Combining random forests and class-balancing to discriminate between three classes of avalanche activity in the French Alps

Pascal Dkengne Sielenou, Léo Viallon-Galinier, Pascal Hagenmuller, Philippe Naveau, Samuel Morin, Marie Dumont, Deborah Verfaillie, Nicolas Eckert

HAL Id: hal-03433337
https://hal.inrae.fr/hal-03433337v1

Submitted on 17 Nov 2021

Combining random forests and class-balancing to discriminate between three classes of avalanche activity in the French Alps

Pascal Dkengne Sielenou [a], Léo Viallon-Galinier [b,a,c], Pascal Hagenmuller [b], Philippe Naveau [d], Samuel Morin [b], Marie Dumont [b], Deborah Verfaillie [b,e], Nicolas Eckert [a,*]

[a] UR ETNA, INRAE, Université de Grenoble Alpes, France
[b] Univ. Grenoble Alpes, Université de Toulouse, Météo-France, CNRS, CNRM, Centre d'Études de la Neige, Grenoble, France
[c] École des Ponts, Champs-sur-Marne, France
[d] Laboratoire des Sciences du Climat et de l'Environnement (CNRS-CEA-UVSQ), EstimR, 91191 Gif-sur-Yvette, France
[e] Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

ARTICLE INFO

ABSTRACT

Determining avalanche activity corresponding to given snow and meteorological conditions is an old problem of high practical relevance. To address it, numerous statistical forecasting models have been developed, but intercomparisons of their efficiency on very large datasets are seldom. In this work, an approach combining random forests with class-balancing is presented and systematically compared with competing methods currently described in the avalanche literature. On more than 50 years of daily avalanche observations, in the 23 massifs of the French Alps, the competing classifiers are evaluated on their ability to distinguish three classes of avalanche activity: non-avalanche days, days with moderate activity, and days with high activity. Moreover, the variables of higher importance in the random forest classifiers are shown to be coherent with current avalanche literature and a clustering based on these variable importance separates massifs which are known to have different avalanche activities. Our approach opens perspectives to support operational avalanche forecasting.

1. Introduction

Snow avalanches are natural phenomena that regularly occur in mountainous regions, putting people, properties and infrastructures at risk (e.g. McClung and Schaerer, 1993; Pudasaini, 2007). Determining where and when avalanche activity is likely to occur is therefore one of the most crucial topics in applied snow and avalanche science. It has attracted scientists' attention for a long time (LaChapelle, 1980; McClung, 2002), but several questions remain unclear, such as the occurrence conditions of large and infrequent avalanches (Schweizer et al., 2009), or of deep slab avalanches resulting from persistent weak layers (Jamieson et al., 2001; Marienthal et al., 2015).

Avalanche activity is controlled by permanent and variable factors (de Quervain, 1981). Permanent factors are related to terrain features (elevation, slope, aspect, roughness of the ground, etc.) and variable factors are related to meteorological conditions (snowfall, rain, wind, temperature, etc.) that progressively build the snowpack. Because of complex non-linear relationships between avalanche release and these factors, it is generally admitted that an exact deterministic prediction of

avalanche release remains out of reach (Schweizer et al., 2003a, 2008). In several countries, mitigation measures such as preventive road closures or the production of *avalanche bulletins* rely on stability indices (Judson and King, 1985; Morin et al., 2019) that provide an avalanche danger level as a function of snow and meteorological data at a given spatial scale (mountain range, valley, etc.), often distinguishing natural from artificial (human) triggers, and weighting differently slopes and aspects (e.g., Schweizer et al. (2020)).

Existing stability indices are more or less physically based, and are defined at different spatial levels. To make it simple, local indices are based on physical stability criteria applied to the snowpack only, leading to so-called *physically based avalanche forecasting approaches*. Regional indices are based on empirical relationships between avalanche observations and snow and meteorological conditions obtained using learning methods, leading to so-called *statistical avalanche forecasting approaches*. However, the boundary between these two classes of approaches is far from being that strict. The transition from local snow stratigraphy to release susceptibility at a larger spatial scale is a tough problem (Jamieson et al., 2009; Bakermans et al., 2010). As a consequence, local

---

physical stability criteria need to be combined with some experience-based knowledge to be upscaled to the operational scale. Besides, even at the very local scale, the physical processes are sufficiently complex to make the use of data-based relationships unavoidable within the modelling of snowpack evolution and avalanche release. Conversely, statistical learning methods basically aim at capturing physics in simplified form, not necessarily universally valid, but, at least, relevant for the location and conditions under study.

Regarding the deterministic or probabilistic nature of the approach, statistical forecasting models are eventually deterministic predictors of a given stability level, interpretable in terms of a given avalanche activity to be expected in the targeted area according to the snow and meteorological conditions. Probabilistic uncertainty considerations are introduced only as soon as one considers uncertainty in snow and meteorological conditions (ensemble of conditions, Vernay et al., 2015) and/or in the model structure (Lafaysse et al., 2017). On the other hand, the probabilistic versions of physically-based local indices that have been developed (Gaume et al., 2014) explicitly model that the response of the snowpack varies strongly in space according to snow spatial variability (Reuter et al., 2016) and to topography. Integrating out this variability leads to a regional release probability. The latter is interpretable, as an index resulting from a statistical avalanche forecasting model, as a given avalanche activity to be expected in the targeted area. Hence, here again, boundaries between physics and empiricism are not as strong as they initially seem to be.

A wide range of statistical avalanche forecasting models relating snow and meteorological data to avalanche activity has been developed over the last decades, following the progress of statistical learning techniques and of the computational power required to implement these on large datasets. Also, simulated snow cover data have progressively supplemented or replaced pointwise snow and meteorological observations in statistical models (Schirmer et al., 2009). Following simple threshold-based approaches (Perla, 1970), linear methods involving normality assumptions have historically been the first proposed and implemented. These include multivariate regressions, as well as discriminant and canonical analyses performed in the (high-dimensional) space of potential avalanche drivers (Bois et al., 1974; Bovis, 1977; Föhn et al., 1977; McGregor, 1989; McClung and Tweedy, 1994; Floyer and McClung, 2003). As a popular non-parametric alternative, nearest neighbour approaches predict avalanche activity by identifying what happened in similar snow and meteorological conditions considering the past. After pioneering developments (Obled and Good, 1980; Buser, 1983, 1989), operational nearest neighbour models were implemented in many countries including France (Navarre et al., 1987), Switzerland (Gassner and Brabec, 2002) Scotland (Purves et al., 2003) and Iran (Choubin et al., 2019). Further refinements such as inclusion of expert knowledge (Schweizer and Föhn, 1996) or explicit consideration of the topography into the analysis (McCollister et al., 2003) have been proposed, without modifying the spirit of the methodology. More recently, non-linear classification methods were introduced, notably classification trees (Davis et al., 1999; Hendrikx et al., 2005; Baggi and Schweizer, 2009; Hendrikx et al., 2014), random forests – also called forest decision trees – (Mitterer and Schweizer, 2013; Möhle et al., 2014; Marienthal et al., 2015; Dreier et al., 2016), and support vector machines (Pozdnoukhov et al., 2008, 2011).

From the review of the important literature devoted to avalanche statistical forecasting models, different critical issues can be identified. First, it is difficult to define avalanche activity classes at a spatial scale where comparison between avalanche observations, avalanche activity predictions and available snow and meteorological data is feasible and meaningful. Second, the evaluation of any developed approach is a crucial question (Schweizer et al., 2003b; Heierli et al., 2004; Schirmer et al., 2010). Yet, this step is rarely investigated in details because i) of the lack of long and comprehensive records of avalanche activity together with snow and meteorological conditions, and ii) of the local-to regional-scale transition problem previously mentioned. Third, even

if the inferred relationships between avalanche activity and covariates remain essentially empirical, they are often considered as sufficiently robust to be physically interpretable. Specifically, these relationships are used to identify the main snow and meteorological drivers of avalanche activity, and, for trees, the succession of processes leading to release (Conway and Raymond, 1993; Conway and Wilbour, 1999).

In this study, we aim at evaluating competing classification techniques and the necessary pre-processing of the data to have an idea of the potential of such techniques for future use as new indicators of avalanche-prone situations. To this end, we take advantage of the comprehensive and high-quality data available all over the French Alps: more than 50 years of daily reanalysed snow and meteorological conditions and avalanche activity observations in 23 massifs (Durand et al., 2009a, 2009b). We develop a statistical approach to predict avalanche activity and evaluate it, in a leave-one-out cross validation scheme, against the main learning methods currently described in the avalanche statistical forecasting field. Specifically, we consider avalanche prediction as a classification problem, where we wish to classify, at the massif scale, days into three classes defined as follows: non-avalanche days (null class), days with a moderate number of recorded avalanches (moderate class) and days with a large number of recorded avalanches (high class). This classification is made given current and past snow and meteorological conditions. Note that this definition of classes results into unbalanced classes because, in any winter season, the number of days with avalanche occurrences, and, even more, of days with high avalanche activity, is indeed (very) small compared to the number of non-avalanche days. This makes most classification approaches less efficient (if not inefficient), another difficulty generally overlooked in applied snow science. To this aim, we introduce in Section 3 a scheme to balance avalanche activity classes. Section 4 contains the main results from the application of our approach to the 23 massifs constituting the French Alps over the 1958–2010 period. Section 5 discusses strengths, weaknesses and outcomes for operational avalanche forecasting. Section 6 concludes.

2. Processed data

The covariates used in this study consist of reanalysed snow and meteorological data while the response variables are the observed avalanche occurrences in the 23 massifs used for operational avalanche forecasting in the French Alps (see Fig. 1). The area of each massif ranges from 500 to 1500 km^2 and the key assumption regarding snow and meteorological numerical simulations is their spatial homogeneity, i.e., within each massif, meteorological and snowpack properties are assumed to depend only on elevation, slope and aspect (see Durand et al., 1999).

2.1. Snow and meteorological conditions

Reanalysed snow and meteorological conditions are outputs from the SAFRAN-Crocus model chain (Durand et al., 2009a, 2009b; Vionnet et al., 2012). The meteorological reanalysis was performed at the massif scale using 44 years of analyzed large-scale atmospheric data from the 40-year European Centre for Medium-Range Weather Forecast (ECMWF) reanalysis (ERA-40) project (Uppala et al., 2004), together with meteorological observations including numerous mountain stations. For several years beyond the end date of the ERA-40 dataset, large-scale meteorological fields from Météo-France operational numerical weather prediction models were used. All in all, the reanalysis covers the period from 1958 to 2009. Vernay et al. (2019) provide an updated version of the SAFRAN-Crocus reanalysis, sharing many characteristics with the dataset used in this study.

The specific variables used, similar to those described by Castebrunet et al. (2012, 2014) concern three elevations (low, mid and high - 1800, 2400, and 3000 m a.s.l.):
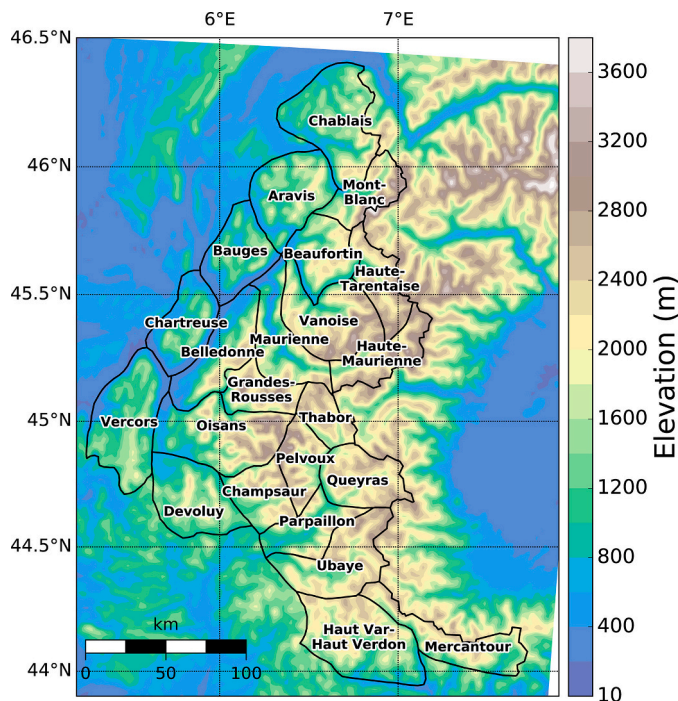
Fig. 1. Map of the French Alps. The area studied is divided into 23 massifs reported on the map.

