What weather variables are important for wet and slab avalanches under a changing climate in low altitude mountain range in Czechia?

Markéta Součková[1,2], Roman Juras[1], Kryštof Dytrt[1], Vojtěch Moravec[1,2], Johanna Ruth Blöcher[1], and Martin Hanel[1]

[1]Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha – Suchdol, 165 00, Czechia
[2] T.G. Masaryk Water Research Institute, Department of Hydrology, Podbabská 2582/30, 160 00 Prague 6, Czechia

Correspondence: Markéta Součková (souckovamarketa@fzp.czu.cz)

Abstract.

Climate change impact on avalanches is ambiguous. Fewer, wetter, and smaller avalanches are expected in areas where snow cover is declining, while in higher altitude areas where snowfall prevails, snow avalanches are frequently and spontaneously triggered. In the present paper, we assess 39 years (winters of 1979–1999 to 2002–2020) of avalanche activity related to meteorological and snow drivers in the Krkonoše Mountains, Czechia, Central Europe. The analysis is based on an avalanche occurrence dataset for mostly south, south-easterly oriented 60 avalanche paths and related meteorological and snowpack data. Since 1979, 179 / 531 wet-snow / slab avalanches have been recorded. The aim is to analyze changes in avalanche activity: frequency and magnitude, and detect driving weather variables of wet and slab avalanches with quantification of variable importance. Especially, the number of wet avalanches in February and March has increased in the last three decades, while the number of slab avalanches has decreased with decadal variability. Medium, large, and very large slab avalanches seem to decline with decadal variability since 1961. The results indicate that wet avalanches are influenced by 3-day maximum and minimum air temperature, snow depth, wind speed, wind direction, and rainfall. Slab avalanche activity is determined by snow depth, rainfall, new snow, and wind speed. Air temperature-related variables for slab avalanches were less important than rain and snow-related variables based on the balanced random forest (RF) method. Surprisingly, the RF analysis revealed less significant relationship between new snow sum and slab avalanche activity. This could be because of the wind redistributing snow in storms in low altitude mountains. Our analysis allows the use of the identified wet and slab avalanche driving variables to be included in the avalanche danger levels alerts. Although it cannot replace operational forecasting, machine learning can allow for additional insights for the decision-making process to mitigate avalanche hazard.

1 Introduction

Snow avalanches are major natural hazards. As rapidly moving snow masses, snow avalanches pose a serious threat to people, property, and infrastructure. The growth in popularity of winter tourism has led to an increase in numbers of avalanche acci-

dents (Techel et al., 2016). Although climate change is happening, its effect on snow avalanches remain unclear (Strapazzon et al., 2021). The frequency and types of snow avalanches may change (Hock et al., 2019). A wetter, warmer climate could exacerbate the consequences of burial (Strapazzon et al., 2021). This is why it is vital to analyse how a changing climate impacts

25 avalanches.

Meteorological conditions, snow stratigraphy, land cover, and land use all determine avalanche activity and risk. Changes in land use and land cover, such as deforestation (García-Hernández et al., 2017), as well as demographics (Giacona et al., 2018) caused by climate change will affect avalanche risk. Avalanche types and avalanche frequency are affected by a combination of precipitation amounts and air temperatures during storms, and prior snow stratigraphy (Schweizer et al., 2009). Winter

30 recreationists face a significant threat due to the presence of persistent weak layers (Statham et al., 2018; Techel et al., 2015). Despite a likely decrease in overall avalanche frequency (Strapazzon et al., 2021), more extreme precipitation events during winter storms can cause more intensive avalanche activity. Avalanche activity is governed by both variable and permanent factors (Quervain et al., 1973). Whereas, variable factors are attributed to meteorological conditions (for instance rain, air temperature, wind, snowfall) that progressively build the snowpack, permanent factors are attributed to terrain features (elevation,

35 slope, aspect, roughness of the ground, etc.) (Dkengne Sielenou et al., 2021).

The overall number and runout distance of snow avalanches will reduce in regions and elevations experiencing a significant reduction in snow cover. There is medium evidence and high agreement that observed changes in avalanches in mountain regions will be exacerbated in the future (Hock et al., 2019), with generally a decrease in hazard (Martin et al., 2001) at lower elevations. However less snow does not necessarily result in fewer avalanches (Ballesteros-Cánovas et al., 2018; Peitzsch

40 et al., 2021; Reuter et al., 2020). At higher elevations, mixed changes are expected, with more wet snow avalanches. Wet snow avalanches will occur more frequently even in winter, which has already been shown for recent decades from December to February (Naaim et al., 2016). Since wet-snow avalanches are more likely to occur early in the season, spring avalanche activity at lower elevations is likely to decrease, while avalanche activity at higher elevations is likely to increase (Castebrunet et al., 2014; Strapazzon et al., 2021). There is no clear direction in the trend for overall avalanche activity (Hock et al., 2019).

45 Ballesteros-Cánovas et al. (2018) reported increased frequency of wet avalanche activity on some slopes of the Western Indian Himalaya over recent decades. In the European Alps and Tatra mountains, avalanche mass and run-out distance have decreased at lower elevation as well as powder avalanches. Avalanche numbers have decreased below 2000 m a.s.l., and increased above this elevation (Eckert et al., 2013; Lavigne et al., 2015; Gądek et al., 2017).

The snow avalanche releases on potentially sliding layers which are weaknesses 1) in new snow stratigraphy, 2) on old snow

50 surface, and 3) in temperature gradient layers (McGregor, 1990). Focusing on wet-snow avalanches, three triggering mechanisms exists (or their combinations) due to loss of strength, overloading and gradual weakening (Baggi and Schweizer, 2009). More specifically, loss of strength can be caused by infiltration and accumulation of water at capillary barriers. Overloading can occur due to precipitation of partially wet and weakened snowpack. Lastly, the gradual weakening of the basal snowpack can occur as the snowpack becomes isothermal, causing failure of the basal layers. This can be caused by heat stored in the

55 ground that melts the lowermost snow layer.

Slab avalanches are wet or dry. Natural slab avalanches are triggered either due to gradual uniform loading by precipitation and wind (or by a combination of both), or due to a non-loading situation that changes the snowpack properties, such as surface warming (Schweizer et al., 2003).

Climate change influences mountain snow cover by increase in air temperature and rainfall during the winter. Depending on elevation, air temperature increases may cause changes in the type, intensity, and frequency of snowfall (Strapazzon et al., 2021). At higher elevations, air temperatures will rise and rain will occur more often. At lower elevations, snowfall is less frequent and intense, resulting in a thinner, wetter snowpack with a higher average density according to Intergovernmental Panel on Climate change (IPCC) special report Hock et al. (2019); Strapazzon et al. (2021). When snow cover decreases, avalanche hazard areas also decrease (Strapazzon et al., 2021). At moderate and high elevations, the likelihood of more dynamic changes in temperature and precipitation is higher, with accelerated fluctuations between extremes, and with less prominent trends because of local effects (Hock et al., 2019). The avalanche regime may be less impacted at higher elevations, where snowfall is still abundant and may increase in intensity (Hock et al., 2019; Laute and Beylich, 2018; Strapazzon et al., 2021). More avalanches occur during snowstorms. During major snowstorms, rain intensity is predicted to increase gently (Strapazzon et al., 2021). The observed shift from solid to liquid precipitation is likely to move the position of seasonal snow lines to higher elevations, and shorten snow seasons (Beniston et al., 2018; Marty et al., 2017; Strapazzon et al., 2021). Globally, snowfall has reduced as a result of increasing temperatures, especially at lower elevations (Hock et al., 2019). Regional trends of increasing liquid precipitation during winter were confirmed (Bintanja, 2018; Feng and Hu, 2007). Moreover, a decrease in snowfall fraction (Sf) of $-5.5\,\% \times \mathrm{decade}^{-1}$ in low altitude mountain catchments (> 900 m) in Czechia and Slovakia has been observed since 1966 (Blahušiaková et al., 2020). Over the last 35 years, the precipitation phase in the cold season has partially shifted from solid to mixed precipitation, with the most substantial decrease in snowfall in February ($-10.5\,\% \times \mathrm{decade}^{-1}$) and January ($-6.3\,\% \times \mathrm{decade}^{-1}$) (Hynčica and Huth, 2019).

As an avalanche is a rapid snow mass, therefore the trends of snowpack properties such as snow depth (SD), snow water equivalent (SWE), and snow characteristics (e.g. snow cover extend (SCE), and snow cover duration (SCD)) serve as the main proxies for the detection of future snow avalanche activity in many regions. Changes projected for the mountain cryosphere refer to a decrease in SCE and SCD (Notarnicola, 2020). SWE and SD have declined in nearly all regions on average by $-5\,\% \times \mathrm{decade}^{-1}$. This trend is apparent especially at lower elevations, although year-to-year variation is high (Hock et al., 2019). Overall, snow cover duration has shortened in Czech mountain catchments over recent decades by up to -6.8 days$\times \mathrm{decade}^{-1}$, principally due to earlier melt out (Blahušiaková et al., 2020). Snowpack at elevations below 1200 m a.s.l. was more sensitive to changes in air temperature, while precipitation influenced the snowpack more at elevations above 1200 m a.s.l. (Nedelcev and Jeníček, 2021). In central Europe, Blahušiaková et al. (2020) reported that snow depth has been declining in recent years. Among higher elevations ($\sim$2300 m a.s.l.), snow depth has decreased by 1 %, while at lower elevations ($\sim$800 m a.s.l.) has decreased by 6.3 % since 1966 (Blahušiaková et al., 2020). According to climatological models, the snow season will be shortened by 25 days in Czechia during the period 2021 to 2040; simulations can be seen at Klimatická Změna.cz.

Avalanche analyses related to meteorological and snow parameters of a) wet snow avalanche were investigated by Baggi and Schweizer (2009); Bel; Peitzsch et al. (2012); b) slab avalanches were investigated by Bellaire et al. (2017); Eckerstorfer

and Christiansen (2011); Marienthal et al. (2015). Different approaches have been used: classification trees (e.g. Hendrikx et al., 2014; Marienthal et al., 2015; Dreier et al., 2016), logistic regression (e.g. Gauthier et al., 2017; Dreier et al., 2016), and random forest (e.g. Möhle et al., 2014; Dkengne Sielenou et al., 2021). The separate climatic, snow, and avalanche releases in Krkonoše were described in Spusta et al. (2020) during winters 2006/07–2018/19; however, the relationship between the

95   individual daily meteorological snow conditions related to avalanches has not yet been done in Czechia. Krkonoše was one of the first non-Alpine regions that established regular snow monitoring and avalanche records in 1961 (Blahůt et al., 2017; Vrba and Spusta, 1975, 1991; Juras et al.; Spusta and Kociánová, 1998; Spusta et al., 2003, 2006). As the effects of climate change on avalanche types is ambiguous and regional differences are vast, there is a need for understanding how avalanche types change over multi-decadal time scales, as well as gaining knowledge of triggering variables of wet and slab avalanche

100   activity in a low altitude mountain range (Krkonoše mountains, NE border of Czechia and Poland). Although snow avalanches do not present a significant risk to the population in Czechia, the rising popularity of winter sports (off piste skiing and ski touring) in recent years has led to an increase in social exposure to snow avalanches and thus a growing number of victims (11 fatalities, 15 injured, and 28 people pulled down since 2005) (Dlouhý) and, rarely, road accidents. Furthermore, a decision support model helping avalanche forecasting is missing.

105   In this study, firstly, we hypothesize that there will be more wet avalanches because of rising air temperature and increasing rainfall in Czechia. Sunlight duration determines the warming of snowpack, resulting in snow-wet avalanche releases. Secondly, we suggest slab avalanches are mainly driven by snow depth, new snow sum, and wind speed redistributing snow. Therefore in this study, we aim to 1) analyze changes in avalanche activity: assessment of frequency and magnitude over 60 winter seasons (1961–2021), and 2) determine the main meteorological and snow drivers governing snow avalanche activity of a)

110   wet avalanches and b) slab avalanches for a daily time scale (1979–1999, 2002–2020) within the altitude mountain ranges in Central Europe, specifically the Krkonoše mountains.

The paper is organized as follows: Section 2 describes the data and methods used in this study. In Section 3, we present the results. Trends in two avalanche types and the skill of models obtained by machine learning are described and discussed, and future work is pointed out in Sect. 4. In Section 5, we conclude the findings of our study.

115   ## 2   Data and methods

In the following section we describe the Krkonoše Mountains, a low altitude mountain range in Central Europe. We present its geology, geomorphology, land cover, and meteorological conditions. Subsequently, we report on the avalanche activity dataset, meteorological and snow data, and methods used for estimating the main weather and snow variables determining avalanche activity.

120   ### 2.1   Study area

The Krkonoše Mountains (internationally known as the Giant Mountains), with the highest peak Sněžka at 1602 m a.s.l., is the area with the most frequent snow avalanche activity in Czechia. The Krkonoše mountain range extends between Czechia and

Poland, with the larger part is located in north-east Czechia. Most of the mountain range belongs to Krkonoše National Park (KRNAP), which covers an area of 550 km^2 and has been protected since 1963.

125   As part of the Variscan / Hercynian mountain ranges in Europe, Krkonoše is mainly comprised of crystalline schists with several quartzites and crystalline limestones. The central part (border with Poland) is formed of granites, with Alpine orogeny and Quaternary glaciations that carved out several plateaus at an elevation between 1300 and 1450 m a.s.l. (Blahůt et al., 2017). The plateaus host several headwaters (e.g. Elbe River) and glacial cirques (Engel et al., 2010), where small brooks originate in the vicinity of several avalanche triggering zones and mainly affect avalanche activity mainly the snowmelt period. The

130 combination of geomorphology and land cover creates a natural environment for avalanche occurrence. The mean slope of avalanche release zones is 31 ° and mean elevation ranges between 1072 to 1575 m a.s.l.; the avalanche paths are mostly facing south, south-east (mean aspect 168 °) (Fig. 1).

     The biogeographical location of the Krkonoše Mountains is created by a varied mosaic of montane spruce and mixed forests, tall herb meadows, dwarf pine communities, Nardus grasslands, sub-arctic peat bogs and lichen tundra. Arcto-alpine tundra

135 covers 4 % of the territory. According to the KRNAP Green Infrastructure map the avalanche release zones consist mainly of alpine meadows (39.7 %), natural cypress (32.65 %) and rocks and scree (20.95 %) (Erlebach). A few, spruces, peat bogs, and springs are spread in avalanche release zones < 3 %. The tree line lies between 1200–1350 m a.s.l. (Štursa et al., 2010).

