha.2020.102149, 2020.

Zgheib, T., Giacona, F., Granet-Abisset, A.-M., Morin, S., Lavigne, A., and Eckert, N.: Spatio-temporal variability of avalanche risk in the French Alps, Regional Environmental Change, 22, https://doi.org/10.1007/s10113-021-01838-3, 2022.

Code and data availability. The meteorological and snow cover reanalysis used in this study is freely available at http://dx.doi.org/10.25326/37#v2020.2. The whole EPA avalanche database is freely available at https://www.avalanches.fr/.

Author contributions. Léo Viallon-Galinier developed the model code and performed the simulations. Léo Viallon-Galinier prepared the manuscript with contributions from all co-authors.

https://doi.org/10.5194/tc-2022-108
Preprint. Discussion started: 3 June 2022

Competing interests. The authors declare no competing interests.

Acknowledgements. INRAE and CNRM/CEN are members of LabEx OSUG.