- **Daily meteorological variables**: daily **precipitation** (rain and snow), **temperature** (daily minimum, maximum, and mean), maximum daily **wind** speed and the associated direction.
- **Additional meteorological variables to take into account recent meteorological conditions** as 3-day moving sums (for rainfall and snowfall) and 3-day moving averages (for air temperature and wind speed) of their original reanalysed daily series. This 3-day time scale is generally considered as being the most relevant for characterizing the severity of winter storms (Schweizer et al., 2009).
- **Snowpack-related variables**: the **snow depth**, **thickness of surface wet snow** and the **thickness of recent surface dry snow**, computed for the four main aspects (north, east, south, and west), on a 40 degree slope. These variables are derived from outputs of the detailed snowpack model Crocus fed by SAFRAN reanalysed meteorological conditions (Durand et al., 1993; Brun et al., 1992; Vionnet et al., 2012).

The thickness of surface wet snow is defined, starting from the top of the snowpack downwards, as the sum of the vertical component of the thickness of the contiguous wet-snow layers characterized by a liquid water content greater than 0.5% by volume. The thickness of the recent surface dry snow is defined as the vertical distance between the snowpack surface and the deepest snow layer characterized by a dendricity greater than 0.25. The threshold expressed in terms of dendricity (Brun et al., 1992) ensures that the considered snow layer still features characteristics of precipitation particles or decomposed fragments (Fierz et al., 2009), and it accounts for the impact of snow metamorphism on snow layers.

Total snow depth corresponds to the entire snowpack, while the thickness of surface wet snow and recent surface dry snow are computed from the uppermost snow layers. Consequently, these variables are not necessarily correlated. Wind direction is the only categorical variable whose categories are the directions *north*, *south*, *east* and *west*. All other variables are continuous.

Eventually, the list of all our variables abbreviated by *dirwmax, Precip.rain, Precip.snow, Tmin, Tmax, Tmean, wmax, hsnow.at.ground, thickness.snow.wet, thickness.snow.dry*, is: wind direction, rain

precipitation, snow precipitation, minimum temperature, maximum temperature, average temperature, wind speed, snow height, thickness of wet snow and thickness of dry snow, respectively. In this list, the names of variables standing for the 3-day moving sum series end with the extension *.cumul.3 days* and those of the variables standing for the 3-day moving average series end with the extension *.average.3 days*. In the variable names, the letters N, S, E, W stand for the aspects north, south, east, west, respectively. The quantities 1800 m, 2400 m, 3000 m stand for the elevations in meters (Table 1).

2.2. Avalanche activity in the French Alps

2.2.1. Systematic records from EPA in the French Alps

In the French Alps, daily observed avalanche data can be found in the *Enquête Permanente sur les Avalanches (EPA)* which describes avalanche events on approximately 3900 paths since the beginning of the 20th century (see Mougin, 1922; Bourova et al., 2016). Quantitative (runout elevations, deposit volumes, etc.) and categorical (flow regime, snow quality, etc.) information is recorded in the report. Obviously, EPA registers only a small proportion of all avalanches that occur in the French Alps. A recent study has indeed showed that during a unique (major) avalanche cycle, about 19,000 avalanches can occur over Switzerland (Bühler et al., 2019), whereas in the EPA, the mean number of avalanches per winter in the French Alps is around 1000. The EPA only records the avalanche activity that are visible from the valleys, so that, with regards to the whole natural activity, the proportion of avalanches that occur at high elevation and when visibility or accessibility is insufficient for the rangers in charge of observation may be underestimated. More generally, EPA's quality varies in time and space due to individual differences in local observations.

Despite these known limitations, the EPA clearly stands among the worldwide longest and most comprehensive records of natural avalanche activity. Studies by Jomelli et al. (2007); Eckert et al. (2010a, 2010c, 2013); Lavigne et al. (2012, 2015) that include further discussions of EPA record's strengths and weaknesses already demonstrated its interest for geophysical inference. Specifically, the joint analysis of these avalanche records together with the Durand et al. (2009a, 2009b) reanalyses supplemented by snow-climate change scenarios showed its relevance for grasping the main patterns of real avalanche activity in space and time in the French Alps and understanding their main climatic drivers (Castebrunet et al., 2012, 2014). Emerging specific studies of severe avalanche episodes at shorter time scales (see Eckert et al., 2010b; Dkengne Sielenou et al., 2016) lead to similar conclusions. All in all, even if some discrepancies with regards to the spatio-temporal characteristics of real activity certainly exist, it can be assumed that they are small enough for the EPA record to be considered as a reasonable surrogate. Hence, their main features, and notably where and when avalanches occur in the EPA, are worth being learned from their snow and weather drivers in order to get insights regarding the variations of real activity in space and time over the French Alps.

2.2.2. Partitioning daily avalanche occurrence data in three activity classes

For this study, only the information about the daily numbers of avalanches is used as the most straightforward variable to describe avalanche activity. We use daily counts aggregated at the massif scale for the same period as the snow and meteorological variables, namely the 51 winter seasons from 1958/59 to 2008/09. This corresponds to 46,610 avalanches, with 220 to 1,844 avalanches per winter throughout the Alps and 369 to 4,554 avalanches per massif during the entire period. For certain avalanches, the exact date of occurrence is unknown, and the observer then provides a time window of a few days. We used here the end of this time window as we are then sure that the avalanche actually occurred. This temporal uncertainty is mitigated by the use of cumulated/averaged snow and meteorological conditions over the three past days.

Formally, we denote S the dataset containing, for a given massif, the

Table 1

Snow and meteorological variables used in this study.

	Variable type	Variable name
1		dirwmax.1800 m
2	Wind direction	dirwmax.2400 m
3		dirwmax.3000 m
4		Precip.rain.1800 m
5	Rain precipitation	Precip.rain.2400 m
6		Precip.rain.3000 m
7		Precip.rain.1800 m..cumul.3 days
8	3-day moving sum of rain precipitation	Precip.rain.2400 m..cumul.3 days
9		Precip.rain.3000 m..cumul.3 days
10		Precip.snow.1800 m
11	Snow precipitation	Precip.snow.2400 m
12		Precip.snow.3000 m
13		Precip.snow.1800 m..cumul.3 days
14	3-day moving sum of snow precipitation	Precip.snow.2400 m..cumul.3 days
15		Precip.snow.3000 m..cumul.3 days
16		Tmin.1800 m
17	Minimum temperature	Tmin.2400 m
18		Tmin.3000 m
19		Tmin.1800 m..average.3 days
20	3-day moving average of minimum temperature	Tmin.2400 m..average.3 days
21		Tmin.3000 m..average.3 days
22		Tmean.1800 m
23	Mean temperature	Tmean.2400 m
24		Tmean.3000 m
25		Tmean.1800 m..average.3 days
26	3-day moving average of mean temperature	Tmean.2400 m..average.3 days
27		Tmean.3000 m..average.3 days
28		Tmax.1800 m
29	Maximum temperature	Tmax.2400 m
30		Tmax.3000 m
31		Tmax.1800 m..average.3 days
32	3-day moving average of maximum temperature	Tmax.2400 m..average.3 days
33		Tmax.3000 m..average.3 days
34		wmax.1800 m
35	Wind speed	wmax.2400 m
36		wmax.3000 m
37		wmax.1800 m..average.3 days
38	3-day moving average of wind speed	wmax.2400 m..average.3 days
39		wmax.3000 m..average.3 days
40		hsnow.at.ground..1800..aspects.N.
41		hsnow.at.ground..2400..aspects.N.
42		hsnow.at.ground..3000..aspects.N.
43		hsnow.at.ground..1800..aspects.E.
44		hsnow.at.ground..2400..aspects.E.
45	Snow depth	hsnow.at.ground..3000..aspects.E.
46		hsnow.at.ground..1800..aspects.S.
47		hsnow.at.ground..2400..aspects.S.
48		hsnow.at.ground..3000..aspects.S.
49		hsnow.at.ground..1800..aspects.W.
50		hsnow.at.ground..2400..aspects.W.
51		hsnow.at.ground..3000..aspects.W.
52		thickness.snow.wet..1800..aspects.N.

Table 1 (*continued*)

	Variable type	Variable name
53		thickness.snow.wet..2400..aspects.N.
54		thickness.snow.wet..3000..aspects.N.
55		thickness.snow.wet..1800..aspects.E.
56		thickness.snow.wet..2400..aspects.E.
57	Thickness of wet snow	thickness.snow.wet..3000..aspects.E.
58		thickness.snow.wet..1800..aspects.S.
59		thickness.snow.wet..2400..aspects.S.
60		thickness.snow.wet..3000..aspects.S.
61		thickness.snow.wet..1800..aspects.W.
62		thickness.snow.wet..2400..aspects.W.
63		thickness.snow.wet..3000..aspects.W.
64		thickness.snow.dry..1800..aspects.N.
65		thickness.snow.dry..2400..aspects.N.
66		thickness.snow.dry..3000..aspects.N.
67		thickness.snow.dry..1800..aspects.E.
68		thickness.snow.dry..2400..aspects.E.
69	Thickness of recent surface dry snow	thickness.snow.dry..3000..aspects.E.
70		thickness.snow.dry..1800..aspects.S.
71		thickness.snow.dry..2400..aspects.S.
72		thickness.snow.dry..3000..aspects.S.
73		thickness.snow.dry..1800..aspects.W.
74		thickness.snow.dry..2400..aspects.W.
75		thickness.snow.dry..3000..aspects.W.

series of daily snow avalanche counts z_t. Since we are dealing with snow avalanches, we consider in the sequel only elements of S corresponding to the winter season, defined as the period ranging from the 15th of October of a given year to the 15th of May of the following year. Given a threshold h, the set S can be divided into three classes as:

$$\begin{cases} \text{Null class}: & \{z_t \in S : z_t = 0\} \\ \text{Moderate class}: & \{z_t \in S : 1 \le z_t \le h\} \\ \text{High class}: & \{z_t \in S : z_t > h\} \end{cases}$$

These three classes correspond to days with no, moderate or high avalanche activity. In each massif, we define the threshold h as a 2-year return level, namely the number of avalanches per day exceeded on average once every two years/winters. This two-year threshold was chosen as a compromise: sufficiently high to isolate high activity days and not too high to keep enough days in the high class. Table 2 illustrates the variability from one massif to another of the threshold h according to the specific avalanche activity of each massif. Table 2 also shows the size (number of elements) of each class, highlighting that, in each massif, they are highly unbalanced, with a factor 6 to 60 between the null and the moderate class, 10 to 60 between the moderate and the high class and 400 to 600 between the null and the high class. Note that by construction, the size of the high class is about half the number of years of the considered time period.