     Krkonoše are characterised by mean annual air temperature of 0 °C at its highest parts, prevailing westerly winds (Blahůt et al., 2017), and annual precipitation of about 1200 mm (Tolasz, 12007), of which about 34 % consists of snow (Juras et al.,

140 2021). The mean annual temperature has increased in Krkonoše over the last two decades by 1 °C in comparison with the period 1961–2000 (Kliegrová and Kašičková, 2019). The wind blows from west to east, resulting in relatively low snow accumulation on the west-facing, wind-ward slopes while steep, lee-ward slopes accumulates much more snow. The snowpack depths usually range between 100 and 300 cm (Blahůt et al., 2017).

2.2   Avalanche activity dataset and data manipulation

145 Avalanches, as rapidly moving snow masses with a minimum length of 50 metres, have been monitored on 60 avalanche paths in the Czech part of Krkonoše since 1961 (Spusta et al., 2020) (Fig. 1) by KRNAP administration, Krkonoše Mountain Rescue Service, and web camera records. During 60 winter seasons (from 1961 to 2021) 1246 avalanches were recorded in the Czech avalanche dataset. Snow avalanches are classified by international codes (Quervain et al., 1973), with a little modification for Krkonoše. The dimensions of each avalanche are listed in Vrba and Spusta (1975, 1991); Spusta and Kociánová (1998); Spusta

150 et al. (2003, 2020).

     In order to know the trend of wet and slab avalanche activity, we filtered two types of avalanches from the avalanche dataset according to criteria: zone of origin (known as release zone) a) manner of starting (A2 line release zone (271 avalanche records: Aval); A3 soft slab (514 Aval); A4 hard slab (45 Aval)), and b) liquid water in snow (C=2, 186 Aval) according to Quervain et al. (1973) Avalanche classification, to perform closer analysis. Wet avalanche dataset, defined by wet snow presence in a release

155 zone, was chosen as an indicator of changing climate, and slab avalanches were chosen as the most frequent and dangerous

Table 1. Avalanche size classification

Size category	Category number	Relative to max length
Very small	1	$< 20\%$
Small	2	$> 20\% < 40\%$
Medium	3	$> 40\% < 60\%$
Large	4	$> 60\% < 80\%$
Very large	5	$> 80\%$

avalanche type for skiers in the Krkonoše Mountains. The selection of two avalanche types is also based on avalanche danger models suggested by Mair and Nairz (2011).

Long term trends in frequency and size, as well as basic weather parameters were analysed for both selected groups. We processed avalanche (Aval) data characteristics for wet and slab snow avalanches: count and magnitude (avalanche size) of

160 avalanche length. Avalanche activity trends in the avalanche dataset were explored over decadal (1961–2021) time scales.

Avalanche size category was estimated based on American classification (Greene et al., 2010). Avalanche size was considered relative to the path; each avalanche length was related to the maximum avalanche length recorded at the specific path since 1961. The classification has five size categories graded in 20 % steps of the maximum length (Table 1).

2.3 Meteorological and snowpack data preparation

165 Daily data were provided by the Czech Hydrometeorological Institute (CHMI) which are now freely accessible. We used meteorological data from an automated weather station in the Krkonoše Mountains: Labská bouda (LBOU, 1320 m a.s.l.; Fig. 1). For the purpose of the study, we created two wet and slab datasets with the variables listed in (Table 2). Beside the measured values, we used two additional variables of rainfall. There are more methods to determine rainfall from total precipitation. First we used rainfall (Rain_Tw) based on single threshold wet bulb temperature calculated according to Stull

170 (2011) formula ($Tt = -0.5\,^\circ\text{C}$). Second, rainfall separated from the total precipitation (Rain_Ta) based on single threshold air temperature ($Tt = +0.46\,^\circ\text{C}$), which was calibrated for the Elbe catchment (Juras et al., 2021). Apart from station measured variables, we generated 3-day moving average windows of the input variables and sums of selected variables (Table 2). By using a moving average, the curve is smoothed and it helps to better identify a trend or trend change; sums highlight the effect. Even though snow water equivalent would be a promising predictor of avalanches, we excluded it from the dataset as

175 it is, unfortunately, measured only weekly in Czechia, and the interpolated data could be misleading. The winter season was considered as a period from 1 / 10 to 31 / 5 when snow can be observed at the study site.

For the purpose of machine learning analysis, the wet and slab datasets contain the Avalanche day (Ad) and Non avalanche day (NAd). Ad was defined within the winter season when at least one avalanche was recorded and NAd when no avalanche was recorded. We explored the occurrence of avalanches based on the course of hydroclimatic variables during the previous
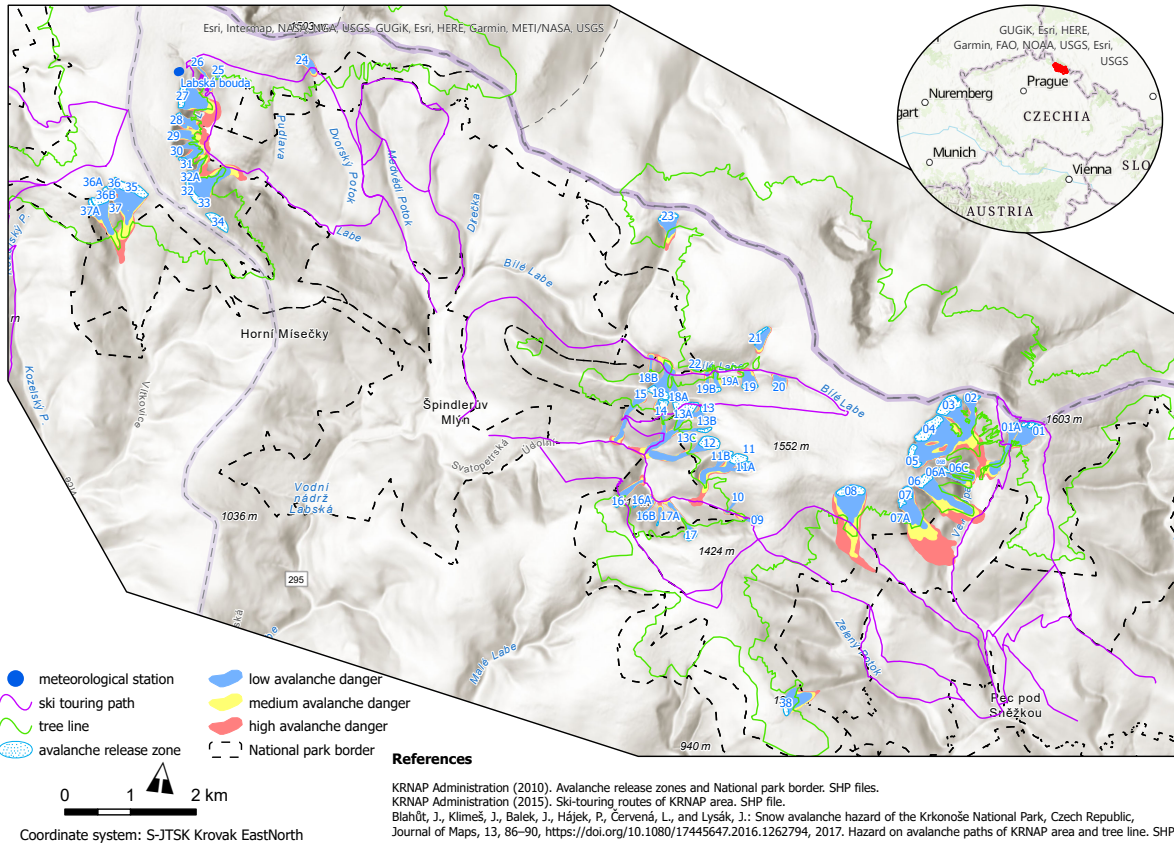
Fig. 1. Study location of avalanche activity in the Krkonoše mountains.

180　six days. NAds that occurred six days after an avalanche record were deleted to minimize the dependency on preconditions between Ad and NAd. Datasets should not contain any missing values as some machine learning algorithms do not deal with them correctly. Therefore, we excluded them using 39 year-long data time series (1979–2020 data period; 2001–2002 was omitted) as the station Labská bouda (LBOU) did not operate during this time window. In the end, we excluded the three-day moving average / sum of variables from (Table 2) as they were highly correlated, and for a linear approach like Principle

185　component analysis, duplications have to be removed. Correlated variables are displayed by dendrograms for wet in Fig. A1 and slab avalanches in Fig. A2. For nonlinear models, colinear variables can remain.

2.4 Balancing data, machine learning methods

We analyzed the explanatory power of several meteorological and snow variables to explain avalanche triggers for the daily time scale (1979–1999, 2002–2020). We applied tree-based models (Decision Tree, Random Forest) and Principle component

190　analysis (PCA) to a) determine the set of the most relevant combination of explanatory variables for the avalanche occur-

Table 2. Description of the weather variables used in this study.

Variable	Abbreviation	Model	Explanation
Snow depth [cm]	SD	SD_value	Daily mean snow depth
		SD_value3	3 day moving average of snow depth before avalanche release (day when Aval occured)
		SD_dif2,3,4	2,3,4 days snow dept difference from the day when avalanche released; SDdif = SD_value - SD2,3,4
New snow sum [cm]	NSS	NSS_value	New snow fallen in a day
		NSS_value3	3 day moving average of new snow before avalanche release (day when Aval occured)
		NSSsum3	Sum of 3 day new snow before avalanche release
Relative humidity [%]	H	H_value	Daily relative humidity
		H_value3	3 day moving average of relative humidity before avalanche release (day when Aval occured)
Air temperature [$^\circ C$]	Tair	Tair_value	Daily mean air temperature [$^\circ C$]
		Tair_value3	3 day moving average of air temperature before avalanche release (day when Aval occured)
		Tmin3	3 days minimum air temperature [$^\circ C$]
		Tmax3	3 days maximum air temperature [$^\circ C$]
		Tamp3	Thermal amplitude the day before and up to 3 days before the avalanche release [$^\circ C$]; Tamp3 = Tmax3-Tmin3
Sunlight duration [$hour$]	SLd	SLd_value	Daily sun light duration
		SLd_value3	3 day moving sum of sunlight duration before avalanche release (day when Aval occured)
Precipitation [mm]	P	P_value	Daily total precipitation
		P_value3	3 day moving sum of daily precipitation
Rainfall [mm]	Rain_Tw	Rain_Tw_value	Daily rainfall separated from total precipitation (P) based on single threshold (Tt = -0.5°C) wet bulb temperature (calculated according to Stull (2011) formula).
		Rain_Tw_value3	3 days moving average of Rain_Tw_value
Rainfall [mm]	Rain_Ta	Rain_Ta_value	Daily rainfall, separated from the total precipitation (P) based on single threshold air temperature (Tt = +0.46 °C) Tt was calibrated by (Juras et al., 2021).
		Rain_Ta_value3	3 days moving average of Rain_Ta_value
		Rain_Ta_sum3	3 days sum of Rain_Ta_value prior the avalanche event
Wind speed [m/s]	Wsavg	Wsavg_value	Daily mean wind speed
		Wsavg_value3	3 day moving average of wind speed before avalanche release (day when Aval occured)
Wind direction [°]	WD	WD_value	Daily circular mean of wind direction
		WD_value3	3 day moving circular average of wind direction before avalanche release (day when Aval occured)

rence represented as Ad and NAd in the model, and b) quantify how much important each variable is and test the model's performance.

A severely imbalanced dataset (91 / 271 of wet / slab Ad and 6588 / 6643 of NAd, respectively) makes the learning process difficult. Hence we created balanced datasets to enhance the skill of the model. We tried to balance the datasets by upscaling
195   avalanche records and synthetic data generation (SMOTE); the latter method was more suitable for evaluating model efficiency, hence we used the approach for Random forest method (RF) - see further in Sect. 3.4. The former upscaled method was better for descriptive purposes and we used it only for one Decision Tree (DT) - see further in Sect. 3.2. The synthetic data generation method (Lunardon et al., 2014) overcame imbalances by generating artificial data via synthetic minority oversampling technique (SMOTE) (Chawla et al., 2002). A SMOTE algorithm creates artificial data based on feature space similarities from
200   minority samples. It uses bootstrapping and k-nearest neighbour (KNN) algorithm and works as follows:

- KNN takes the difference between the feature vector (sample) under consideration and its nearest neighbour,

- differences between neighbours are multiplied by a random number ranging from 0 to 1 (can be adjusted to maintain dispersion in data - our case),

- new data are added to the feature vector under consideration.

205   Synthetic data had much higher dispersion than the original dataset (after application of aforementioned SMOTE approach); therefore, we generated data closer to our input data by adjusting the natural neighbour algorithm with Kernel density (ranging in our case from 0 to 0.5). Thanks to the SMOTE method, we can generate data with similar statistical distribution along with the two classes of avalanche types.

To check distributions of initial datasets of weather variables of original and synthetic data, we used boxplots divided by each
210   variable, Quantile–quantile (Q–Q) probability plot (Loy et al., 2016), and PCA (Maćkiewicz and Ratajczak, 1993). Box plots of all variables are essential to understand the data for further modelling. By plotting quantiles of original and synthetic data against each other, we compared probability distributions. To see the distance of correlation variables we created dendrograms (Murtagh and Legendre, 2014). PCA reduces the dimensionality of datasets by creating components where standardized data are projected on the variable axis. PCA suggests the main differences in data for both datasets.

215   Prior to evaluation of the models, we split datasets into train and test sets. For evaluation we used confusion matrix (CM), Receiver operating characteristic curve (ROC), and Area under the curve (AUC). CM makes a two way frequency table and compares predicted versus actual classes (Table A1). ROC plots true / false positive rate at every possible threshold. The ROC curve is the ratio of the sensitivity and specificity; by increasing one measure, the other is decreased. The AUC coefficient is defined as an area under the ROC curve and is a single-number summary of a model's predictability, ranging from 0 to 1; when
220   AUC == 1 it means that the model is correct all the time at predicting (Biecek and Burzykowski, 2021; James et al., 2013).

While we use the Decision Tree (DT) (Therneau and Atkinson, 2019) method to find out some variable threshold separating Ad and NAd, Random Forest (RF) (Liaw and Wiener, 2002) is used to get the most significant variables of importance determining wet and slab avalanches. Whereas RF focuses on predictive performance, DTs present relationships in a way re-

sembling to human decision-making processes and are thus a helpful tool to understand the avalanche problem identification

225 and assessment process (Horton et al., 2020).

The RF method was used to predict a binary target variable, specifically the occurrence of unique Ad and NAd, with multiple predictor variables; in this case, weather and snow variables. Generally, RF is robust to overfitting and yields an accurate, non-linear model due to bootstrapping observations from the bag which generates error rate: (OOB) and random pick of variables being compared at the moment of splitting data for classification. After training enough trees, stabilization the model's OOB

230 error could be reached, gaining credible results. Hence we did not have to conduct cross validation to prevent overfitting. RF classifiers can improve prediction accuracy, however at the cost of interpretability (Dkengne Sielenou et al., 2021). We obtained an overall summary of the importance of each predictor in the RF due to bagging. RF's benefit is that it can also perform feature importance of variables by conducting permutations on all the variables. Therefore we can distinguish what influence each of the variables has on the accuracy of the decision classifier process if we excluded a variable. A larger value

235 of variable importance indicates a more important predictor (mean decrease of accuracy, MDA hereafter) (Gregorutti et al., 2017). To gain better performance of the RF, hyperparameters should be tuned. Hyperparameters impact model fit and vary from dataset to dataset (Probst et al., 2019). We picked the best parameters based on a grid search technique. A search grid is a set of hyperparameter combinations to be tuned for an algorithm and, by cross validating and re-evaluating the model, we found the best performing parameters based on AUC.

240 ## 3   Results

This section presents statistical analyses of the avalanche datasets and relates meteorological and snow variables to wet and slab avalanches. Firstly, we show the results of long-term avalanche activity in the study location. We investigate decadal changes in wet and slab avalanches. Furthermore, we check the statistical distribution of datasets and assess the explanatory power of the variables concerning wet and slab avalanche activity using three methods - Principal Component analysis, Decision Tree

245 and Random Forest models.

3.1   Long-term wet and slab avalanche activity in the study location

In the Krkonoše Mountains, an average of 20 snow avalanches are reported each year. This number varies greatly year–to–year and ranges from 0–77 records (2010–2011, 2005–2006 respectively). We focus on wet avalanches, defined by wet snow presence in a release zone (186 records 14.9 % of all snow avalanches in the dataset) and slab avalanches, the most frequent

250 avalanches (830 records, 68 %) in Krkonoše mountain range. Mountain regions worldwide are susceptible to these types of avalanches (Soteres et al., 2020). The percentages of avalanche activity are related to 1246 avalanches recorded over the period 1961–2021.

It was revealed that the number of wet avalanches classified in the cadastre as wet, i.e., C=2 (186 Aval), has increased since the winter season 1961–2011. However, it has slightly decreased again in the last decade, 2011–2021 (Fig. 2). The

255 highest number of wet avalanches was observed in 2005 and the 2001–2011 decade. The decadal avalanche length magnitude
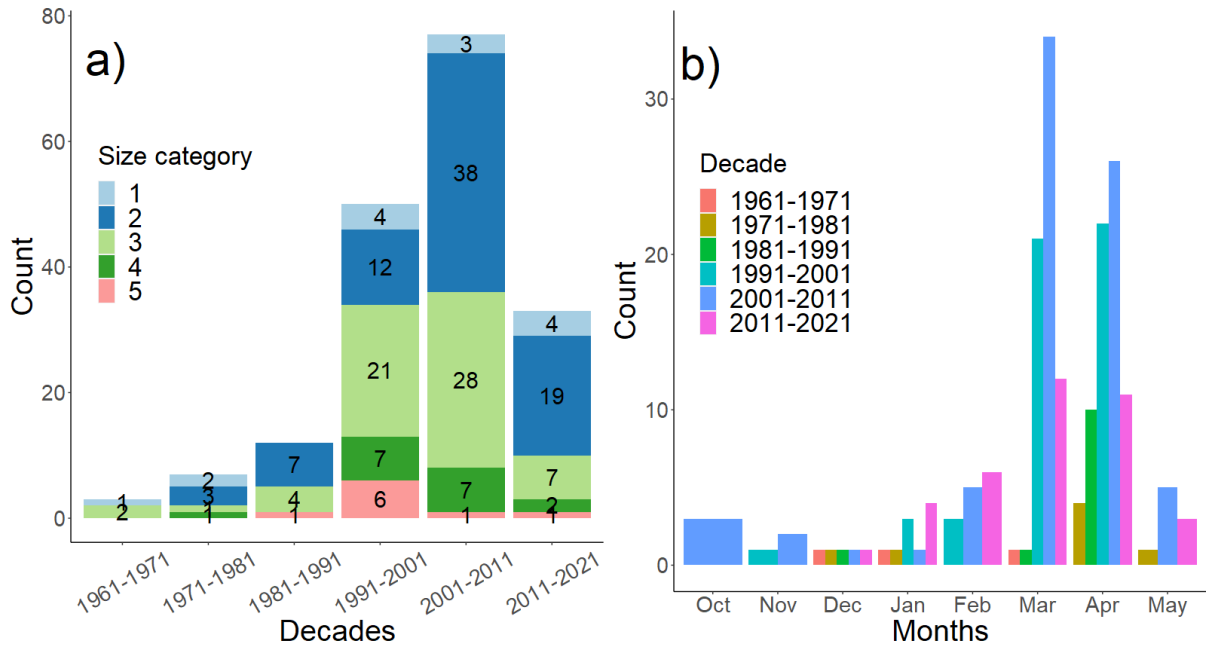
Natural Hazards
and Earth System
Sciences
Discussions
EGU
Open Access

Fig. 2. Occurrence of wet avalanches over the last six decades. a) Total number and avalanche size over decades. Each decade represents the snow season starting the 1st of October and ending the 31st of May. Four events are omitted due to missing value of length. b) Total numbers over months.

denotes rising avalanche sizes (2,3) (see Table 1) from 1971 until 2011, then it decreases again. Very large wet avalanches (size category 5) appeared during 1991–2001. There is a less extreme magnitude of avalanches in the older period (1961–1991) than in the period 1991–2021. Over the last three decades there were about seven times more wet avalanches than in 1961–1991. Wet avalanches mainly occur in March and April and reach small (2) to medium (3) size. The biggest increase in avalanche

260   occurrence in the last three decades was in March, followed by April and also February (Fig. 2). Mean snow depth increased from 1961 until 1999. Since 2002 it has slightly declined (Fig. 10).

There might be a decreasing trend in slab avalanche records, with alternating decade increases and decreases (830 records). Avalanche magnitude (medium - 3, large - 4, and very large - 5) has declined since 1961, with decadal variability. Avalanche size 3 is the most frequent (see Fig. 3). Slab avalanches releases dominate in March and mainly occur from December to March.

265   The overall mean snow depth has gently decreased since 1991 (Fig. 10). When comparing two 30 year periods (1961–1991, 1991–2021), slab avalanches diminish. In the last three decades, slab avalanches occur also in April which was not that typical in the older period (Fig. 3).

3.2 Exploratory datasets analysis of original and synthetic data

Initial dataset distributions of weather variables of original and synthetic data show similar distributions of box plots of wet

270   (Fig. A3) and slab avalanche (Fig. A4) datasets. Although some distortion can be found in the Q–Q plots, the most significant
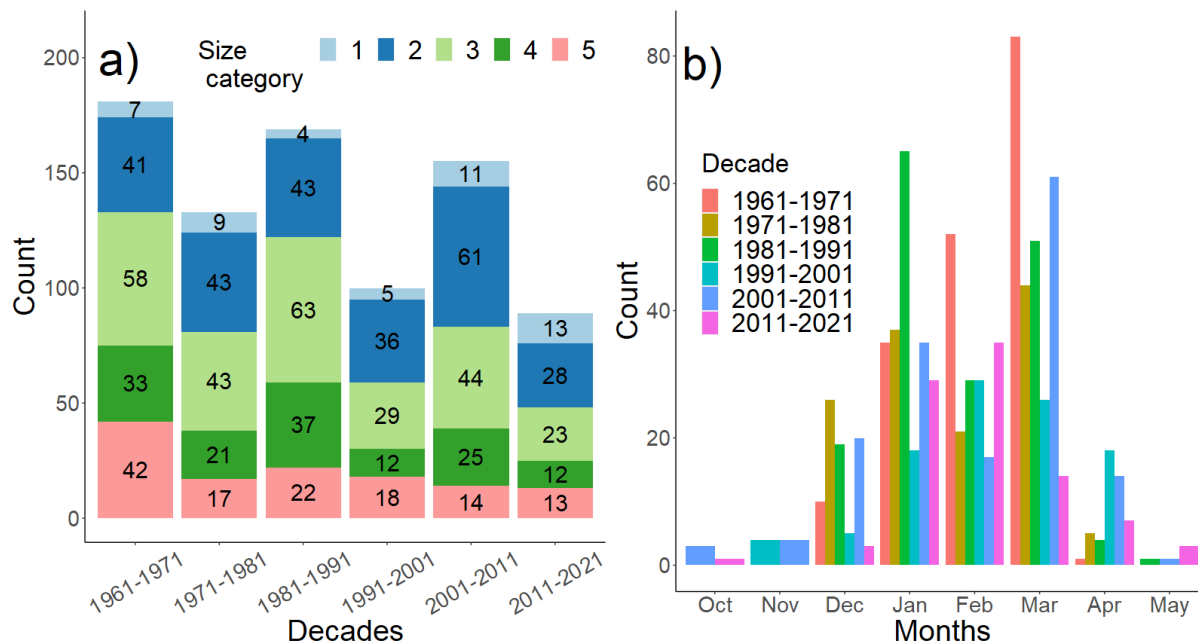
Fig. 3. Occurrence of slab avalanches over the last six decades. a) Total number and avalanche size over decades. Each decade represents the snow season starting the 1st of October and ending the 31st of May. Three events are omitted due to missing value of length. b) Total numbers over months.

parts overlap in the wet (Fig. A5) and slab (Fig. A6) avalanche dataset. PCA of Ad and NAd suggests the main differences in data for both datasets of wet (Fig. A7) and slab (Fig. A8) avalanches.

According to PCA clusters displaying only avalanche records of both datasets, wet avalanches are probably explained by snow depth (SD), rain ($Rain\_Ta$), sunlight duration (SLd) and precipitation (P). In contrast, slab avalanches are apparently driven by snow depth difference (SD_dif2) two days prior to the Ad, new snow (NSS), air temperature ($Tair$), and partly by humidity (H), and precipitation (P). Wind direction and speed did not have a substantial effect on PCA analyses that it why these variable's labels were eliminated in (Fig. 4) for better clarity of the figure. As our PCA describes only partly the variability of data (three direction PCA components explain 64 % of variability), we employed another methods.

3.3 Decision Tree of wet and slab avalanches

We analyse weather variables determining triggering of wet and slab avalanches in the period 1979–2020. The wet avalanche dataset contains 91 unique wet and slab avalanche dataset contains 271 avalanche days.

The daily mean snow depth was the primary split in the decision tree of wet avalanches that splits days with and without slab avalanches (Fig. 5). The group of days with more than or equal to 99 cm had a probability ($p = 0.38$) that an avalanche would occur. When 3-day moving average air temperature Tair_value3 >= -3.7 °C we obtained a 0.69 likelihood of avalanche (Aval) trigger, using 48 % of the observational data. If SDvalue3 is higher than 177 cm there is high probability of avalanche release
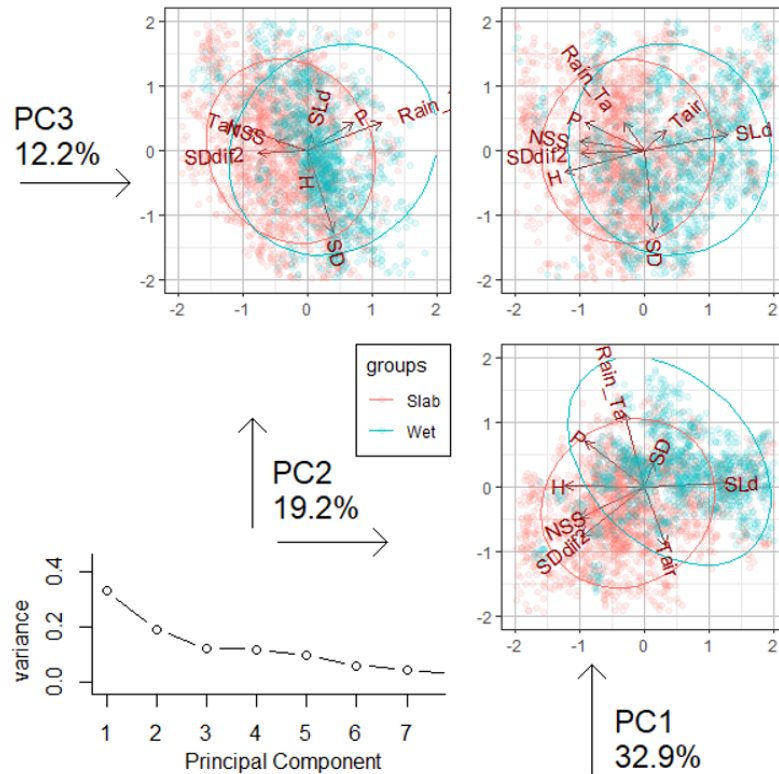
Fig. 4. The best predictors of avalanche events assessed by PCA using wet and slab synthetic datasets.

($p = 0.79$). When SDvalue3 < 177 cm slightly above zero 3 day minimum air temperature is likely to trigger an avalanche (Tmin3 >= 0.45 °C, $p = 0.68$). Other significant splits ($p > 0.9$) are mean wind speed (WS_avg_value3 >= 5.1 $m \times s^{-1}$) and wind direction (WD_avg_value3 >= 282) (Fig. 5). The higher the snow depth, the higher the probability of avalanche trigger. This might be because of the fraction of wet snow compared to dry within the snowpack.