Our definition of the high activity class by a return level has the

Table 2
Massifs names and identifiers in the French Alps refer to Fig. 1. For each massif: number of monitored avalanche paths in the EPA record, number of avalanches per day exceeded on average every two years (threshold h), number of days in each avalanche activity class.

Massif identifiers (i)	Massif names	Number of paths	Threshold (h)	Size of classes		
				High	Moderate	Null
1	Chablais	193	14	26	1515	9335
2	Aravis	207	14	26	1503	9347
3	Mont-Blanc	167	11	25	703	10,148
4	Bauges	97	10	25	635	10,216
5	Beaufortain	106	5	25	524	10,327
6	Haute-Tarentaise	166	10	24	725	10,127
7	Vanoise	247	14	23	929	9924
8	Maurienne	222	14	23	1449	9404
9	Chartreuse	55	3	17	228	10,631
10	Belledonne	91	7	26	493	10,357
11	Haute-Maurienne	186	12	26	848	10,002
12	Vercors	35	4	17	160	10,699
13	Grandes-Rousses	101	10	24	619	10,233
14	Thabor	74	6	20	279	10,577
15	Oisans	297	22	26	969	9881
16	Pelvoux	107	11	24	434	10,418
17	Queyras	205	7	26	476	10,374
18	Devoluy	107	5	25	315	10,536
19	Champsaur	124	14	24	673	10,179
20	Embrunais-Parpaillon	89	3	20	274	10,582
21	Ubaye	92	11	25	630	10,221
22	Mercantour	191	12	24	636	10,216
23	Haut-Var - Haut-Verdon	65	4	26	326	10,524

advantage to be independent of the chosen spatial scale and can be applied to different types of data, to allow comparisons between regions and variables. With this definition, a few avalanche occurrences may be considered a significant episode in an area where avalanche activity is usually low. Conversely, a particularly high number of avalanches is required to belong to the high activity class in an area that normally sees high avalanche activity. This approach, solely based on massif-scale avalanche numbers, deviates from the usual implementation of the international avalanche danger scale which relates danger to avalanche activity in a more absolute way, independently of local considerations, and also includes information related to avalanche types and size (Morin et al., 2019; Schweizer et al., 2020). However, the goal of this study is not to automatically predict the avalanche danger level which also considers potential accidental triggers. Instead, we aim to develop a refined avalanche statistical forecasting model whose outputs could, in the future, assist forecasters in their daily duties to predict natural avalanche activity (Sect. 5). Besides, our choice can be seen as advantageous for operational purposes. In areas where high avalanche activity is usual, operational services are generally well prepared, whereas a few events can cause considerable problems in areas where activity is usually low. Eventually, our relative definition of classes facilitates the comparison from one massif to another of the main drivers of moderate to high activity. Indeed, even in two massifs where activity is rather different (say a massif that is usually affected by many avalanches and another by very few), we thus compare snow and meteorological drivers that induce a similarly rare avalanche activity, whereas comparing the drivers leading to the same number of avalanches would be, in such a case, less meaningful.

3. Methods

3.1. Data preparation

Our initial dataset of reanalysed snow and meteorological conditions contained a few missing values resulting from various numerical (leap years, etc.) and physical (wind direction when the wind speed is zero, etc.) artifacts. Although not numerous, these are sufficient to preclude the use of classification techniques that are not able to handle missing values. Hence, in each massif, these few missing values were imputed

using the method of Audigier et al. (2016) described in Appendix A.1.

The resulting set of snow and weather covariates is a mixture of continuous and categorical variables, most of them highly correlated. This is not suitable for many statistical learning methods as it may lead to unstable classifiers. To overcome this difficulty, we transformed in each massif the imputed table of snow and weather data into a new dataset of only continuous variables by performing a Factorial Analysis for Mixed Data (FAMD) (see Pagès, 2004; Husson et al., 2010b; James et al., 2013). This preliminary step presented in Appendix A.2 denoises the covariates, puts all of them on equal footing (standardized quantitative variables) and decorrelates them in order to get stable classifiers.

3.2. Class balancing

Unbalanced dataset problems occur in classification when the number of instances in certain classes is much lower than the instances in the other classes. The main challenge is that the small classes are often more useful, so that it is costly to misclassify examples of these minority classes. Most machine learning algorithms work well with balanced datasets since they aim at optimizing the overall classification accuracy or a related measure. For unbalanced datasets, however, decision rules established by standard machine learning algorithms tend to be biased towards the majority classes. Therefore, the minority classes are likely to be misclassified.

Many attempts have been made to deal with classification of unbalanced datasets. They can be categorized into two groups (see Nguyen et al., 2009; Ramyachitra and Manikandan, 2014) namely *internal approaches* and *external approaches*. Internal approaches are algorithmically explicit, meaning that one creates innovative classification algorithms (or at least adapts existing ones) in order to take the class unbalanced problem into consideration. For example, an internal approach is used in Möhle et al. (2014) where avalanche forecasting is attempted using a modified class-balancing random forest algorithm. On the contrary, in external approaches, one pre-processes the data in order to equilibrate the size of each class without modifying the classification algorithm.

In this work, we wanted to compare several classification methods. To balance the dataset once for all, an external approach was therefore preferable. We implemented a simple generic external approach to

balance our dataset before applying machine learning techniques. For each massif, the data belonging to moderate and high classes was extracted and copied as many times as requested to become as close as possible to the size of the null class. Eventually, a subset of the initial table was copied to generate a fully balanced dataset with the same number of lines attributed to the three classes. Formal details on the procedure and the choice of the subset to complete the process are available in Appendix A.3.

3.3. Classification

3.3.1. Competing methods

In this work, the following classification methods were compared: the linear discriminant analysis (LDA) (see Venables and Ripley, 2002), the support vector machine with linear kernel (SVM_LK), polynomial kernel (SVM_PK) and radial kernel (SVM_RK) (see Meyer et al., 2014), the classification trees (Trees) (see Ripley, 2014), the K-nearest neighbors (K-NN) with $K = 1$ (see Venables and Ripley, 2002), and the random forests for classification (RF) (see Liaw and Wiener, 2002). The latter is a large collection of decision trees so that the most frequent class predicted by these decision trees is the one which is predicted by the random forest (see Breiman, 2001; Liaw and Wiener, 2002; James et al., 2013; Salford-Systems, 2014). It is out of the scope of this paper to go into details in their mathematical description but a general introduction relevant for most of them can be found, e.g. in James et al. (2013).

As explained before, all these methods were combined with the same external balancing scheme (Appendix A.3). However, for comprehensiveness and consistency of the demonstration, an internal balancing approach, namely the weighted random forests (wRF, Chen et al., 2004) already used for avalanche forecasting (Möhle et al., 2014), was also implemented. The different classifiers were implemented on transformed avalanche drivers which are the first principal components responsible, in each massif, for at least 95% of the total variability in the meteorological and snow variables (Appendix A.2).

3.3.2. Classification evaluation

Predictive performance of all classifiers was assessed using Leave-One-Out Cross-Validation (LOO_CV) error rates (see Stone, 1974; James et al., 2013). It is worth noting that there is a straightforward way to estimate the test error of a RF model (see James et al., 2013), but we did not use it to make comparison with other classifiers easier. Specifically, predictive performance is evaluated on the ability to discriminate the null, moderate and high classes measured by the confusion matrix introduced below for each of the 23 massifs of the French Alps.

In a multi-class application of k classes (with $k \geq 2$), the confusion matrix is the simplest way of measuring or visualizing the performance of a classifier. It can be presented in the form of Table 3, where C_i denotes the class label of the i-th class. It is a square matrix $k \times k$, where k is the number of classes (in our example of application, k equals 3). Each row corresponds to one of the actual categories of the cases. Each column corresponds to one of the categories in which an element is classified by the system (maybe incorrectly). Each cell (i, j) of the matrix contains the number $n_{i, j}$ of cases from the evaluation (or testing) set that actually belongs to the category i and were classified as category j. In this way, $n_{i, i}$ is the number of well-classified instances in the category i.. The **class i success rates** (CSR$_i$) are defined by.

Table 3
Confusion matrix.

		Predicted class		
	Classes	C_1	⋯	C_k
True Class	C_1	$n_{1, 1}$	⋯	$n_{1, k}$
	⋮	⋮	⋱	⋮
	C_k	$n_{k, 1}$	⋯	$n_{k, k}$

$$\text{CSR}_i = \frac{n_{i,i}}{\sum_{j=1}^{k} n_{i,j}}, \quad i = 1, 2, ..., k$$

from which one deduces the class i error rates as $1 - \text{CSR}_i$, $i = 1, 2, ..., k$. The **overall success rate** (OSR) is then calculated as the mean of the three classes.

3.4. Variable importance in avalanche activity classes and subsequent massif clustering

Subsequently, in order to work with interpretable drivers, we repeated the RF classification with the original covariates table in all 23 massifs. We then quantified variable importance in the random forests classification in the 23 massifs. Eventually, we clustered the 23 massifs with respect to the snow and meteorological variable importance in the null, moderate and high avalanche activity class.

3.4.1. Variable importance in random forests

When applying the RF model for classification, it is no longer possible to represent the resulting statistical learning procedure. Hence, random forest classifiers improve prediction accuracy at the expense of interpretability. However, although the collection of many decision trees is much more difficult to interpret than a single tree, one can obtain an overall summary of the importance of each predictor in a RF. The method is based on measuring the damage that would be done if one loses access to true values of a given predictive variable. To simulate losing access to a predictor, one randomly scrambles its values in the data. That is, one moves the value belonging to a specific day of the data to another day. One permutes just one predictor at a time and measures the consequential loss in predictive accuracy (mean decrease in accuracy). A large value indicates an important predictor (Genuer et al., 2010; Gregorutti et al., 2016).

3.4.2. Clustering

For interpreting the results of variable importance in each massif, a clustering method is proposed to identify groups of massif with similar repartition of variable importance. In this study, as we do not want to presume a number of clusters, we first use hierarchical clustering. In addition, it has the advantage of resulting in an attractive tree-based representation of the observations, called a dendrogram (James et al., 2013; Kaufman and Rousseeuw, 1990). Then, the obtained clusters are refined with the k-means clustering algorithm.

The *hierarchical clustering* dendrogram is obtained via an extremely simple algorithm based on a defined dissimilarity measure between each pair of observations or between two groups of observations (Murtagh and Legendre, 2014). Hierarchical trees considered in this paper use the Ward's criterion (Ward, 1963). This criterion is based on the Huygens theorem which allows one to decompose the total inertia (total variance) in between groups and within groups variance. The Ward's method consists in aggregating two clusters such that the growth of within-inertia is minimum (in other words minimizing the reduction of the between-inertia) at each step of the algorithm. The within-inertia characterizes the homogeneity of a cluster. The hierarchy is represented by a dendrogram which is a hierarchical tree considered as a sequence of nested partitions from the one in which each individual is a cluster to the one in which all the individuals belong to the same cluster.

Choosing the *number of clusters* is a core issue and several approaches have been proposed. The rule used here is based on the growth of inertia (Husson et al., 2010a). It suggests a division into K clusters when the increase of between-inertia between $K - 1$ and K clusters is much greater than the one between K and $K + 1$ clusters. An empirical criterion can formalize this idea. Let $\Delta(K)$ denote the between-inertia increase when moving from $K - 1$ to K clusters. The criterion proposed is the number K which minimized the ratio $\Delta(K)/\Delta(K + 1)$.