290   Snow depth was the primary split in the decision tree that splits days with and without slab avalanches (Fig. 6). In the group of days that had SD_value more than 47 cm, there is a 0.36 probability that an avalanche will occur. However, if SD < 47 cm, not releasing Aval is insecure ($p = 0.05$). The second split node (using 61 % of observations used for splitting each node) separates with 0.57 likelihood Ad and NAd. When 3-day mean new snow (NSS_value3) >= 3.8 cm, an avalanche occurs ($p = 0.77$), but when it is lower an avalanche will not be released. The higher the 3-day mean snow depth SD_value3 is (>= 134 and

295   195 cm), the higher probability of avalanche release. If snow depth difference between four days before Aval record (SDdif4) is higher than 13 cm, the avalanche hazard increases. Avalanches do not usually occur on easterly-north easterly oriented slopes (WS_value3) > 108 °, when wind speed is lower than 11 $m \times s^{-1}$ ($p = 0.24$) and 3 day air temperature amplitude Tamp3 < 6.6 °C in the Krkonoše Mountains.
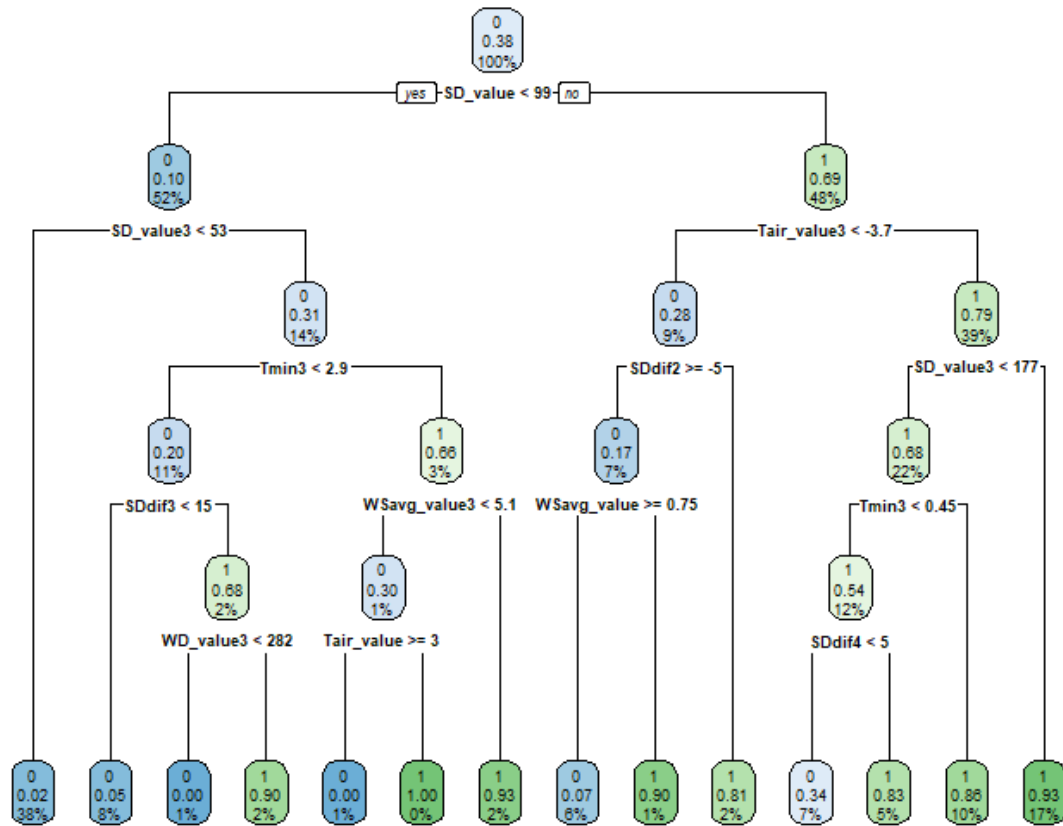
Fig. 5. Decision Tree of weather variables triggering wet avalanches.

3.4 Random forest model of wet and slab avalanches

300 The random forest model ranked the most important variables based on Variable Importance (VIP) using MDA. The most important variables for wet avalanches seem to be 3-day maximum and minimum air temperature (Tmax3, Tmin3), snow depth (SD_value3, SD_value), wind speed (WS_value, WSavg_value3), wind direction (WD_value3, WD_value), and rainfall separated by wet bulb temperature (Rain_Tw_value) (Fig. 7). Sunlight duration (Sld_value) and precipitation are almost of 1.6 times less importance than 3-day maximum air temperature. From snow depth difference variables the most important is when 305 it is 2 days different from avalanche day (SDdif2). Wet bulb temperature is counted from humidity so humidity also plays role; however, its importance is 25.

The most important variables for slab avalanches in the daily forest are the most likely snow depth (SD_value, SD_value3), rainfall variables based on threshold temperature (Rain_Ta_sum3, Rain_Ta_value, Rain_Ta_value3), new snow (NSS_value3,
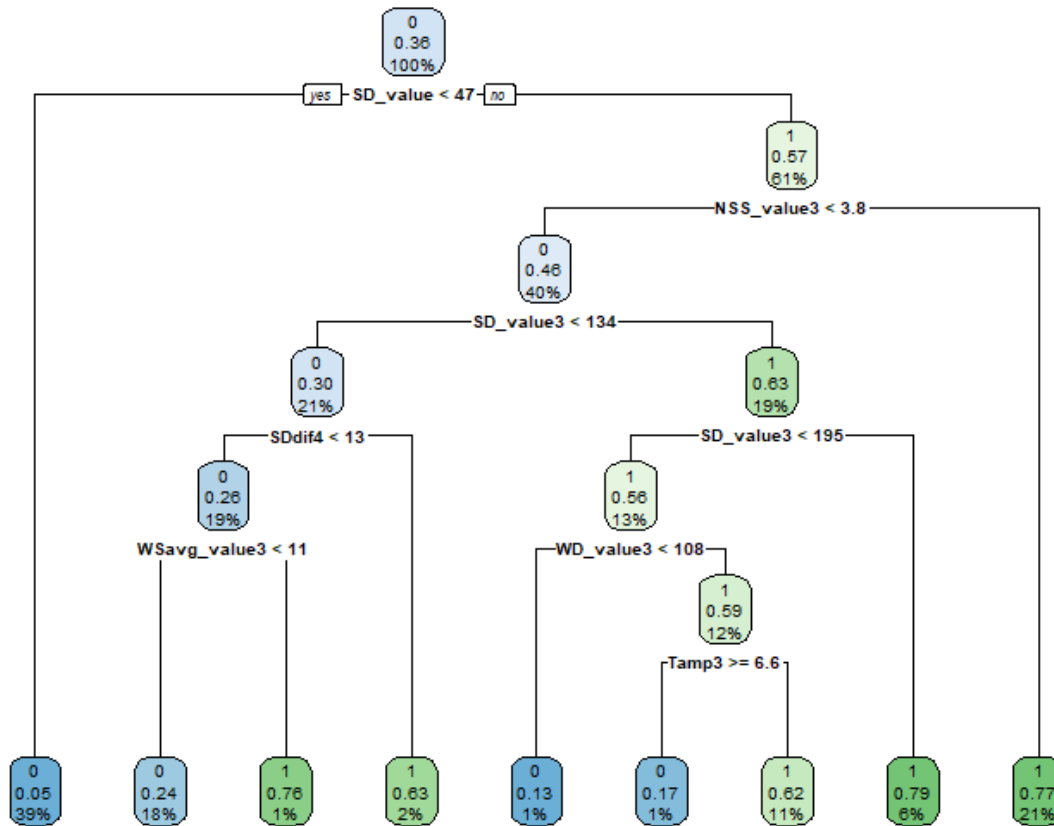
Fig. 6. Decision Tree of weather variables triggering slab avalanches.

NSS_value), wind speed (Wavg_value3, Wavg_value), and air temperature (Tair_value). Daily mean air temperature was about

310   1.3 times less important than daily mean snow depth (Fig. 8). The results show that rain and snow related variables are more

important than air temperature (Tair_value).

3.5   Random forest model performance on original data

The wet avalanche slab model predicts 84 from 91 avalanches and 6555 non-avalanches from 6588. RF model correctly predicts

slab avalanches 254 (true positives) / 271 and 5813 (false negatives) / 6643 slab avalanche days. There were 33 / 830 false-

315   positives wet / slab avalanches respectively (Fig. 9), which means that the model tends to falsely predict wet / slab avalanches

that were not marked as non-avalanches. In the actual scenario, we would falsely inform that there is a high probability of

avalanche occurrence. The performance of the model according to AUC values is very good 0.992 and 0.97 for wet and slab
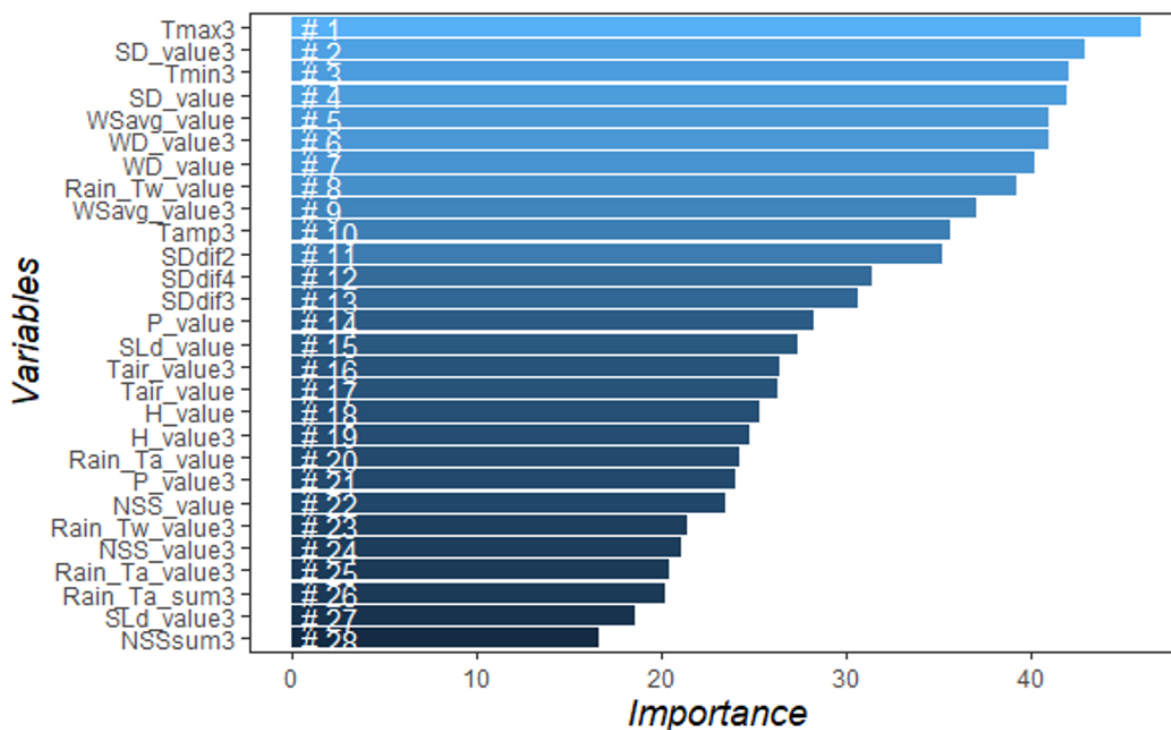
avalanches (Fig. 9).

Fig. 7. Variable importance using the Mean Decrease Accuracy (MDA) for each variable of wet avalanche dataset in the random forests. Variables are described in Table 2.

4 Discussion

320 Even though there is a lot of evidence of anthropogenic impacts on climate (e.g. Hock et al., 2019; Strapazzon et al., 2021), the effect of climate change on snow avalanches is poorly understood. Moreover, snow avalanche forecasting is still challenging (e.g. in Czechia, the avalanche forecasting model is missing and avalanche danger levels are assessed by human expertise). Therefore, in this work, we aimed to determine the main predictors of wet and slab avalanche release and revisited the importance rating of snow and meteorological variables according to their significance in avalanche triggering in a low altitude

325 mountain range. This might help incorporate identified driving variables into avalanche warning decisions and to awareness of how warming will influence the type and number of avalanches.

4.1 Avalanches days determined by weather variables using PCA and RF method

The first hypothesis is that "there might be more wet avalanches because of rising air temperature and decreasing snowfall / rainfall ratio in Czechia, and avalanches are also determined by sunlight duration (a form of shortwave radiation)." The