One further *consolidates the clusters* obtained via the hierarchical

clustering method by using the k-means clustering algorithm where the initial starting centroids are the centres of gravity of the k clusters resulting from the hierarchical clustering. Doing this, it is possible that a few individuals move from one cluster to another. Hence, the hierarchical clustering is used here as a pre-processing step for the identification of the relevant number of clusters and of the initial centroids, which guarantees the convergence of the k-means algorithm to the optimal (highest between-cluster inertia) and stable clusters.
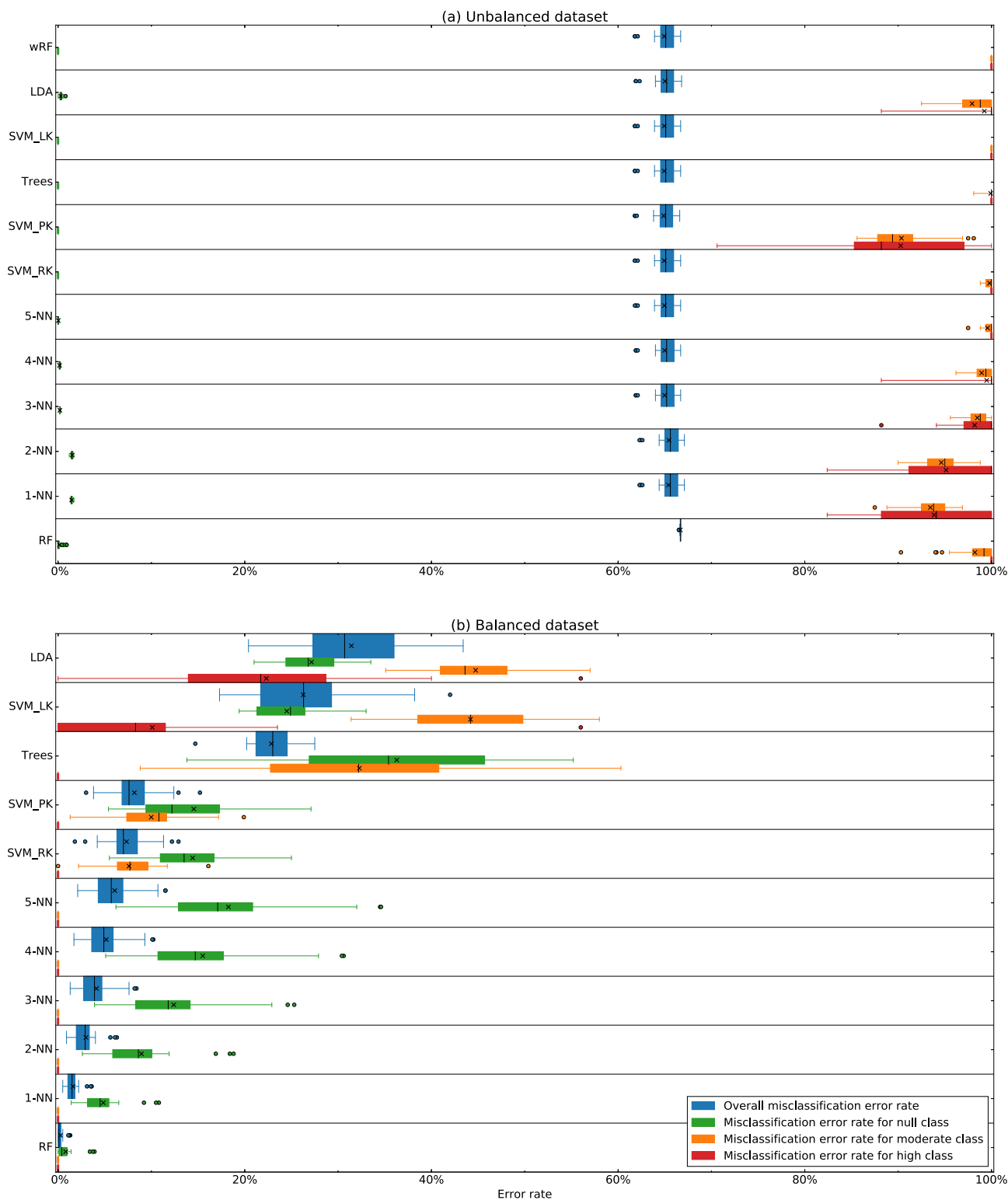

Fig. 2. Inter-massif variability in misclassification error rates, for (a) unbalanced and (b) balanced datasets. Are reported the overall misclassification error rate (blue) and the corresponding values for the three classes (null, in green, moderate in orange and high in red). Refer to Table B.1 for the full names of the competing classification methods. The boxes span the inter-quartile range from the 1st to 3rd quartile with the vertical line showing the median. The whiskers show the range of observed values that fall within 1.5 times the interquartile range and the black dots are outliers above or below it. Black crosses denote the mean value.

4. Results

4.1. Comparing classifiers on transformed avalanche drivers

4.1.1. Unbalanced case

We started by applying all classification methods listed in Section 3.3 without balancing instances from the three avalanche classes. Overall misclassification error rates are shown in Table B.1. Graphical overviews of the misclassification error rates including the class error rates are displayed in Fig. 2a. Generally, in the French Alps, and regardless of the learning method used, observations correctly classified are mainly instances from the majority class (the null class). On the contrary, all classifiers including the weighted random forests (wRF, internal approach) fail to correctly identify days belonging to the minority classes (the moderate and high classes), namely avalanche days. The balancing step (external approach) appears therefore a key and a necessary step for the success of the prediction.

4.1.2. Balanced case

We carried out the external balancing method described in Section 3.2 and then applied again each of the various classification methods presented in Section 3.3. Resulting overall error rates are given in Table B.2. The box plots in Fig. 2b summarize by-class misclassification error rates. With regards to the unbalanced case, they are much lower. Among the competing methods, the linear discriminant analysis, the classification trees and the support vector machine with linear kernel perform poorly. By contrast, the random forest and the nearest neighbour classifiers appears to be more relevant here. In what follows, we therefore only consider the random forest for classification.

4.2. Variable importance in Random forests and massifs clustering

The box plots in Fig. 3 show the inter-massif variability of the mean decrease in accuracy when loosing access to a variable, which is a way of evaluating the inter-massif variability in the importance of each variable with respect to the overall accuracy. In the French Alps, the weakest predictors of avalanche activity class membership are the variables related to the thickness of wet snow while the most important predictors are the variables related to the total height of snow, the thickness of dry snow and precipitation. However, variability between the different massifs with respect to the variable importance in the classification of snow avalanche days by the random forest is high.

To sum up results, we performed a clustering analysis as detailed in Section 3.4. Hierarchical clustering led to three clusters of massifs (Fig. 4a), and the subsequent k-means algorithm to the final map of Fig. 4b. The obtained three groups of massifs are very homogeneous in the sense that the within cluster variable importances are very close. Note that only one massif moves from one cluster to another in the k-means step, the Thabor massif. Strikingly, in each group, the massifs are also very homogeneous in terms of spatial/and or elevation characteristics. For instance, cluster three (massifs in green) includes the lower elevation massifs of the French Prealps, whereas cluster 1 (in blue) consists in four contiguous massifs of higher, rather similar, topographical characteristics. This demonstrates that in massifs similar in terms of localisation and elevation characteristics, main snow and meteorological drivers of avalanche activity are very similar, an intuitive but consistent result.

5. Discussion

In this work, we revisited the problem of classifying snow and meteorological conditions according to their consequences in terms of natural avalanche occurrences. Specifically, we considered three avalanche activity classes, namely the null, the moderate and the high classes, the two latter on the basis of a two-year return level threshold. We implemented an external class-balancing scheme and then built
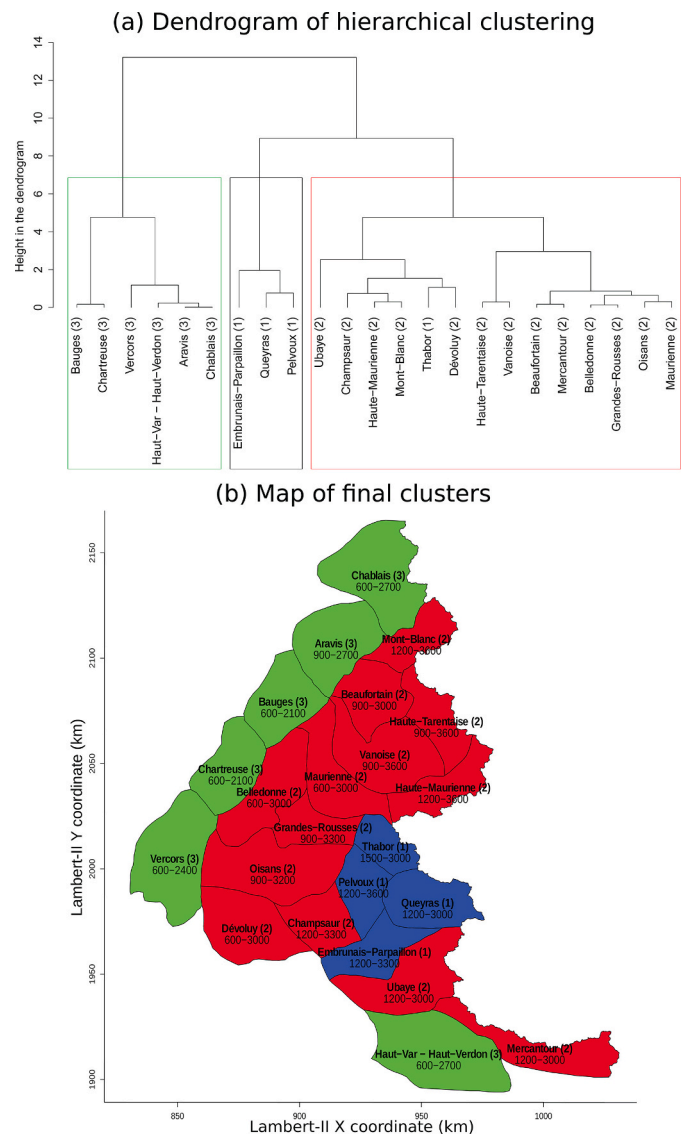

(a) Dendrogram of hierarchical clustering

(b) Map of final clusters

Fig. 3. Mean decrease in accuracy when loosing access to a variable (as described in Section 3.4.1). It is a measure of the variable importance in the retained classification approach with respect to the overall accuracy in the classification of avalanche days. Variables are grouped by types separated by dotted lines. Each box represents a variable for one elevation and one aspect (W, S, E, N) if relevant. Additional variables to take into account recent conditions (3-days average for wind and temperature and sums for precipitation) are identified with"3d". These box plots highlight the inter-massif variability in variable importance. For box plot definition, see Fig. 2.

various classifiers to predict the class membership on the basis of the snow and meteorological data available at the massif scale. We also ran all the classification approaches without the class-balancing step.

5.1. Random forest as a relevant machine learning method for avalanche activity classification

The results of the different machine learning methods points out random forest as a relevant method for learning avalanche activity classes (Fig. 2b and Table B.2) with a very low error rate according to our metric in each of the 23 massifs of the French Alps. These results are significantly different from those obtained with the other tested machine learning methods, especially for misclassification of non-avalanche days. In each of the 23 massifs, the comparison was made using more than 50 years of high quality daily data. To use it with similar efficiency
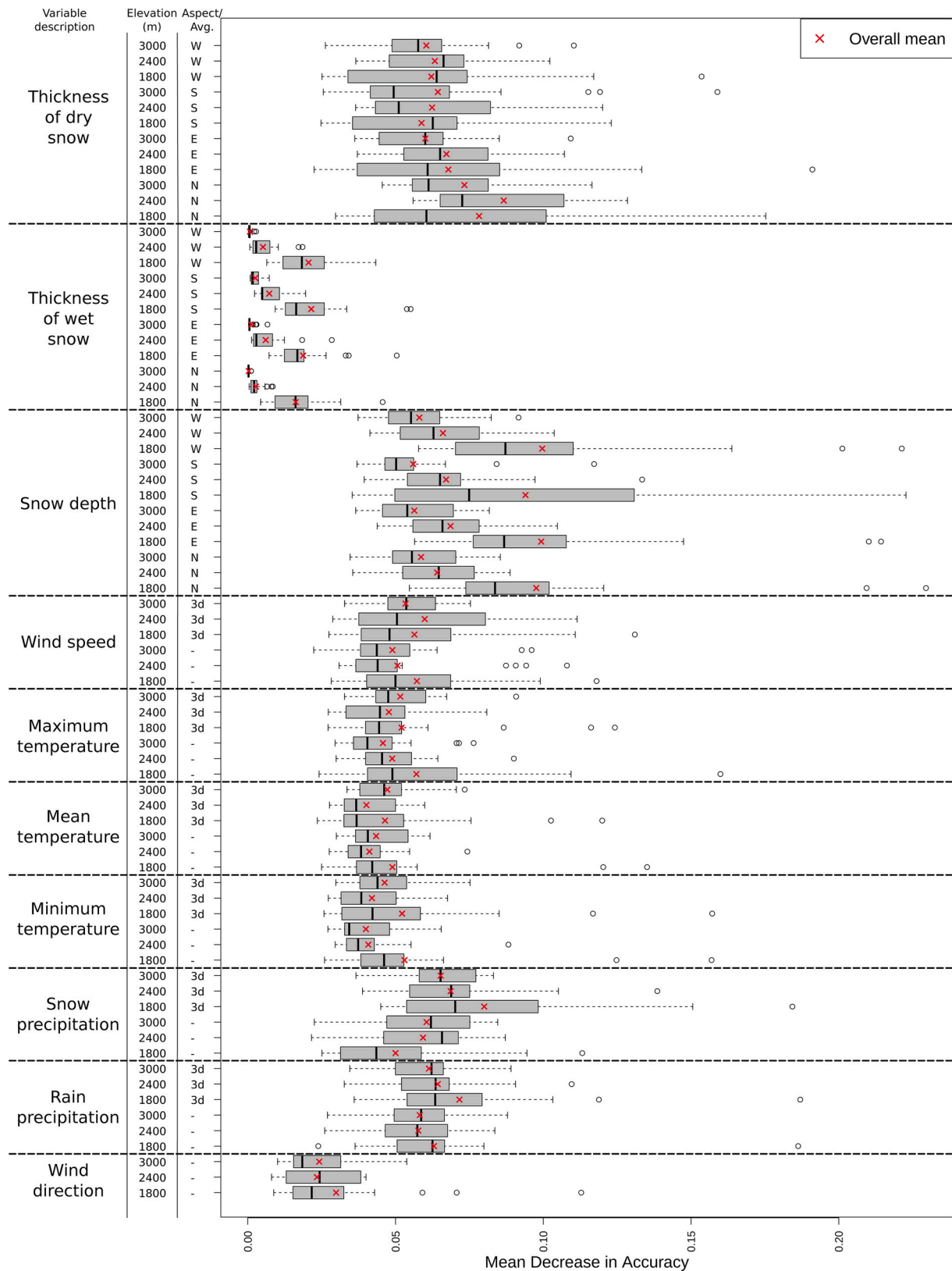
Fig. 4. (a) Hierarchical clustering of the 23 massifs in the French Alps with respect to the variable importance results obtained with the retained approach. Subsequent application of the k-means algorithm as detailed in Section 3.4.2 led to the final three groups of massifs identified by the numbers in brackets. (b) Map of the clusters identified in (a) using the hierarchical clustering and the k-means algorithm. In each massif, the number after the massif name stands for the cluster label while the values below the name refer to the elevation ranges (in meters) over which the numerical SAFRAN meteorological reanalysis is performed. Note that the cluster belonging is modified by the k-means algorithm for one massif only (Thabor, from cluster 2 to cluster 1).

in other regions, since machine learning methods need a quite important dataset and are sensitive to input data, it would be necessary to have similar data at hand. In such cases, we believe (also this remains to be checked) that the classification approach we propose would be powerful/flexible enough to perform well.

However, the results have to be handled with care as the evaluation is done leaving only one observation out. Hence, with our class balancing approach, some copies of this observation may be present in the evaluation set for moderate and high classes. As a consequence, in the balanced case, nearest neighbors for moderate and high class (at least for low values of *N*) are the same observation, which explains the error rate of zero for these two classes. More widely, from this perspective, results of Fig. 2a and Table B.1 are not really comparable with those of Fig. 2b and Table B.2. Values of error rates for Fig. 2b should therefore be considered as indicators of the relative interest of the competing methods, at least for the moderate and high classes, rather than rates to be expected in a true"blind prediction context". Hence, further work could deepen the evaluation of the different classifiers using even more stringent cross-validation approaches. More generally, our work may constitute a good starting point for further inter-comparisons of avalanche statistical forecasting models, a key question for their usefulness in an operational context (Sect. 5.4).

5.2. Classifiers and clustering compared to main drivers of avalanche activity

The importance of variables in the classification could be compared to the main known drivers of avalanche activity. To do so, we moved, with the retained random forest model, from the principal component space to the space of the original snow and meteorological variables. The main variables identified (Fig. 3) correspond to the current literature (Schweizer et al., 2003a; Castebrunet et al., 2012; Birkeland et al., 2001): total snow depth, thickness of dry snow and precipitation are the main drivers while wind and temperature are of lower importance. This contributes to the confidence it is possible to have in the ability of random forest methods to reproduce avalanche activity.

Applied to each massif, this method led to point out the variables that predominantly generate moderate or high avalanche activity days in each massif. This further allowed clustering the 23 massifs of the French Alps into 3 groups rather homogeneous in terms of main explanatory variables of avalanche activity. Within each group, massifs are relatively close in terms of geographical location and topography, notably elevation. This determines their exposure to synoptic patterns and the characteristics of prevailing snow conditions as a function of elevation, aspect and latitude/longitude, which explains why they behave similarly in terms of avalanche activity response to snow and meteorological drivers. Notably, the clusters' localisation roughly mimics spatial patterns in extreme snowfall in the French Alps (Gaume et al., 2013).

Much more could certainly be said from these results but this is arguably enough to depict the potential of our approach to infer the main drivers of avalanche activity. Similar studies could therefore be conducted in various contexts (snow-climate conditions, topography, etc.) to guide further research on physical understanding of avalanche release. However, one needs to remain careful when trying to infer underlying physical processes driving avalanche activity as such statistical methods rely on correlations which may not be mirrored to causality links.

5.3. Definition of classes of avalanche activity

The analysis performed in this study was facilitated by the use of a threshold separating the high and moderate activity classes corresponding to the same frequency of occurrence in all massifs. However, our separation between the null and moderate class does not correspond to the same occurrence frequency in all massifs. To overcome this difficulty, a threshold of a few avalanches in the most active massifs could

have been chosen instead of a universal"more than zero" threshold. Our choice was made to clearly separate no activity from sporadic to moderate activity in all massifs, since the information"fully safe" versus"-potential activity, even very limited" may always be useful, even if the"no activity conditions" may indeed be more stringent (less frequently encountered) in the massifs with more intense avalanche activity.

Independently of the way these three classes are defined (in our case, raw aggregation of avalanche counts, but, different or more sophisticated variations could be proposed and tested, see below), which of these is to be expected a given day may be an information useful but insufficient for many operational problems. However, note, first that it is already better than the even coarser avalanche / non-avalanche day segregation used by most of the statistical forecasting models in the literature. Second, one could easily think, now that we know that our approach works efficiently with three classes, to expand it to a higher class number, in order to test if it is possible to predict even more precisely the avalanche activity to be expected on the basis of the snow and meteorological conditions. An obvious choice would be to consider a higher number of occurrence number classes. Yet, segregating avalanche activity by elevation range and/or aspect/slope would be another appealing possibility which may help bridging the gap between the outputs of our approach and the information usually required to assess avalanche danger and prepare an *avalanche bulletin*.

5.4. Operational usefulness and limitations

Avalanche operational forecasting and *an avalanche statistical forecasting model* remain two different things. However, a statistical forecasting model has clearly the ambition to provide useful entries to operational forecasting, and the developed approach could be, in the future, systematically implemented to assist avalanche forecasters in preparing *avalanche bulletins*, as a complement to their expertise and to the different stability indices already at hand. For instance, in France, MEPRA (Giraud et al., 1998) is an expert system integrated into SAFRAN-Crocus model outputs for operational avalanche forecasting (Durand et al., 1999). Formally, MEPRA provides local indices for each massif, elevation range and aspect on the basis of physical and empirical rules. The different values obtained by elevation range and aspect are also aggregated to obtain a single index value at the massif scale. The overall efficiency and physical consistency of the results we obtained in the 23 massifs under study suggests that our approach could be a useful complement for the practitioners to these standard MEPRA indices (possibly after subsequent developments regarding the class definition). Specifically, for each daily snow and meteorological condition, our approach would provide, for each massif, a reliable hint about the avalanche activity to be expected: very low (null class), standard (moderate class), or (very) marked (high class). In turn, such a practical implementation could help to further understand why, for some massifs, the classification rates remain slightly lower than for others.

A limitation of the present work for operational use is that the considered snow and meteorological data resulted from reanalyses, by definition not available for real-time forecasting. Hence, the classification efficiency should now be tested in forecast mode (Durand et al., 1999; Vernay et al., 2015) to study the impact of forecasted snow and meteorological information on the method accuracy. Another limitation of our study is that it makes full use of the extensive data available in France: very detailed snow and meteorological data all over the French Alps at the massif scale from the SAFRAN-Crocus chain supplemented by the comprehensive EPA avalanche database, allowing an efficient learning of the conditions in which moderate and high avalanche activity actually occur. The approach could now be tested in contexts that differ in terms of data structure, spatial coverage and/or exhaustivity to evaluate whether or not the proposed classification method is really able to outperform other more classical techniques.

Eventually, it could be objected that avalanche activity is measured

in our approach in terms of avalanche occurrence numbers from the EPA database only. This raises the question of the representativity of EPA records with regards to the real activity discussed in Sect. 2. Also, the class predicted in our approach will not give any information about the severity of avalanche activity in terms of the magnitude (runout, velocity field, etc.) of the avalanches themselves, and will, for instance, consider an episode with many small avalanches as more severe than an episode with a lower number of larger avalanches. A first improvement would be to integrate additional avalanche activity data within the approach, which would be a step towards classifiers accounting for real activity in an even more comprehensive way. For instance, Castebrunet et al. (2012) proposed a composite index combining EPA counts with the MEPRA stability index, more reliable at high elevation and during severe winter conditions (snow storms, cold temperatures, etc.). Building new classifiers with this composite index as target would be straightforward, presumably reducing the biases towards real activity. Other information sources such as remote sensing observations of deposits could be used in a similar spirit (Eckerstorfer and Malnes, 2015; Bühler et al., 2019). Second, as said before, more complex summaries of avalanche activity more relevant for specific operational problem could be defined and targeted, such as avalanche numbers weighted by avalanche size (deposit volume or size class, McClung and Schaerer, 1981) that are routinely used in avalanche *bulletins* in several countries.

6. Conclusion and outlooks

In this work, we tackled the problem of classifying snow and meteorological conditions according to their consequences in terms of natural avalanche occurrence, and more precisely in three classes (no-avalanche days, and two classes of avalanche days separated by a two-year return level). We compared all the main statistical classifying approaches currently described in the literature and demonstrated that random forest classifiers could be relevant. The variables of higher importance in the random forest classifiers have been shown to be coherent with current literature and a clustering based on these variable importances separates massifs which are known to have different avalanche activities (due to elevation, geographical and geomorphological particularities).