330 second hypothesis is that slab avalanches are mainly driven by snow depth, new snow sum and wind speed redistributing snow. Using the random forest method, we found that the most important triggering factors of wet avalanche activity in the Krkonoše
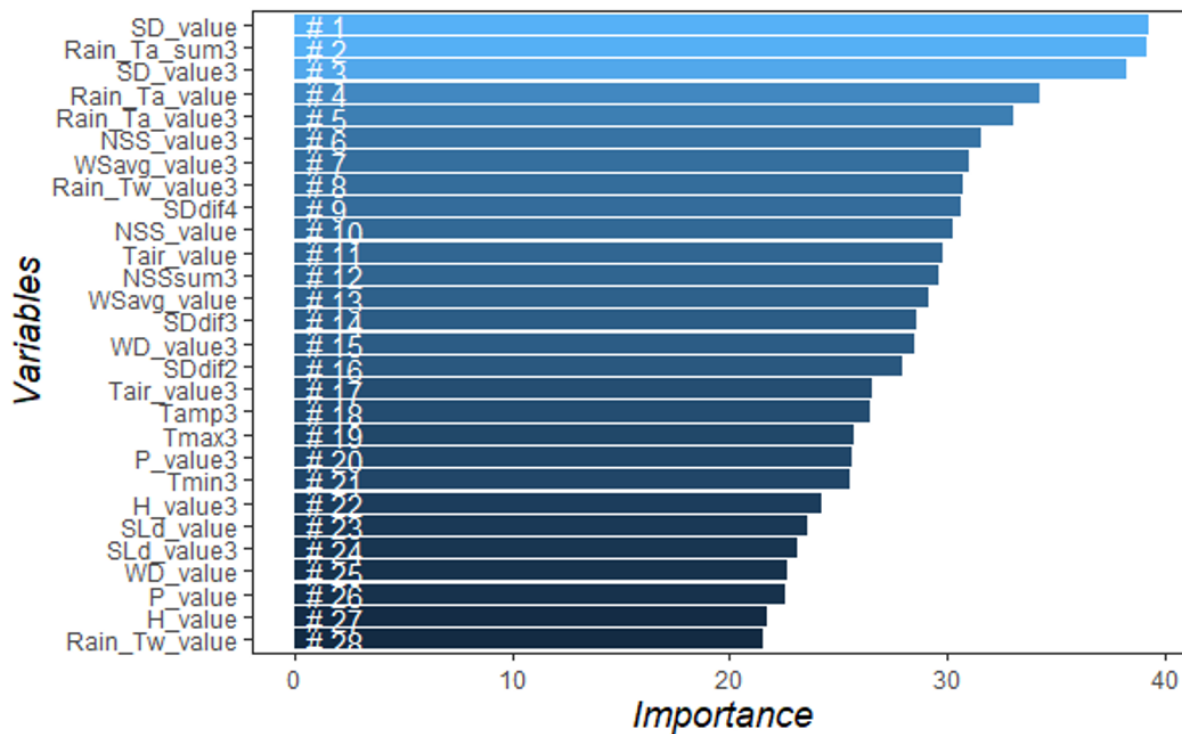
Fig. 8. Variable importance using the Mean Decrease Accuracy (MDA) for each variable of slab avalanche dataset in the random forests. Variables are described in Table 2.
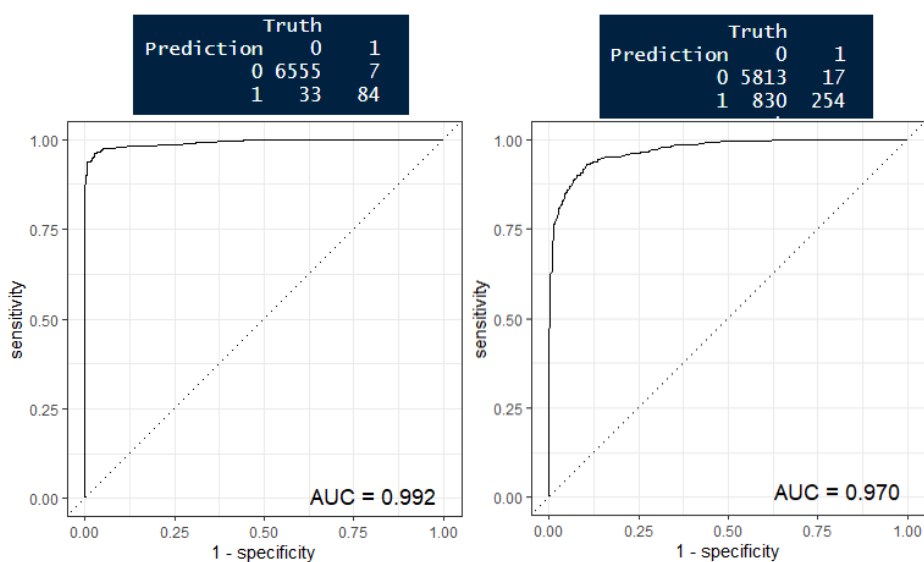

Fig. 9. Random forest model fit on original wet and slab avalanche datasets using CM, ROC, AUC metrics.

Mountains are mainly determined by: 3-day air minimum and maximum air temperature, 3-day moving average snow depth, and both 3-day circular mean wind direction and wind speed. Slab avalanches are more influenced by snow depth, rainfall related variables, new snow variables, and wind speed. PCA rather suggested the explanatory variables: snow depth, sunlight
335 duration and combination of liquid, solid precipitation, and air temperature. It is likely that wet avalanches are more influenced by air temperature and slab avalanches by rain rather than air temperature. PCA points at snow depth difference, new snow, and mixed precipitation and air temperature related effect. The results partially confirmed the hypothesis.

    As we expected, both wet and slab avalanches are dependent on snow depth. Moreover wet avalanches are presumably triggered by air temperature related variables (more likely maximum and minimum than mean air temperature) than liquid or
340 solid precipitation (Rain_Tw_value, P_value). Precipitation and sun light duration influence wet avalanches; however, from our RF results their contribution is almost twice lower than for air temperature in the Krkonoše Mountains (Fig. 7). Slab avalanches are most likely triggered by rainfall separated by threshold air temperature, new snow, and wind speed (Fig. 8). Air temperature does play a role to some extent; however, daily mean air temperature was about 10 % less important than daily mean snow depth. It seems that slab avalanches are triggered more because of rain rather than snow. Furthermore, wet and slab avalanches
345 are influenced by wind speed; this might be caused by the LBOU meteorological station position. Around the station there are open plains of level surfaces, which accelerate the westerly winds. Afterwards, the wind falls into the deep cirques behind the plains, causing massive turbulence which influences snow conditions of avalanche paths. Also, wet avalanches are determined by wind direction; wind usually redistributes snow, so new snow might be a less significant variable for slab avalanches in the Krkonoše Mountains. The dry warm wind, known as "föhn", can cause very intense melting or avalanches. On the Czech side
350 of the Krkonoše Mountains, it arises when the wind blows from the north (from Poland).

    The main variables possibly triggering wet avalanches are air temperature, snow depth, wind speed, and wind direction in RF and DT. Sunlight duration and rainfall (Rainfall_Tw_value) were important variables in both RF and PCA (Fig. 7), (Fig. 5). Wind speed and direction did not have a considerable effect on PCA.

    Daily mean snow depth (SD_value) was the most important variable in random forest for predicting slab avalanches, and it
355 was the primary split for decision trees. DT and RF are in accord in most of the RF importance variables higher than 25 MDA (Fig. 8), (Fig. 6) except for rainfall variables (Rain_Ta_value, Rainfall_Ta_value3, Ranfaill_sum3). However, precipitation is significant in PCA as well as new snow, snow difference, air temperature and humidity (Fig. 4). In contrast, in PCA results the importance of wind speed was not significant in comparison with RF and DT.

    Wet avalanches are also part of the slab avalanche dataset so-called "wet slab avalanches" as the datasets were created
360 according to criteria, zone of origin (known as release zone), a) manner of starting, and b) liquid water in snow according to Quervain et al. (1973) Avalanche classification. Wet and slab avalanche datasets have 53 avalanche days with the same weather conditions within the inner join. This fact makes their model results similar to some extent and explains analogous Variable Importance of RF plots.

4.2 Changes in long term avalanche activity

365 As within the alpine environment, more road accidents happen due to avalanche fencing, so attention was paid to changes in
the flow characteristics, such as the formation of shorter and less predictable runouts (Eckert et al., 2013; Naaim et al., 2016).
Moreover, it is becoming more common that very large avalanches which originally released in dry snow drag along warm
snow in the avalanche path below (Eckert et al., 2013; Naaim et al., 2016; Dkengne Sielenou et al., 2021). This is consistent
with medium confidence in an increase in avalanche activity involving wet snow, and a decrease in the size and runout distance

370 of snow avalanches over recent decades, particularly in Europe (Hock et al., 2019). This is confirmed by our results of the
accelerated increase of wet avalanche occurrences for the last 30 years (7 times more avalanches than in the previous 30 year
period) (Fig. 2). According to Naaim et al. (2016), wet avalanches occur more frequently, even in winter from December
through February. In contrast, we have not found an increasing number of wet avalanches in December. However, in January
and mostly in February, there are more occurrences in the last three decades than there were in rest of the studied period. Slab

375 avalanche sizes (medium - 3, large - 4, and very large - 5) have slightly decreased with decadal changes since 1961 (see Fig. 3).

Our data indicate that wet avalanches in the Krkonoše Mountains are more related to increasing air temperature causing
snowmelt rather than rainfall. Beside the RF model (see Fig. 7), this also shows rising temperature during last two decades
(about 1.85 °C), which is depicted in Fig. 10. Total precipitation, and thus rainfall, have also increased over the last two decades,
but a detailed analysis of the data shows a decreasing rainfall amount a few days prior to the wet avalanche events. Rising

380 temperature in the study region was also documented by Kliegrová and Kašičková (2019) and, therefore, snowpack has become
thinner since 2002 (Fig. 10). Snowmelt is, beside temperature, also driven by solar radiation or wind speed; our results show
increasing total sunlight duration (a proxy of solar radiation). Wind speed is usually considered as an important driver governing
turbulent heat exchange between snowpack and atmosphere, but it seems that wind speed has slightly decreased in the Krkonoše
Mountains over recent snow seasons (Fig. 10). The decreasing trend in wind speed probably affects snow deposition, snowmelt,

385 and thus wet and slab avalanche activity. Moreover, air temperature or rain prevalence is probably elevation-based. As Nedelcev
and Jeníček (2021) suggest, snowpack is more sensitive to changes in air temperature at elevations below 1200 m a.s.l. and
precipitation at elevations above 1200 m a.s.l. in Czechia mountain catchments. Moreover, Baggi and Schweizer (2009) related
wet snow avalanches to precipitation within the Eastern Swiss Alps and weather data at an elevation of 2300 m a.s.l., and they
propose that wet instability is strongly influenced by snowpack properties related to the warming of snowpack and meltwater

390 production. Our results confirm the finding of Laute and Beylich (2018) that the probability of wet snow avalanches increase
due to more frequent periods with air temperatures close to or above freezing point during the winter period (Fig. 10). Peitzsch
et al. (2012) emphasize the variable importance on air temperature and more on snowpack settlement and snow water equivalent
loss, explaining wet slab avalanches in Glacier National Park, Montana, USA.

Avalanche activity of north easterly oriented avalanche release zones in the Eastern Swiss Alps was primarily related to

395 snow depth, precipitation, and air temperature (Baggi and Schweizer, 2009). This finding is in line with our study; however, for
Krkonoše leeward S-SE oriented release zones, wind direction and wind speed influence snow deposition and thus avalanche
activity. Baggi and Schweizer (2009) revealed that wet-snow slab avalanche days (primarily observed in May) were signifi-
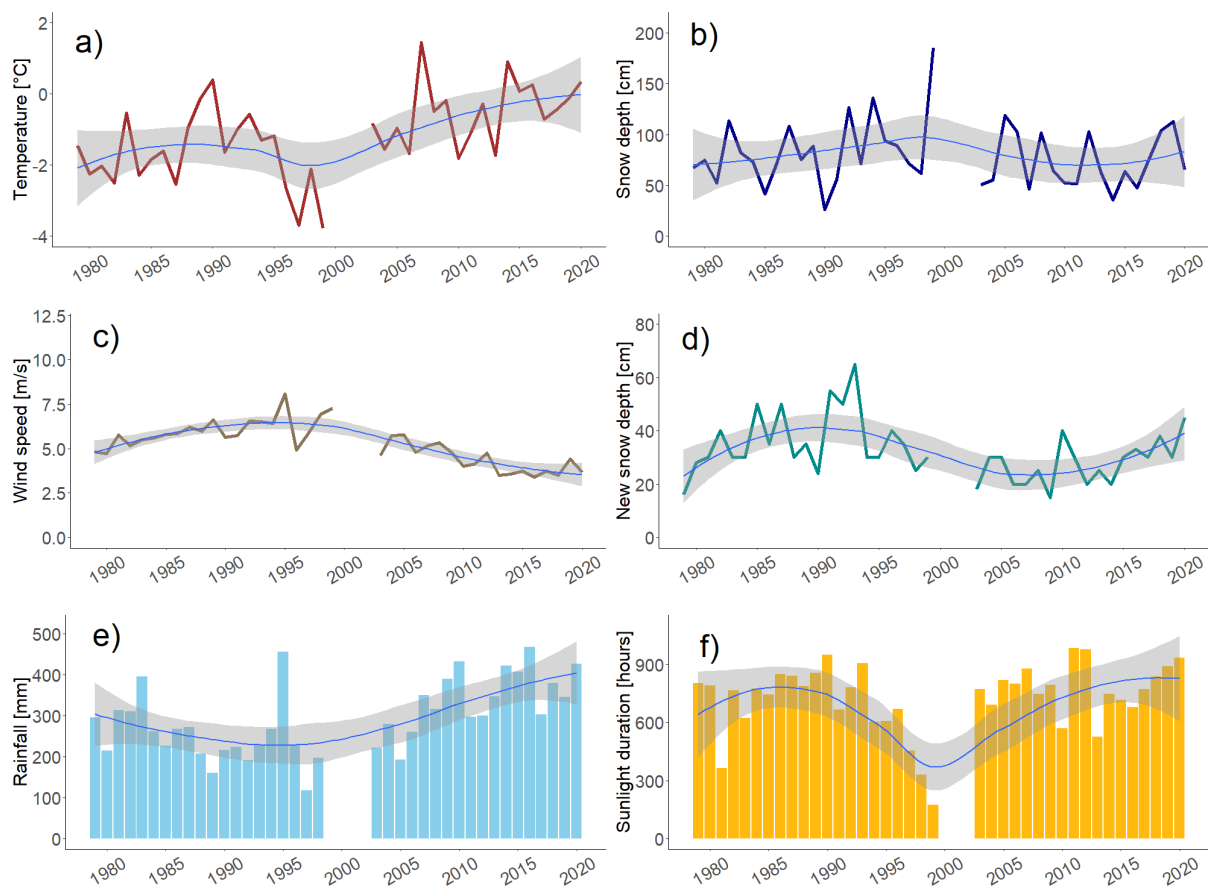
Fig. 10. Long-term regime of meteorological and snow variables from Labská bouda automated weather station, 1979 to 2020. Each year represents snow season data aggregation from 1.10. to 31.5. Specifically: a) mean air temperature, b) mean snow depth, c) mean wind speed, d) maximum new snow depth, e) rainfall sum, and f) total sunlight duration. The time series includes a data gap from seasons 2000–2002 for all variables except rainfall, when a data gap occurred from 1999–2002. Blue line shows local polynomial regression fitting function and grey envelope represents 0.95 level of confidence.

Natural Hazards
and Earth System
Sciences
Discussions
EGU
Open Access

cantly (p.05) related to several variables including to minimum air temperature and the sum of positive air temperatures of 3 or 5 days, as well as rain and decrease in snow depth over 3 days, soon after the snowpack had become isothermal and reached

400   0 °C, and the presence of capillary barriers. In a non-European environment, Bridger Bowl (SW Montana) days with deep slab avalanches on persistent weak layers often had warmer air temperatures for a minimum of 24 hours (days were typically preceded by 3 days above freezing), and more precipitation over the prior 7 days, than days without deep slabs on persistent weak layers (Marienthal et al., 2015). The variables are equivalent to our results, see Fig. 8. Conversely, in an areas with a treeline around 1800–2100 m a.s.l., snowfall in the past 72 hours was the most significant variable for storm slab avalanche problems,

405   and slab density was the most crucial variable for persistent-slab avalanche problems in Glacier National Park, Canada (Horton et al., 2020). In the High Arctic landscape around Svalbard's main settlement of Longyearbyen (78 ° N), precipitation and snowdrift 24, 48, and 72 hours prior to an avalanche and non-avalanche day were the best predictors of avalanche activity, and min, max, and mean wind speeds could be used as indicators of avalanche activity (Eckerstorfer and Christiansen, 2011). Contrary to our findings and those of Baggi and Schweizer (2009), Eckerstorfer and Christiansen (2011) did not find any signif-

410   icance of air temperature. According to Germain et al. (2009), wind is an important snow transport agent governing avalanche activity; wind largely controls avalanche activity in a barren landscape. Thus it is not necessary for large amounts of snowfall to release avalanches as the wind can redistribute the snow quite efficiently, limited by the availability of snow (Eckerstorfer and Christiansen, 2011). In our results, a combination of barren plateaus at the top of the Krkonoše mountains is present and could explain less importance of new snow in triggering slab avalanches.

415   Existing studies focusing on weather conditions triggering snow avalanche release show somewhat contradictory results. It is difficult to relate results and trends of snow avalanche activity to climatic fluctuations due to various environmental factors that control snow avalanche activity (Laute and Beylich, 2018). We claim and have reviewed that the results vary with many geographic characteristics, such as location (latitude, longitude), elevation, type of climate (maritime, continental), local-to-regional topography, position of station: leeward / forward side, scale (small: specific roads sections / regional), and avalanche

420   type definitions. Furthermore, the variables triggering snow avalanches often change with the different models (Gauthier et al., 2017) and, based on our experience and that of Dkengne Sielenou et al. (2021) with statistical methods and machine learning methods (random forest, logistic regression, classification tree). So far, only a limited number of studies exist with long-term avalanche records as well as weather or snowpack data from different elevations relevant to snow avalanche formation (Laute and Beylich, 2018).

425   **4.3 Model performance: Random forest as a relevant machine learning method for avalanche activity**

Our models show interesting forecasting potential. Random forest is a relevant method for targeting either avalanche days or non-avalanche days of wet and slab avalanche activity according to metrics used for our datasets, assessing performance of the model: very low error rate and high accuracy of prediction of wet and slab avalanche dataset. We checked the RF models skill against the original dataset and we achieved satisfactory results for model metrics (Fig. 9) in the Krkonoše Mountains. RF

430   selection as a relevant method ties well with the previous study of Dkengne Sielenou et al. (2021).

RF model using synthetic data might present a good starting point for obtaining a feasible system to complement decision support in estimating snow avalanche hazard Dkengne Sielenou et al. (2021). However the variables with best importance differ with length of the explored data series. Therefore, every year, models should be run using data from the most recent season to ensure optimal performance of the RF method, which is in accord with the claim of Gauthier et al. (2017). Inferring physical

435   processes that drive avalanche activity can be challenging, as these statistical methods are reliant on correlations that do not necessarily represent causal links (Dkengne Sielenou et al., 2021). Since the focus of the analysis was to explore relationships rather than construct predictive models, further enhancing the model's performance is beyond this study. However, another step to get more precise predicting model could be gradient boosting or neural network. Removal of variables with low importance could also result in better RF performance.

440   ## 4.4   Limitations, uncertainty, and scope of study

The avalanche dataset analyzed in this study contains 271 unique slab Ad / 6643 NAd, 91 unique wet Ad / 6588 NAd reported over 39 avalanche seasons of daily data. The quality of the avalanche survey is variable and also depends on human observation expertise (some misclassification of the avalanche type within the dataset might be present) and a few avalanches might be recorded a bit later if the weather is unfavourable. The weather station used in this study is located approximately 0.2–15 km

445   away from the closest (27) and furthest (01) avalanche path in the west and east Krkonoše from the LBOU station (Fig. 1). On the other hand, one professional weather station for such an area extent and data series length is usual in mountain environment. The results might be partly site specific and influenced by the generally south, south-east aspect of the starting zones. In choosing a relevant meteorological station according to data availability, the location position of the station was the only option. Other meteorological stations: have short time series (like Luční bouda (LUCB, 1413 m a.s.l., data since 2009)),

450   Vítkovice (VITK, 1410 m a.s.l.) station was replaced by LBOU station, or they are located further away as Harrachov (HARR, 675 m a.s.l.). If we choose other stations, the ranges of identified variables may differ, and therefore the interpretation may change, but this does not necessarily invalidate the models (Gauthier et al., 2017).

For assessing the danger of avalanches, worldwide avalanche services could use, amongst others, meteorological data measured on a daily basis as well as expert knowledge about avalanche activity (Gauthier et al., 2017; Dkengne Sielenou et al.,

455   2021). Czech avalanche safety currently uses expert knowledge, therefore avalanche prevention might start using machine learning to get additional insights leading to better decision making processes for an avalanche warning system. The weather and snowpack variables offer a snow science perspective on what conditions favoured the formation of different avalanche problem types.

Although we studied wet and slab snow avalanches, our approach might be extended to different avalanche types, avalanche

460   return periods or extended avalanche danger rating as mentioned by Dkengne Sielenou et al. (2021) rather than too broad studies of differences between avalanches and non-avalanches. However, unification and using the same avalanche type definition (within diverse avalanche classification) would help to compare results from different studies. Dkengne Sielenou et al. (2021) suggest that avalanche activity could be separated according to elevation range, aspect, or slope. It might be interesting to relate snow profile properties and weather to avalanche releases, or to explore how the results change with a shorter time scale and

465   separate them according to the position of the meteorological stations and avalanche paths. We aimed to do that, but hourly data are only available since 2004 for LBOU station and since 2009 for Luční bouda (LUCB) station, resulting in a low number of unique avalanche days.

      From our knowledge of exploring different methods (logistic regression) we obtained result with a high p value of variables; however, when we verified the model´s best combination of variables explaining avalanche activity by confusion matrix (True

470   positive Ad, True negative NAd), we had very low predicting power for avalanche and non-avalanche days. If we were to change the length of the explored period the best predictors changes. Therefore it is evident that the skills of the model have to be evaluated by metrics, such as confusion matrix, since we can distinguish between the false and positives rates of predictions which was not reported in some studies (e.g. Baggi and Schweizer, 2009; Dreier et al., 2016; Eckerstorfer and Christiansen, 2011). These metrics provide an assessment of potential forecast reliability.

475   The purpose of this study was to determine wet and slab avalanche triggering variables in Krkonoše and the study fills a gap by adding a regional study of the Krkonoše Mountains, Central Europe, based on daily long-time data series. We presented how climate change (wetter and warmer climate, upward snowline shift (Schattan et al., 2017)) affects the frequency and types of snow avalanches and which weather variables trigger them.

5   Conclusions

480   We investigated the long-term regime of the avalanche dataset and weather variables related to the avalanche activity. Due to climate change, more avalanches involving wet snow (liquid water in the release zone, naturally released) due to snowmelt and decreasing trend of slab avalanche occurrence (the most dangerous type of avalanches for off piste skiers and tourists) have been recorded in mid-elevations of the Krkonoše Mountains, north-east Czechia. We applied the random forest method to quantitatively explain the importance of the meteorological and snow variables of wet and slab avalanche types. We used 28

485   predictors to feed the random forest model. Predictor selection, hyperparameters tuning was performed and RF model yielded high performance. The most important variables for wet avalanches were 3-day maximum and minimum air temperature, snow depth, wind speed, wind direction, and rainfall separated by wet bulb temperature. The most critical variables explaining slab avalanche activity were snow depth, rainfall variables based on threshold temperature, new snow, and wind speed. Air temperature also plays a role; however, rain and snow-related variables are more important than air temperature for the period

490   1979–2020. Our results might provide vital information for avalanche forecasting, and public authorities could use them, such as the Krkonoše National Park administrations, the mountain rescue services of Czechia, or the Institute of Forest Management. Land use management practitioners should adapt their behavior and planned management activities to simultaneously mitigate avalanche hazards and conserve unique ecosystems (requiring avalanche releases in Krkonoše National Park). This information should also be freely available for the general public (the global population is growing and more people are moving into the

495   Czech mountains), which, especially recently due to the pandemic, tends to operate more in the open landscape (e.g., ski touring), where the chances of triggering an avalanche are very high. The methodology has the power to identify driving weather variables of wet and slab avalanches. We recommend a combination of expert knowledge about avalanche activity,

Natural Hazards
and Earth System
Sciences
Discussions

snow profile measuring (stability tests and snowpack meteorological conditions), and the identified daily meteorological and snow variables to assess the avalanche hazard.

500 *Code and data availability.* Meteorological daily data for the case study are freely available and can be found at the Czech Meteorological Institute (CHMI) website: https://www.chmi.cz/historicka-data/pocasi/denni-data/Denni-data-dle-z.-123-1998-Sb. The avalanche dataset was provided for research purposes by the owner Valerián Spusta (permission of the owner is required).

Particular codes will be provided on a request. The datasets and markdown code for machine learning are available at https://github.com/MarketaS/Avalanche/tree/main/MachineLearning. The finalized version will be available at Zenodo upon publication.

505 *Author contributions.* MS: Conceptualization, Methodology, Formal analysis, Writing – original draft, Visualization. RJ: Formal analysis, Methodology, Writing - review editing. KD: Formal analysis, Methodology, Software, Writing - review editing, Validation. VM: Formal analysis, Writing - review editing, Validation. JRB: Formal analysis, Methodology, Writing - review editing, Validation. MH: Supervision, Conceptualization, Methodology, Writing - review editing, Validation.

Competing interests. The authors declare that they have no conflict of interest.

Natural Hazards
and Earth System
Sciences
Discussions

Open Access

EGU

Appendices

Cluster Dendrogram
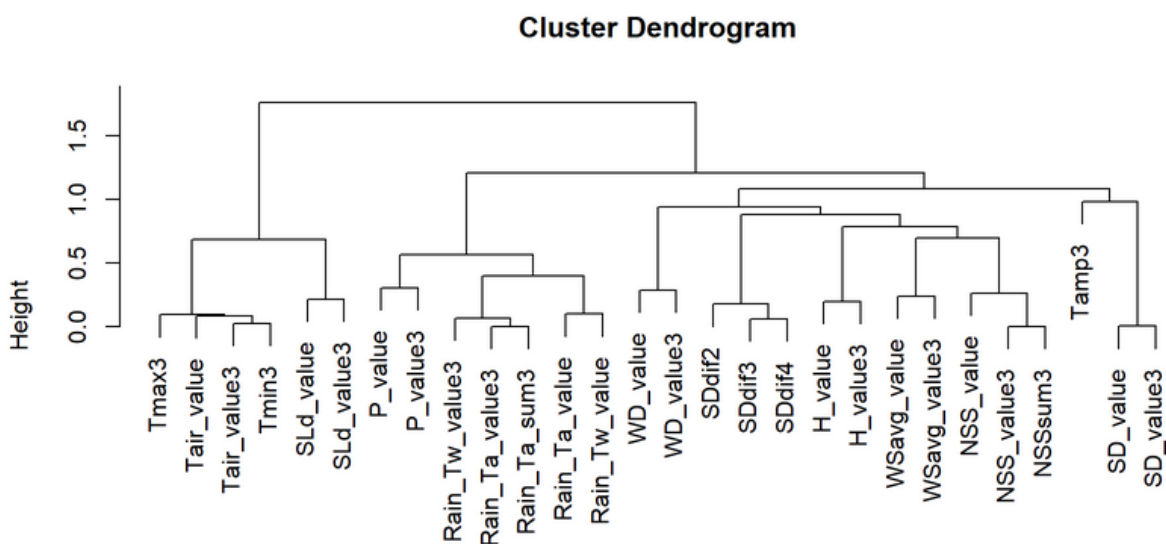
Fig. A1. Dendrograms of covariate variables of wet avalanches.

Cluster Dendrogram
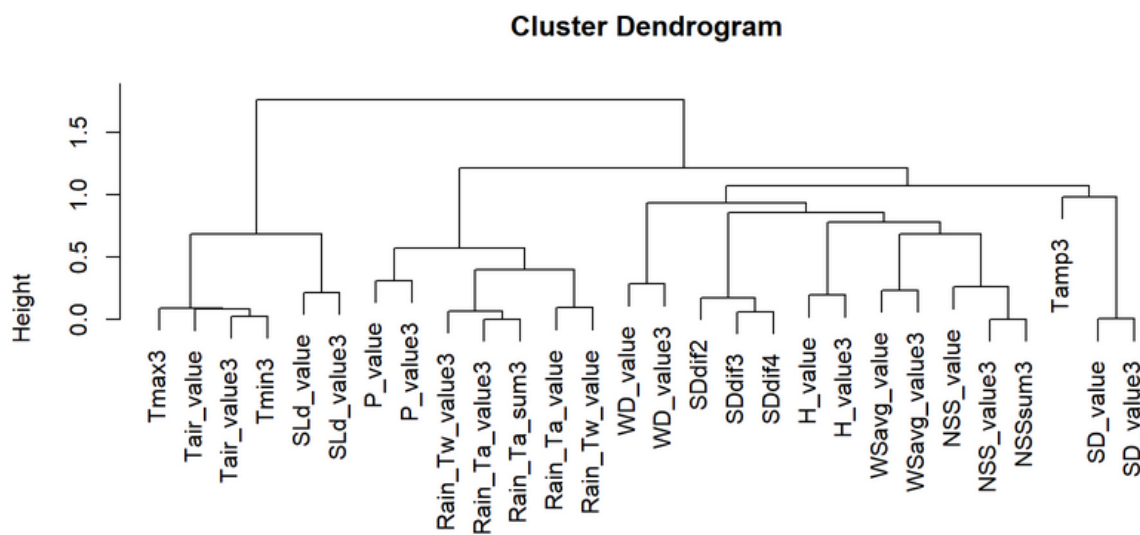
Fig. A2. Dendrograms of covariate variables of slab avalanches.