These results show the potential of statistical methods to interpret snow cover model outputs in terms of expected avalanche activity or return level. This may be used in the future as an additional guiding tool for practitioners or avalanche forecasters in their analysis. We worked here on three classes, but our approach could be extended to different (and more relevant) classes. Eventually, it would be interesting to go beyond the deterministic approach developed here, modelling the probability of belonging to each class rather than attributing only one class to each meteorological and snow condition. Comparing these approaches may help to break the walls, somewhat artificial, that still exist between different forecasting approaches and index classes (physical or empirical, local or regional, deterministic or probabilistic), with the ultimate goal to reach more efficient and reliable prediction of avalanche activity according to snow and meteorological conditions.

Authors contribution

Conceptualization, methodology and funding acquisition: NE, PN and SM; Data curation: NE and PH; Investigation and formal analysis: PS and LVG; Writing - original draft: PS, LVG, PH and NE; Writing - review & editing: all co-authors.

Declaration of Competing Interest

None.

Appendix A. Details on statistical pre-processing

A.1. Imputation of missing data

The method proposed in (Audigier et al., 2016) was used to impute the few missing snow and weather data to generate in each massif the complete set of covariates required for several of the used machine learning techniques. Implementation was performed using the package **missMDA** (Josse and Husson, 2016) of the R software (function ***imputeFAMD()***). The method works as follows:

1) There is initially a table of mixed data with missing values.
2) This table is transformed to obtain the matrix *X* coding categorical variables using an indicator matrix of dummy variables. A missing value on a categorical variable then leads to a row of missing values in the indicator matrix. Continuous variables are left unchanged.
3) This data table is imputed according to the following algorithm:
   i) Set the missing elements at initial values, for example, the mean of the variable for the continuous variables and the proportion of the category for each category using the non-missing entries. Note that concerning the categorical variables, initial values can be non integer ones but the sum of the entries corresponding to one individual and one categorical variable must equal one;
   ii) Perform a Principal Component Analysis (PCA) on the completed dataset;
   iii) Input the missing values with values predicted by the reconstruction formula (defined by the fitted matrix obtained with the axes and components) using a predefined number of dimensions;
   iv) Repeat the steps ii) and iii) on the newly obtained matrix until the total change in the matrix falls below an empirically determined threshold.
4) At the end of the algorithm, imputed values for the missing entries for the categories are not equal to 0 and 1 but are real numbers. However, the constraint in the initialization step makes that categorical variables are imputed with the most plausible values.

A.2. Factorial analysis for mixed data (FAMD)

The principle of FAMD is to obtain independent variables from potentially correlated variables and to balance the influence of the continuous and the categorical variables in the analysis (see Pagès, 2004; Husson et al., 2010b; James et al., 2013). The rationale is to weight the variables in such a

way that each variable of both types contributes equivalently to the construction of the dimensions of variability. It is the same idea as scaling for continuous variables in Principal Component Analysis (PCA), adapted to the use with combined continuous and categorical variables. The technique works as follows:

 i) The first step of FAMD consists in coding categorical variables using the indicator matrix of dummy variables. Continuous variables are conserved;

 ii) The second step is a weighting step: each continuous variable is standardized and each dummy variable is divided by the squared root of the proportion of individuals for which the value of the considered dummy variable is equal to one;

 iii) The final step consists in performing a PCA on the resulting weighted matrix.

FAMD was performed in all massifs on the imputed snow and weather table. In each massif, the smallest number of the first principal components responsible for at least 95% of the total variability were kept as new covariates to be used in the classification step. The last remaining components of the FAMD were considered as noise and eliminated.

A.3. Class balancing

To balance in each massif our dataset of null, moderate and high activity days before applying machine learning techniques, we adopted the generic external approach described below:

1. Let $\mathscr{D}$ denotes our class unbalanced data table in which the first p columns contain the values of snow and meteorological avalanche covariates while the last column contains the avalanche null-moderate-high class labels. The integers n_x, n_y and n_z are the total number of winter days over the studied time period belonging to the null, moderate and high class, respectively.

	Covariates			Class
$\mathscr{D}=$	$X_{1,\,1}$	$\cdots$	$X_{1,\,p}$	null
		$\ddots$	$\vdots$	$\vdots$
	$X_{nx,\,1}$	$\cdots$	$X_{nx,\,p}$	null
	$Y_{1,\,1}$	$\cdots$	$Y_{1,\,p}$	moderate
		$\ddots$	$\vdots$	$\vdots$
	$Y_{ny,\,1}$	$\cdots$	$Y_{ny,\,p}$	moderate
	$Z_{1,\,1}$	$\cdots$	$Z_{1,\,p}$	high
		$\ddots$	$\vdots$	$\vdots$
	$Z_{nz,\,1}$	$\cdots$	$Z_{nz,\,p}$	high

2. Let $\mathscr{D}(X)$, $\mathscr{D}(Y)$ and $\mathscr{D}(Z)$ denote the sub-tables corresponding to the null, moderate and high classes, respectively

	Covariates			Class
$\mathscr{D}(X) =$	$X_{1,\,1}$	$\cdots$	$X_{1,\,p}$	null
		$\ddots$	$\vdots$	$\vdots$
	$X_{n_x,\,1}$	$\cdots$	$X_{n_x,\,p}$	null
	Covariates			Class
$\mathscr{D}(Y) =$	$Y_{1,\,1}$	$\cdots$	$Y_{1,\,p}$	moderate
		$\ddots$	$\vdots$	$\vdots$
	$Y_{n_y,\,1}$	$\cdots$	$Y_{n_y,\,p}$	moderate
	Covariates			Class
$\mathscr{D}(Z) =$	$Z_{1,\,1}$	$\cdots$	$Z_{1,\,p}$	high
		$\ddots$	$\vdots$	$\vdots$
	$Z_{n_z,\,1}$	$\cdots$	$Z_{n_z,\,p}$	high

3. Let $\|\cdot\|$ be a norm on $\mathbb{R}^p$. let u be the permutation of the set $\{1,\cdots,n_y\}$ such that

$$\left\| \left(Y'_{u(1),1} \cdots Y'_{u(1),p} \right) \right\| \leq \cdots \leq \left\| \left(Y'_{u(n_y),1} \cdots Y'_{u(n_y),p} \right) \right\|,$$

where $(Y_{j,\,1}' \cdots Y_{j,\,p}') \in \mathbb{R}^p$, $j = 1, \cdots, n_y$ is the coordinate vector of the j-th row of $\mathscr{D}(Y)$ in the system defined by the p principal components resulting from the FAMD applied on $\mathscr{D}(Y)$ in which the class label variable is added.

Let v be the permutation of the set $\{1,\cdots,n_z\}$ such that:

$$\left\| \left(Z'_{v(1),1} \cdots Z'_{v(1),p} \right) \right\| \leq \cdots \leq \left\| \left(Z'_{v(n_z),1} \cdots Z'_{v(n_z),p} \right) \right\|,$$

where $(Z_{k,\,1}{}'\cdots Z_{k,\,p}{}') \in \mathbb{R}^p$, $k = 1, \cdots, n_z$ is the coordinate vector of the k-th row of $\mathscr{D}(Z)$ in the system defined by the p principal components resulting from the FAMD applied on $\mathscr{D}(Z)$ in which the class variable is added.

4. Extend the sub-tables $\mathscr{D}(Y)$ and $\mathscr{D}(Z)$ of the minority classes (moderate and high classes) to the sub-tables $\widetilde{\mathscr{D}}(Y)$ and $\widetilde{\mathscr{D}}(Z)$ by reordering and recycling successively their row elements up to the size n_x of the majority class (the null class) as described below. The row vectors of the extended sub-tables $\widetilde{\mathscr{D}}(Y)$ and $\widetilde{\mathscr{D}}(Z)$ are given by

$$\left(\widetilde{Y}_{j,1} \cdots \widetilde{Y}_{j,p} | \text{moderate} \right) = \begin{cases} \left(Y_{u(r),1} \cdots Y_{u(r),p} | \text{moderate} \right) & \text{if} \quad 1 \le r < n_y \\ \left(Y_{u\left(n_y\right),1} \cdots Y_{u\left(n_y\right),p} | \text{moderate} \right) & \text{if} \quad r = 0 \end{cases} \quad j = 1, \ldots, n_x,$$

where r is the remainder of the Euclidean division of j by n_y.

$$\left(\widetilde{Z}_{k,1} \cdots \widetilde{Z}_{k,p} | \text{high} \right) = \begin{cases} \left(Z_{v(s),1} \cdots Z_{v(s),p} | \text{high} \right) & \text{if} \quad 1 \le s < n_z \\ \left(Z_{v(n_z),1} \cdots Z_{v(n_z),p} | \text{high} \right) & \text{if} \quad s = 0 \end{cases} \quad k = 1, \ldots, n_x,$$

where s is the remainder of the Euclidean division of k by n_z. Hence, we get:

	Covariates			Class
$\widetilde{\mathscr{D}}(Y) =$	$\widetilde{Y}_{1,1}$	$\cdots$	$\widetilde{Y}_{1,p}$	moderate
		$\ddots$	$\vdots$	$\vdots$
	$\widetilde{Y}_{n_x,1}$	$\cdots$	$\widetilde{Y}_{n_x,p}$	moderate
	Covariates			Class
$\widetilde{\mathscr{D}}(Z) =$	$\widetilde{Z}_{1,1}$	$\cdots$	$\widetilde{Z}_{1,p}$	high
		$\ddots$	$\vdots$	$\vdots$
	$\widetilde{Z}_{n_x,1}$	$\cdots$	$\widetilde{Z}_{n_x,p}$	high

5. Combine row-wise the sub-tables $\mathscr{D}(X)$, $\widetilde{\mathscr{D}}(Y)$ and $\widetilde{\mathscr{D}}(Z)$ to obtain the class-balanced data table $\widetilde{\mathscr{D}}$

	Covariates			Classes
$\widetilde{\mathscr{D}} =$	$X_{1,\,1}$	$\cdots$	$X_{1,\,p}$	null
		$\ddots$	$\vdots$	$\vdots$
	$X_{n_x,\,1}$	$\cdots$	$X_{n_x,\,p}$	null
	$\widetilde{Y}_{1,1}$	$\cdots$	$\widetilde{Y}_{1,p}$	moderate
		$\ddots$	$\vdots$	$\vdots$
	$\widetilde{Y}_{n_x,1}$	$\cdots$	$\widetilde{Y}_{n_x,p}$	moderate
	$\widetilde{Z}_{1,1}$	$\cdots$	$\widetilde{Z}_{1,p}$	high
		$\ddots$	$\vdots$	$\vdots$
	$\widetilde{Z}_{n_x,1}$	$\cdots$	$\widetilde{Z}_{n_x,p}$	high

Appendix B. Full result tables

This appendix reports overall misclassification error rates for all classifications performed in this study. Tables B.1 is related to the unbalanced case while Table B.2 is for the balanced case.

Table B.1: Overall misclassification error rates (%) for all tested statistical learning methods applied to the original unbalanced datasets from the 23 massifs in the French Alps. The leave-one-out score or assessing prediction performance is used for all methods: the K-nearest neighbors (K-NN) method ($K = 1$), the linear discriminant analysis (LDA), the support vector machine with linear kernel (SVM_LK), the support vector machine with polynomial kernel (SVM_PK), the support vector machine with radial kernel (SVM_RK), the classification trees (Trees) method, the random forests (RF) method and the weighted random forest method (wRF).