Natural Hazards
and Earth System
Sciences
Discussions

Open Access

EGU
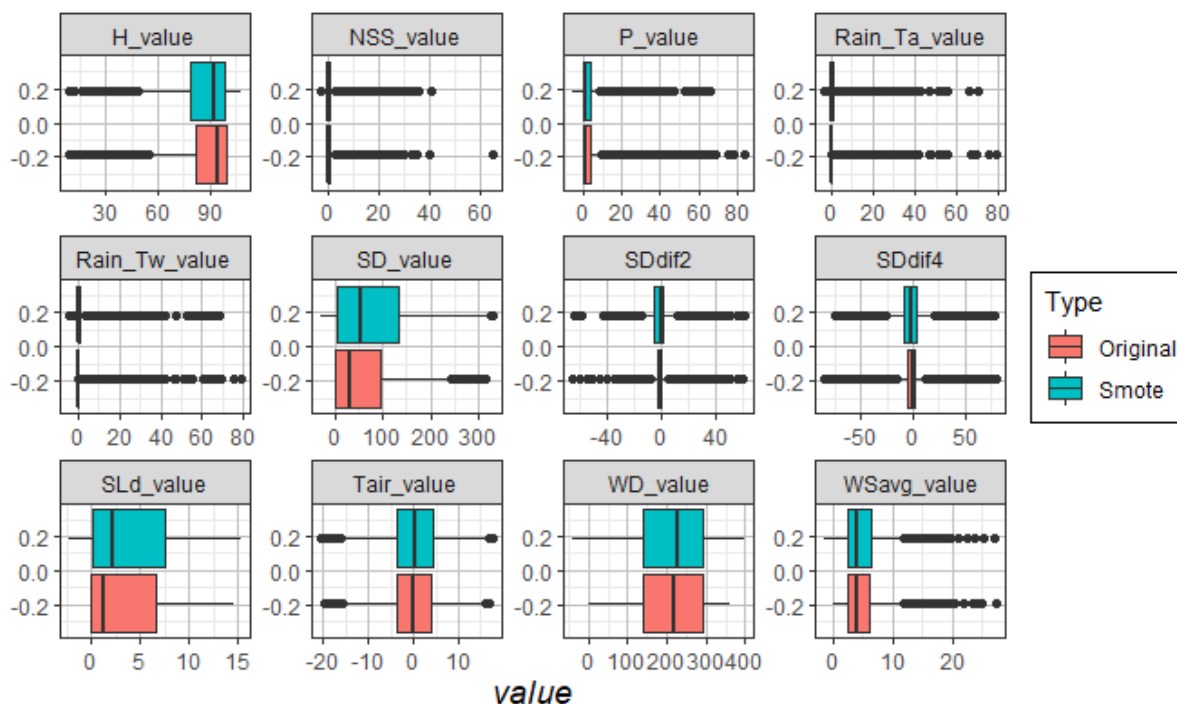
Fig. A3. Boxplot of wet avalanches of original and synthetic (Smote) data.

Natural Hazards
and Earth System
Sciences

Discussions

Open Access

EGU
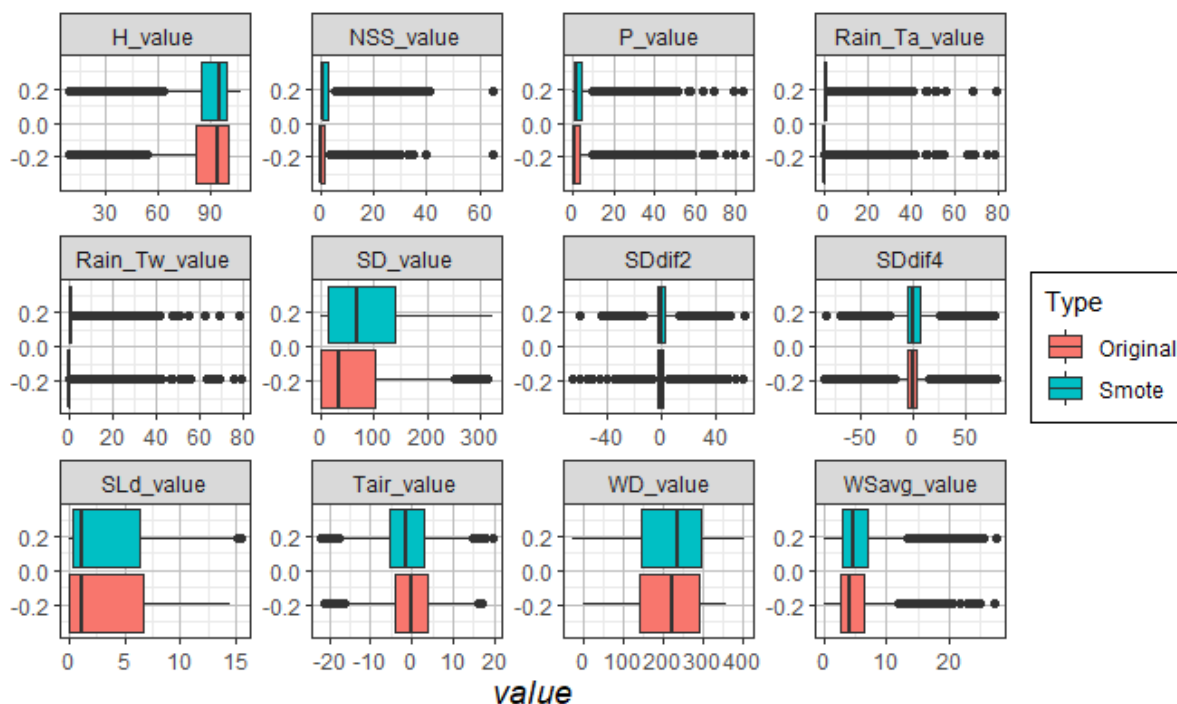
Fig. A4. Boxplot of slab avalanches of original and synthetic (Smote) data.

Natural Hazards
and Earth System
Sciences

Open Access

Discussions

EGU
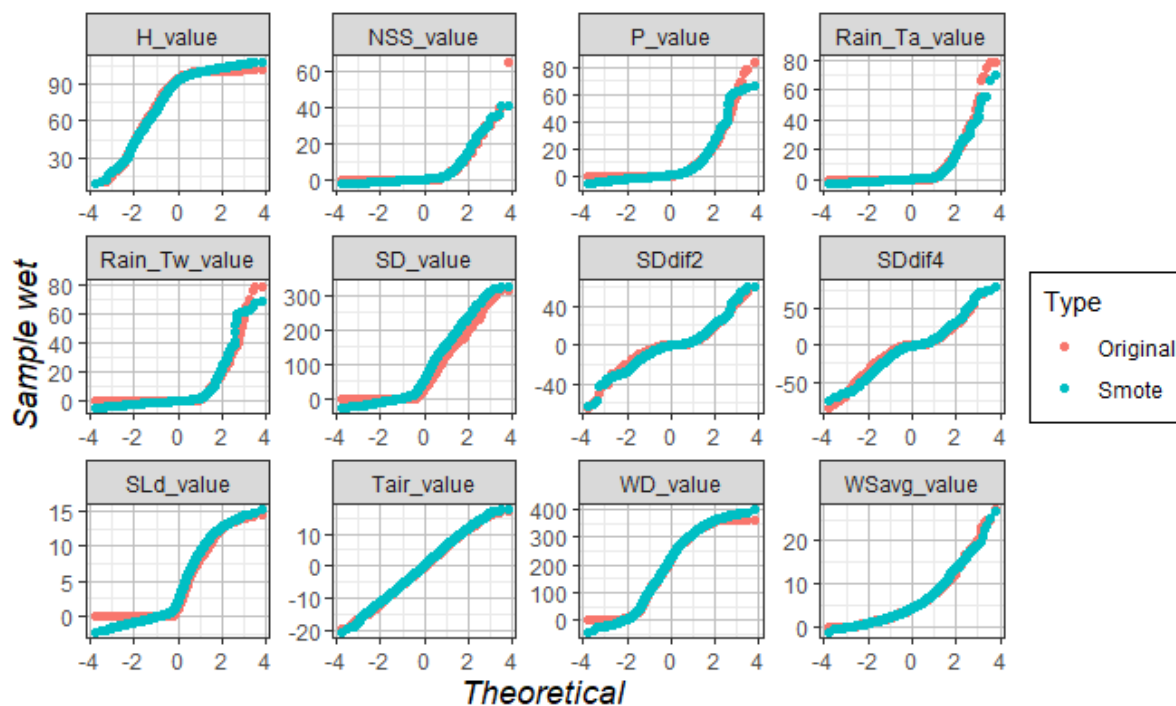
Fig. A5. Quartile-quartile (QQ) of wet avalanches of original and synthetic data (Smote).
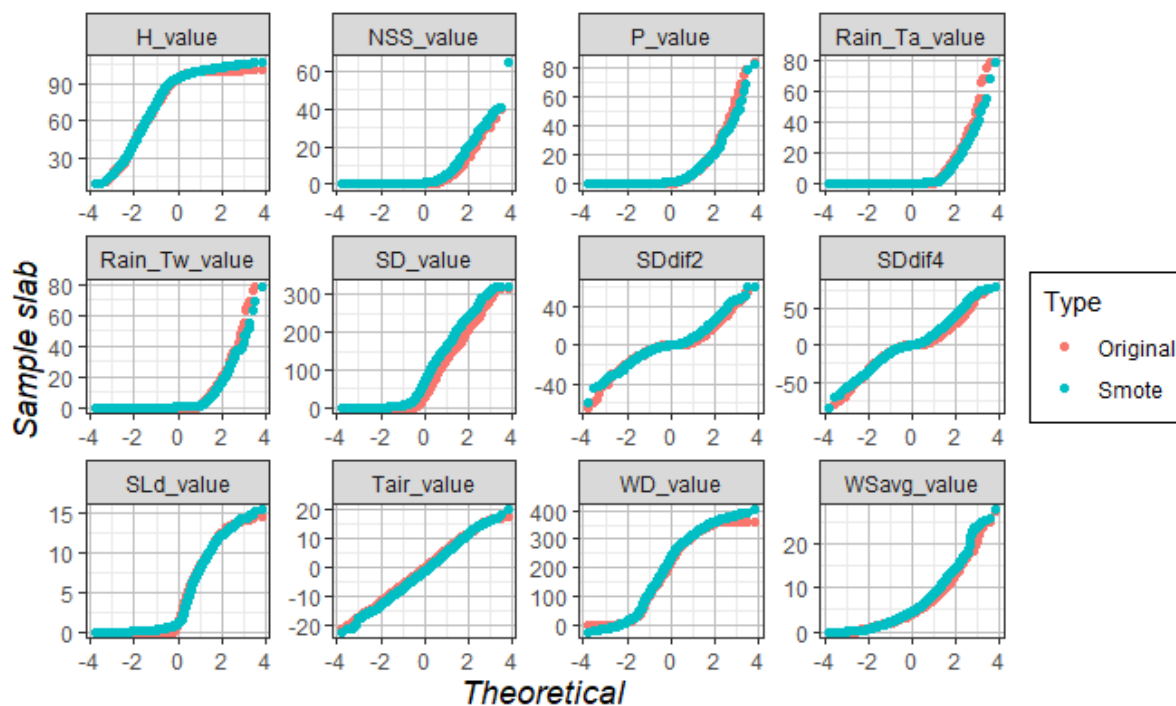
Fig. A6. Quartile-quartile (QQ) of slab avalanches of original and synthetic data (Smote).
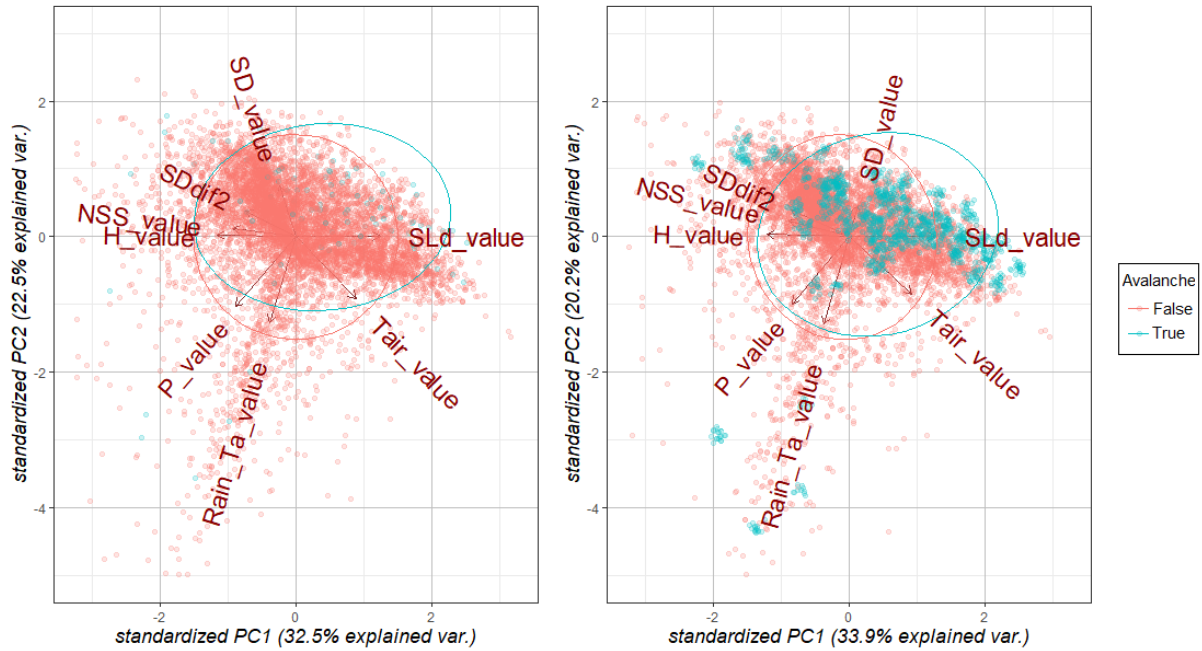
Fig. A7. Comparison of original and synthetic datasets for wet avalanches using PCA.
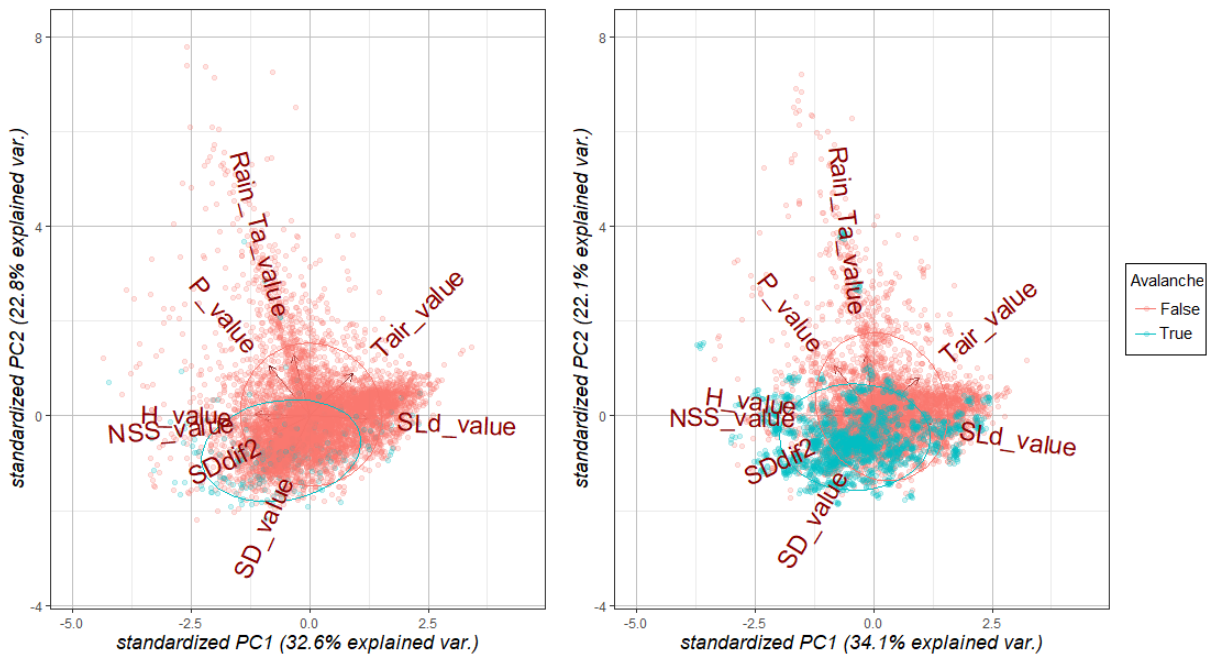

Fig. A8. Comparison of original and synthetic datasets for slab avalanches using PCA.