Massif names	LDA	1-NN	Trees	RF	wRF	SVM_LK	SVM_PK	SVM_RK
Aravis	61.9	62.4	61.8	66.6	61.8	61.8	61.8	61.8
Bauges	65.2	65.5	65.1	66.7	65.1	65.1	65.1	65.1
Beaufortain	65.5	65.9	65.5	66.7	65.5	65.5	65.4	65.5
Belledonne	65.7	66.1	65.6	66.7	65.6	65.6	65.5	65.6
Chablais	61.9	62.3	61.8	66.6	61.8	61.8	61.8	61.8
Champsaur	65.0	65.4	65.0	66.7	65.0	65.0	64.9	65.0
Chartreuse	66.6	66.9	66.5	66.7	66.5	66.5	66.4	66.5
Devoluy	66.4	66.6	66.2	66.7	66.2	66.2	66.1	66.1
Embrunais-Parpaillon	66.4	66.8	66.3	66.7	66.3	66.3	66.2	66.3
Grandes-Rousses	65.3	65.6	65.1	66.7	65.1	65.1	65.1	65.1
Haute-Maurienne	64.4	64.8	64.3	66.7	64.3	64.3	64.3	64.3

(continued)

Massif names	LDA	1-NN	Trees	RF	wRF	SVM_LK	SVM_PK	SVM_RK
Haute-Tarentaise	64.8	65.2	64.8	66.6	64.8	64.8	64.7	64.8
Haut-Var - Haut-Verdon	66.1	66.6	66.1	66.7	66.1	66.1	66.0	66.1
Maurienne	62.3	62.6	62.1	66.6	62.1	62.1	62.0	62.1
Mercantour	65.1	65.7	65.1	66.7	65.1	65.1	65.0	65.1
Mont-Blanc	64.9	65.3	64.9	66.7	64.9	64.9	64.8	64.9
Oisans	64.0	64.4	63.9	66.5	63.9	63.9	63.8	63.9
Pelvoux	65.8	66.2	65.8	66.7	65.8	65.8	65.7	65.8
Queyras	65.7	66.1	65.6	66.7	65.6	65.6	65.6	65.6
Thabor	66.3	66.7	66.3	66.7	66.3	66.3	66.2	66.3
Ubaye	65.2	65.6	65.1	66.7	65.1	65.1	65.0	65.1
Vanoise	64.2	64.5	64.1	66.6	64.1	64.1	64.0	64.1
Vercors	66.8	67.1	66.7	66.7	66.7	66.7	66.6	66.7

Table B.2: Overall misclassification error rates (%) for all tested statistical learning methods applied to the balanced datasets from the 23 massifs in the French Alps. The leave-one-out scheme for assessing prediction performance is used for all methods: the K-nearest neighbors (K-NN) method ($K = 1$), the linear discriminant analysis (LDA), the support vector machine with linear kernel (SVM_LK), the support vector machine with polynomial kernel (SVM_PK), the support vector machine with radial kernel (SVM_RK), the classification trees (Trees) method and the random forests (RF) method.

Massif names	LDA	1-NN	Trees	RF	SVM_LK	SVM_PK	SVM_RK
Aravis	30.7	3.6	24.2	1.3	28.6	12.4	11.3
Bauges	43.4	1.7	20.2	0.3	42.0	15.2	12.9
Beaufortain	36.2	1.5	25.4	0.0	27.4	8.0	7.6
Belledonne	37.2	1.2	26.4	0.0	29.4	6.7	6.0
Chablais	29.3	3.5	23.8	1.2	27.8	12.9	12.2
Champsaur	36.1	1.6	23.3	0.1	33.6	7.6	6.9
Chartreuse	39.4	0.6	22.6	0.0	38.2	11.9	9.1
Devoluy	31.7	0.8	21.2	0.0	24.2	4.9	4.4
Embrunais-Parpaillon	34.7	0.7	25.2	0.0	28.2	3.8	2.9
Grandes-Rousses	24.0	1.5	23.6	0.1	19.7	7.1	6.8
Haute-Maurienne	25.7	1.8	22.5	0.3	25.1	8.1	7.9
Haute-Tarentaise	27.8	1.8	24.4	0.3	22.2	6.9	6.5
Haut-Var - Haut-Verdon	34.5	0.9	21.5	0.0	29.2	6.8	6.1
Maurienne	27.7	3.1	21.0	1.1	22.6	11.1	10.6
Mercantour	25.4	1.5	20.8	0.2	19.4	7.3	7.0
Mont-Blanc	28.4	1.7	23.0	0.2	26.3	8.1	7.3
Oisans	26.9	2.1	22.6	0.5	19.4	7.9	7.0
Pelvoux	35.9	1.2	25.3	0.0	29.7	7.2	6.5
Queyras	35.5	1.2	27.5	0.0	29.6	8.8	7.6
Thabor	36.1	0.8	21.2	0.0	24.8	5.6	4.2
Ubaye	20.4	1.5	20.5	0.1	17.3	7.0	7.0
Vanoise	29.9	2.2	24.7	0.3	21.2	9.7	9.1
Vercors	25.9	0.5	14.7	0.0	18.3	3.0	1.8

References

Audigier, V., Husson, F., Josse, J., 2016. A principal component method to impute missing values for mixed data. Adv. Data Anal. Classif. 10 (1), 5–26.

Baggi, S., Schweizer, J., 2009. Characteristics of wet-snow avalanche activity: 20 years of observations from a high alpine valley (Dischma, Switzerland). Nat. Hazards 50 (1), 97–108.

Bakermans, L., Jamieson, B., Schweizer, J., Haegeli, P., 2010. Using stability tests and regional avalanche danger to estimate the local avalanche danger. Ann. Glaciol. 51 (54), 176–186.

Birkeland, K., Mocks, C., Shinker, J., 2001. Avalanche extremes and atmospheric circulation patterns. Ann. Glaciol. 32, 135–140.

Bois, P., Obled, C., Good, W., 1974. Multivariate data analysis as a tool for day-by-day avalanche forecast. In: Proceedings of the Snow Mechanics Symposium, vol. 114. IAHS Pub, pp. 391–403.

Bourova, E., Maldonado, E., Leroy, J.-B., Alouani, R., Eckert, N., Bonnefoy-Demongeot, M., Deschatres, M., 2016. A new web-based system to improve the monitoring of snow avalanche hazard in France. Nat. Hazards Earth Syst. Sci. 16 (5), 1205–1216.

Bovis, M.J., 1977. Statistical forecasting of Snow Avalanches, San Juan Mountains, Southern Colorado, U.S.a. J. Glaciol. 18, 87–99.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brun, E., David, P., Sudul, M., Brunot, G., 1992. A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting. J. Glaciol. 38, 13–22.

Bühler, Y., Hafner, E.D., Zweifel, B., Zesiger, M., Heisig, H., 2019. Where are the avalanches? Rapid spot6 satellite data acquisition to map an extreme avalanche period over the swiss alps. Cryosphere 13 (12), 3225–3238.

Buser, O., 1983. Avalanche forecast with the method of nearest neighbours: an interactive approach. Cold Reg. Sci. Technol. 8 (2), 155–163.

Buser, O., 1989. Two years experience of operational avalanche forecasting using nearest neighbors method. Ann. Glaciol. 13, 31–34.

Castebrunet, H., Eckert, N., Giraud, G., 2012. Snow and weather climatic control on snow avalanche occurrence fluctuations over 50 yr in the French Alps. Clim. Past 8 (2), 855–875.

Castebrunet, H., Eckert, N., Giraud, G., Durand, Y., Morin, S., 2014. Projected changes of snow conditions and avalanche activity in a warming climate: the French Alps over the 2020-2050 and 2070-2100 periods. Cryosphere 8 (5), 1673–1697.

Chen, C., Liaw, A., Breiman, L., 2004. Using Random Forest to Learn Imbalanced Data. Statistics Technical Reports 666. Department of Statistics, UC Berkeley.

Choubin, B., Borji, M., Mosavi, A., Sajedi-Hosseini, F., Singh, V.P., Shamshirband, S., 2019. Snow avalanche hazard prediction using machine learning methods. J. Hydrol. 577, 123929.

Conway, H., Raymond, C.F., 1993. Snow stability during rain. J. Glaciol. 39 (133), 635–642.

Conway, H., Wilbour, C., 1999. Evolution of snow slope stability during storms. Cold Reg. Sci. Technol. 30 (1–3), 67–77.

Davis, R., Elder, K., Howlett, D., 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 Reg. Sci. Technol. 30 (1–3), 79–89.

Dkengne Sielenou, P., Eckert, N., Naveau, P., 2016. A limiting distribution for maxima of discrete stationary triangular arrays with an application to risk due to avalanches. Extremes 19 (1), 25–40.

Dreier, L., Harvey, S., van Herwijnen, A., Mitterer, C., 2016. Relating meteorological parameters to glide-snow avalanche activity. Cold Reg. Sci. Technol. 128, 57–68.

Durand, Y., Brun, E., Mérindol, L., Guyomarch'h, G., Lesaffre, B., Martin, E., 1993. A meteorological estimation of relevant parameters for snow models. Ann. Glaciol. 18, 65–71.

Durand, Y., Giraud, G., Brun, E., Mérindol, L., Martin, E., 1999. A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting. J. Glaciol. 45, 469–484.

Durand, Y., Giraud, G., Laternser, M., Etchevers, P., Mérindol, L., Lesaffre, B., 2009a. Reanalysis of 47 years of climate in the French Alps (1958-2005): climatology and trends for snow cover. J. Appl. Meteorol. Climatol. 48 (12), 2487–2512.

Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., Mérindol, L., 2009b. Reanalysis of 44 Yr of climate in the French Alps (1958-2002): methodology, model validation, climatology, and trends for air temperature and precipitation. J. Appl. Meteorol. Climatol. 48 (3), 429–449.

Eckerstorfer, M., Malnes, E., 2015. Manual detection of snow avalanche debris using high-resolution radarsat-2 SAR images. Cold Reg. Sci. Technol. 120, 205–218.

Eckert, N., Baya, H., Deschâtres, M., 2010a. Assessing the response of snow avalanche runout altitudes to climate fluctuations using hierarchical modelling: application to 61 winters of data in France. J. Clim. 23, 3157–3180.

Eckert, N., Coleou, C., Castebrunet, H., Deschâtres, M., Giraud, G., Gaume, J., 2010b. Cross-comparison of meteorological and avalanche data for characterizing avalanche cycles: the example of December 2008 in the eastern part of the French Alps. Cold Reg. Sci. Technol. 64, 119–136.

Eckert, N., Parent, E., Kies, R., Baya, H., 2010c. A spatio-temporal modelling framework for assessing the fluctuations of avalanche occurrence resulting from climate change: application to 60 years of data in the northern French Alps. Clim. Chang. 101, 515–553.

Eckert, N., Keylock, C.J., Castebrunet, H., Lavigne, A., Naaim, M., 2013. Temporal trends in avalanche activity in the French Alps and subregions: from occurrences and runout altitudes to unsteady return periods. J. Glaciol. 59, 93–114.

Fierz, C., Armstrong, R.L., Durand, Y., Etchevers, P., Greene, E., McClung, D.M., Nishimura, K., Satyawali, P.K., Sokratov, S.A., 2009. The international classification of seasonal snow on the ground. IHP-VII Technical Documents in Hydrology 83, IACS Contribution No. 1. UNESCO-IHP, Paris.

Floyer, J., McClung, D., 2003. Numerical avalanche prediction: Bear pass, British Columbia, Canada. Cold Reg. Sci. Technol. 37 (3), 333–342.

Föhn, P., Good, W., Bois, P., Obled, C., 1977. Evaluation and comparison of statistical and conventional methods of forecasting avalanche hazard. J. Glaciol. 19, 375–387.

Gassner, M., Brabec, B., 2002. Nearest neighbour models for local and regional avalanche forecasting. Nat. Hazards Earth Syst. Sci. 2, 247–253.

Gaume, J., Eckert, N., Chambon, G., Naaim, M., Bel, L., 2013. Mapping extreme snowfalls in the French Alps using max-stable processes. Water Resour. Res. 49 (2), 1079–1098.

Gaume, J., Schweizer, J., van Herwijnen, A., Chambon, G., Reuter, B., Eckert, N., Naaim, M., 2014. Evaluation of slope stability with respect to snowpack spatial variability. J. Geophys. Res. Earth Surf. 119 (9), 1783–1799.

Genuer, R., Jean-Michel Poggi, J.-M., Tuleau-Malot, C., 2010. Variable selection using random forests. Pattern Recogn. Lett. 31 (14), 2225–2236.

Giraud, G., Brun, E., Durand, Y., Martin, E., 1998. Safran/crocus/mepra models as an helping tool for avalanche forecasters. Publik. - Norges Geotek. Inst. 203, 108–112.

Gregorutti, B., Michel, B., Saint-Pierre, P., 2016. Correlation and variable importance in random forests. Stat. Comput. 1–20.

Heierli, J., Purves, R., Felber, A., Kowalski, J., 2004. Verification of nearest neighbours interpretations in avalanche forecasting. Ann. Glaciol. 38, 84–88.

Hendrikx, J., Owens, I., Carran, W., Carran, A., 2005. Avalanche activity in an extreme maritime climate: the application of classification trees for forecasting. Cold Reg. Sci. Technol. 43 (1–2), 104–116.

Hendrikx, J., Murphy, M., Onslow, T., 2014. Classification trees as a tool for operational avalanche forecasting on the Seward highway, Alaska. Cold Reg. Sci. Technol. 97, 113–120.

Husson, F., Josse, J., Pages, J., 2010a. Principal Component Methods-Hierarchical Clustering-Partitional Clustering: Why Would we Need to Choose for Visualizing Data. Applied Mathematics Department.

Husson, F., Le, S., Pages, J., 2010b. Exploratory Multivariate Analysis by Example Using R. Chapman and Hall/CRC Computer Science and Data Analysis.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning with Applications in R. Springer-Verlag New York.

Jamieson, B., Geldsetzer, T., Stethem, C., 2001. Forecasting for deep slab avalanches. Cold Reg. Sci. Technol. 33 (2–3), 275–290.

Jamieson, B., Haegeli, P., Schweizer, J., 2009. Field observations for estimating the local avalanche danger in the Columbia mountains of Canada. Cold Reg. Sci. Technol. 58 (1–2), 84–91.

Jomelli, V., Delval, C., Grancher, D., Escande, S., Brunstein, D., Hetu, B., Filion, L., Pech, P., 2007. Probabilistic analysis of recent snow avalanche activity and climate in the French Alps. Cold Reg. Sci. Technol. 47, 180–192.

Josse, J., Husson, F., 2016. missMDA: a package for handling missing values in multivariate data analysis. J. Stat. Softw. 70 (1), 1–31.

Judson, A., King, R., 1985. An index of regional snow-pack stability based on natural slab avalanches. J. Glaciol. 31, 67–73.

Kaufman, L., Rousseeuw, P.J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley and Sons.

LaChapelle, E.R., 1980. The fundamental processes in conventional avalanche forecasting. J. Glaciol. 26, 75–84.

Lafaysse, M., Cluzet, B., Dumont, M., Lejeune, Y., Vionnet, V., Morin, S., 2017. A multiphysical ensemble system of numerical snow modelling. Cryosphere 11 (3), 1173–1198.

Lavigne, A., Bel, L., Parent, E., Eckert, N., 2012. A model for spatio-temporal clustering using multinomial probit regression: application to avalanche counts. Environmetrics 23 (6), 522–534.

Lavigne, A., Eckert, N., Bel, L., Parent, E., 2015. Adding expert contributions to the spatiotemporal modelling of avalanche activity under different climatic influences. J. Royal Statist. Soc. Ser. C: Appl. Statist. 64 (4), 651–671.

Liaw, A., Wiener, M., 2002. Classification and regression by randomforest. R News 2 (3), 18–22.

Marienthal, A., Hendrikx, J., Birkeland, K., Irvine, K., 2015. Meteorological variables to aid forecasting deep slab avalanches on persistent weak layers. Cold Reg. Sci. Technol. 120, 227–236.

McClung, D., 2002. The elements of applied avalanche forecasting part ii: the physical issues and the rules of applied avalanche forecasting. Nat. Hazards 26 (2), 131–146.

McClung, D., Schaerer, P., 1981. Snow avalanche size classification. Nat. Res. Council Can. Tech. Memoran. 133, 12–30.

McClung, D., Schaerer, P., 1993. The Avalanche Handbook. The Mountaineers, Seattle, Washington, p. 271.

McClung, D., Tweedy, J., 1994. Numerical avalanche prediction: Kootenay Pass, British Columbia, Canada. J. Glaciol. 40 (135), 350–358.

McCollister, C., Birkeland, K., Hansen, K., Aspinall, R., Comey, R., 2003. Exploring multi-scale spatial patterns in historical avalanche data, Jackson Hole Mountain Resort, Wyoming. Cold Reg. Sci. Technol. 37 (3), 299–313.

McGregor, G., 1989. Snow avalanche forecasting by discriminant function analysis. Weather Clim. 9, 3–14.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2014. e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R Package Version 1.6-4.

Mitterer, C., Schweizer, J., 2013. Analysis of the snow-atmosphere energy balance during wet-snow instabilities and implications for avalanche prediction. Cryosphere 7 (1), 205–216.

Möhle, S., Bründl, M., Beierle, C., 2014. Modeling a system for decision support in snow avalanche warning using balanced random forest and weighted random forest. In: Proceedings of the 16th Artificial Intelligence: Methodology, Systems, and Applications (AIMSA 2014), pp. 80–91.

Morin, S., Horton, S., Techel, F., Bavay, M., Coléou, C., Fierz, C., Gobiet, A., Hagenmuller, P., Lafaysse, M., Lizar, M., Mitterer, C., Monti, F., Müller, K., Olefs, M., Snook, J.S., van Herwijnen, A., Vionnet, V., 2019. 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.

Mougin, P., 1922. Les avalanches en savoie. In: Technical report, Ministère de l'Agriculture, Direction Générale des Eaux et Forêts. Service des Grandes Forces Hydrauliques, Paris.

Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J. Classif. 31, 274–295.

Navarre, J.P., Guyomarc'h, G., Giraud, G., 1987. Un modèle statistique pour la prévision locale des avalanches. IAHS Publ. 162, 571–580.

Nguyen, G.H., Bouzerdoum, A., Phung, S., 2009. Learning pattern classification tasks with imbalanced data sets. Pattern Recogn. 193–208.

Obled, C., Good, W., 1980. Recent developments of avalanche forecasting by discriminant analysis techniques: a methodological review and some applications to the persenn area (Davos). J. Glaciol. 25 (92), 315–346.

Pagès, J., 2004. Analyse factorielle des données mixtes. Rev. Statist. Appl. 52 (4), 93–111.

Perla, R., 1970. On contributory factors in avalanche hazard evaluation. Can. Geotech. J. 7 (4), 414–419.

Pozdnoukhov, A., Matasci, G., Kanevski, M., Purves, R.S., 2008. Applying machine learning methods to avalanche forecasting. Ann. Glaciol. 49, 107–113.

Pozdnoukhov, A., Matasci, G., Kanevski, M., Purves, R.S., 2011. Spatio-temporal avalanche forecasting with support vector machines. Nat. Hazards Earth Syst. Sci. 11, 367–382.

Pudasaini, S., 2007. Avalanche Dynamics. Springer-Verlag, Berlin Heidelberg.

Purves, R.S., Morrison, K.W., Moss, G., Wright, D.S.B., 2003. Nearest neighbours for avalanche forecasting in Scotland - development, verification and optimisation of a model. Cold Reg. Sci. Technol. 37, 343–355.

de Quervain, M.R., 1981. Avalanche Atlas: Illustrated International Avalanche Classification. Unesco.

Ramyachitra, D., Manikandan, P., 2014. Imbalanced dataset classification and solutions: a review. Intern. J. Comput. Bus. Res. 5 (4) (ISSN: 2229-6166).

Reuter, B., Richter, B., Schweizer, J., 2016. Snow instability patterns at the scale of a small basin. J. Geophys. Res. F: Earth Surf. 121 (2), 257–282.

Ripley, B., 2014. tree: Classification and Regression Trees. R Package Version 1.0-35.

Salford-Systems, 2014. Random Forests for Beginners. ©Salford Systems.

Schirmer, M., Lehning, M., Schweizer, J., 2009. Statistical forecasting of regional avalanche danger using simulated snow-cover data. J. Glaciol. 55 (193), 761–768.

Schirmer, M., Schweizer, J., Lehning, M., 2010. Statistical evaluation of local to regional snowpack stability using simulated snow-cover data. Cold Reg. Sci. Technol. 64 (2), 110–118.

Schweizer, J., Föhn, P., 1996. Avalanche forecasting - an expert system approach. J. Glaciol. 42 (141), 318–332.

Schweizer, J., Jamieson, J., Schneebeli, M., 2003a. Snow avalanche formation. Rev. Geophys. 41 (4), 2–1 – 2–25.

Schweizer, J., Kronholm, K., Wiesinger, T., 2003b. Verification of regional snowpack stability and avalanche danger. Cold Reg. Sci. Technol. 37 (3), 277–288.

Schweizer, J., Kronholm, K., Jamieson, J., Birkeland, K., 2008. Review of spatial variability of snowpack properties and its importance for avalanche formation. Cold Reg. Sci. Technol. 51 (2–3), 253–272.

Schweizer, J., Mitterer, C., Stoffel, L., 2009. On forecasting large and infrequent snow avalanches. Cold Reg. Sci. Technol. 59, 234–241.

Schweizer, J., Mitterer, C., Techel, F., Stoffel, A., Reuter, B., 2020. On the relation between avalanche occurrence and avalanche danger level. Cryosphere 14 (2), 737–750.

Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. J. Royal Statist. Soc. Ser. B (Methodol.) 36 (2), 111–147.

Uppala, S., Kallberg, P.W., Hernandez, A., Saarinen, S., Fiorino, M., Li, X., Onogi, K., Sokka, N., Andrae, U., Da Costa Bechtold, V., 2004. Era-40: Ecmwf 45-year reanalysis of the Global Atmosphere and Surface conditions 1957-2002. Newsletter, p. 101 (ECMWF).

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, Fourth ed. Springer, New York. ISBN 0-387-95457-0.

Vernay, M., Lafaysse, M., Merindol, L., Giraud, G., Morin, S., 2015. Ensemble forecasting of snowpack conditions and avalanche hazard. Cold Reg. Sci. Technol. 120, 251–262.

Vernay, M., Lafaysse, M., Hagenmuller, P., Nheili, R., Verfaillie, D., Morin, S., 2019. The s2m Meteorological and Snow Cover Reanalysis in the French Mountainous Areas (1958–Present).

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E., Willemet, J.-M., 2012. The detailed snowpack scheme crocus and its implementation in surfex v7.2. Geosci. Model Dev. 5 (3), 773–791.

Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244.