Table A1. Explanation of a contingency table of the confusion matrix for the final model performance fit see Fig. 9.

		Observed	
		No	**Yes**
Modelled	No	true falses	misses
	Yes	false alarms	true positives

Natural Hazards
and Earth System
Sciences
Discussions

References

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

520 Ballesteros-Cánovas, J. A., Trappmann, D., Madrigal-González, J., Eckert, N., and Stoffel, M.: Climate warming enhances snow avalanche risk in the Western Himalayas, Proceedings of the National Academy of Sciences, 115, 3410–3415, 2018.

Bellaire, S., van Herwijnen, A., Mitterer, C., and Schweizer, J.: On forecasting wet-snow avalanche activity using simulated snow cover data, Cold Regions Science and Technology, 144, 28–38, https://doi.org/10.1016/j.coldregions.2017.09.013, 2017.

Beniston, M., Farinotti, D., Andreassen, L. M., Coppola, E., Eckert, N., Fantini, A., Giacona, F., Hauck, C., Huss, M., Huwald, H.,
525 Lehning, M., López-Moreno, J.-I., Magnusson, J., Marty, C., Morán-Tejeda, E., Morin, S., Naaim, M., Provenzale, A., and Vincent, C.: The European mountain cryosphere: A review of its current state, trends, and future challenges, The Cryosphere, 12, 759–794, https://doi.org/10.5194/tc-12-759-2018, 2018.

Biecek, P. and Burzykowski, T.: Explanatory model analysis: explore, explain, and examine predictive models, CRC Press, 2021.

Bintanja, R.: The impact of Arctic warming on increased rainfall, Scientific Reports, 8, 6–11, https://doi.org/10.1038/s41598-018-34450-3,
530 2018.

Blahušiaková, A., Matoušková, M., Jenicek, M., Ledvinka, O., Kliment, Z., Podolinská, J., and Snopková, Z.: Snow and climate trends and their impact on seasonal runoff and hydrological drought types in selected mountain catchments in Central Europe, vol. 0, Taylor & Francis, https://doi.org/10.1080/02626667.2020.1784900, 2020.

Blahůt, J., Klimeš, J., Balek, J., Hájek, P., Červená, L., and Lysák, J.: Snow avalanche hazard of the Krkonoše National Park, Czech Republic,
535 Journal of Maps, 13, 86–90, https://doi.org/10.1080/17445647.2016.1262794, 2017.

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

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P.: SMOTE: synthetic minority over-sampling technique, Journal of artificial intelligence research, 16, 321–357, 2002.

540 Dkengne Sielenou, P., 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, 187, 103 276, https://doi.org/https://doi.org/10.1016/j.coldregions.2021.103276, 2021.

Dlouhý, R.: Avalanche danger in Krkonoše National Park.

Dreier, L., Harvey, S., van Herwijnen, A., and Mitterer, C.: Relating meteorological parameters to glide-snow avalanche activity, Cold
545 Regions Science and Technology, 128, 57–68, https://doi.org/10.1016/j.coldregions.2016.05.003, 2016.

Eckerstorfer, M. and Christiansen, H.: Relating meteorological variables to the natural slab avalanche regime in High Arctic Svalbard, Cold Regions Science and Technology, 69, 184–193, https://doi.org/https://doi.org/10.1016/j.coldregions.2011.08.008, international Snow Science Workshop 2010 Lake Tahoe, 2011.

Eckert, N., Keylock, C., Castebrunet, H., Lavigne, A., and Naaim, M.: Temporal trends in avalanche activity in the French Alps and subre-
550 gions: from occurrences and runout altitudes to unsteady return periods, Journal of Glaciology, 59, 93–114, 2013.

Engel, Z., Nývlt, D., Křížek, M., Treml, V., Jankovská, V., and Lisá, L.: Sedimentary evidence of landscape and climate history since the end of MIS 3 in the Krkonoše Mountains, Czech Republic, Quaternary Science Reviews, 29, 913–927, https://doi.org/10.1016/j.quascirev.2009.12.008, 2010.

Erlebach, M.: Krkonoše National Park and surroundings, Czech Republic, https://www.interreg-central.eu/Content.Node/555 MaGICLandscapes.html.

Feng, S. and Hu, Q.: Changes in winter snowfall/precipitation ratio in the contiguous United States, Journal of Geophysical Research Atmospheres, 112, 1–12, https://doi.org/10.1029/2007JD008397, 2007.

Gądek, B., Kaczka, R. J., Rączkowska, Z., Rojan, E., Casteller, A., and Bebi, P.: Snow avalanche activity in Żleb Żandarmerii in a time of climate change (Tatra Mts., Poland), Catena, 158, 201–212, 2017.

560 García-Hernández, C., Ruiz-Fernández, J., Sánchez-Posada, C., Pereira, S., Oliva, M., and Vieira, G.: Reforestation and land use change as drivers for a decrease of avalanche damage in mid-latitude mountains (NW Spain), Global and Planetary Change, 153, 35–50, 2017.

Gauthier, F., Germain, D., and Hétu, B.: Logistic models as a forecasting tool for snow avalanches in a cold maritime climate: northern Gaspésie, Québec, Canada, Natural hazards, 89, 201–232, 2017.

Germain, D., Filion, L., and Hétu, B.: Snow avalanche regime and climatic conditions in the Chic-Choc Range, eastern Canada, Climatic 565 Change, 92, 141–167, 2009.

Giacona, F., Eckert, N., Mainieri, R., Martin, B., Corona, C., Lopez-Saez, J., Monnet, J.-M., Naaim, M., and Stoffel, M.: Avalanche activity and socio-environmental changes leave strong footprints in forested landscapes: a case study in the Vosges medium-high mountain range, Annals of Glaciology, 59, 111–133, 2018.

Greene, E., Atkins, D., Birkeland, K., Elder, K., Landry, C., Lazar, B., Mccammon, I., Moore, M., Sharaf, D., Sterbenz, C., Tremper, B., and 570 Williams, K.: Snow, Weather, and Avalanches: Observational Guidelines for Avalanche Programs in the United States, 2010.

Gregorutti, B., Michel, B., and Saint-Pierre, P.: Correlation and variable importance in random forests, Statistics and Computing, 27, 659–678, 2017.

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, 2014.

575 Hock, R., Rasul, G., Adler, C., Cáceres, B., Gruber, S., Hirabayashi, Y., Jackson, M., Kääb, A., Kang, S., Kutuzov, S., et al.: High mountain areas, 2019.

Horton, S., Towell, M., and Haegeli, P.: Examining the operational use of avalanche problems with decision trees and model-generated weather and snowpack variables, Natural Hazards and Earth System Sciences, 20, 3551–3576, 2020.

Hynčica, M. and Huth, R.: Long-term changes in precipitation phase in Czechia, Geografie-Sbornik CGS, 124, 41–55, 580 https://doi.org/10.37040/geografie2019124010041, 2019.

James, G., Witten, D., Hastie, T., and Tibshirani, R.: An introduction to statistical learning, vol. 112, Springer, 2013.

Juras, R., Pavlásek, J., Blahůt, J., Bašta, P., Klose, Z., Moudrý, V., and Balek, J.: https://doi.org/10.13140/2.1.1591.6483.

Juras, R., Blöcher, J. R., Jenicek, M., Hotovy, O., and Markonis, Y.: What affects the hydrological response of rain-on-snow events in low-altitude mountain ranges in Central Europe?, Journal of Hydrology, p. 127002, 2021.

585 Kliegrová, S. and Kašičková, L.: Changes in air temperature and precipitation totals in the period 1961–2016 in the Giant Mountains, in: Meteorological Bulletin, vol. 72, p. 11, 2019.

Laute, K. and Beylich, A. A.: Potential effects of climate change on future snow avalanche activity in western Norway deduced from meteorological data, Geografiska Annaler: Series A, Physical Geography, 100, 163–184, 2018.

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, 2015.

Liaw, A. and Wiener, M.: Classification and Regression by randomForest, R News, 2, 18–22, https://CRAN.R-project.org/doc/Rnews/, 2002.

Loy, A., Follett, L., and Hofmann, H.: Variations of Q–Q plots: the power of our eyes!, The American Statistician, 70, 202–214, 2016.

Lunardon, N., Menardi, G., and Torelli, N.: ROSE: a Package for Binary Imbalanced Learning, R Journal, 6, 82–92, 2014.

Maćkiewicz, A. and Ratajczak, W.: Principal components analysis (PCA), Computers & Geosciences, 19, 303–342, 1993.

Mair, R. and Nairz, P.: Lawine, Die, 10, 10–11, 2011.

Marienthal, A., Hendrikx, J., Birkeland, K., and Irvine, K. M.: Meteorological variables to aid forecasting deep slab avalanches on persis-
tent weak layers, Cold Regions Science and Technology, 120, 227–236, https://doi.org/https://doi.org/10.1016/j.coldregions.2015.08.007,
2015.

Martin, E., Giraud, G., Lejeune, Y., and Boudart, G.: Impact of a climate change on avalanche hazard, Annals of Glaciology, 32, 163–167,
2001.

Marty, C., Tilg, A. M., and Jonas, T.: Recent Evidence of Large-Scale Receding Snow Water Equivalents in the European Alps, Journal of
Hydrometeorology, 18, https://doi.org/10.1175/JHM-D-16-0188.1, 2017.

McGregor, G. R.: Snowpack structure and avalanching, Craigieburn Range, New Zealand, New Zealand Journal of Geology and Geophysics,
33, 405–417, 1990.

Möhle, S., Bründl, M., and Beierle, C.: Modeling a system for decision support in snow avalanche warning using balanced random forest
and weighted random forest, in: International Conference on Artificial Intelligence: Methodology, Systems, and Applications, pp. 80–91,
Springer, 2014.

Murtagh, F. and Legendre, P.: Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?, Journal
of classification, 31, 274–295, 2014.

Naaim, M., Eckert, N., Giraud, G., Faug, T., Chambon, G., Naaim-Bouvet, F., and Richard, D.: Impact of climate warming on avalanche
activity in French Alps and increase of proportion of wet snow avalanches, HOUILLE BLANCHE-REVUE INTERNATIONALE DE L
EAU, pp. 12–20, 2016.

Nedelcev, O. and Jeníček, M.: Trends in seasonal snowpack and their relation to climate variables in mountain catchments in Czechia,
Hydrological Sciences Journal, 2021.

Notarnicola, C.: Hotspots of snow cover changes in global mountain regions over 2000–2018, Remote Sensing of Environment, 243, 111 781,
2020.

Peitzsch, E. H., Hendrikx, J., Fagre, D. B., and Reardon, B.: Examining spring wet slab and glide avalanche occurrence along the Going-to-
the-Sun Road corridor, Glacier National Park, Montana, USA, Cold regions science and technology, 78, 73–81, 2012.

Peitzsch, E. H., Pederson, G. T., Birkeland, K. W., Hendrikx, J., and Fagre, D. B.: Climate drivers of large magnitude snow avalanche years
in the US northern Rocky Mountains, Scientific reports, 11, 1–13, 2021.

Probst, P., Wright, M. N., and Boulesteix, A.-L.: Hyperparameters and tuning strategies for random forest, Wiley Interdisciplinary Reviews:
Data Mining and Knowledge Discovery, 9, e1301, 2019.

Quervain, M., De Crecy, L., LaChapelle, E., Losev, K., and Shoda, M.: Avalanche classification, Hydrological Sciences Bulletin, 18, 391–
402, 1973.

Reuter, B., Mitterer, C., and Bellaire, S.: 7.3 WARNSIGNAL KLIMA: Die Lawinengefahr im Klimawandel, 2020.

Schattan, P., Baroni, G., Oswald, S. E., Schöber, J., Fey, C., Kormann, C., Huttenlau, M., and Achleitner, S.: Continuous monitoring of snowpack dynamics in alpine terrain by aboveground neutron sensing, Water Resources Research, 53, 3615–3634, 2017.

Schweizer, J., Bruce Jamieson, J., and Schneebeli, M.: Snow avalanche formation, Reviews of Geophysics, 41, 2003.

Schweizer, J., Mitterer, C., and Stoffel, L.: On forecasting large and infrequent snow avalanches, Cold Regions Science and Technology, 59, 234–241, 2009.

Soteres, R. L., Pedraza, J., and Carrasco, R. M.: Snow avalanche susceptibility of the Circo de Gredos (Iberian Central System, Spain), Journal of Maps, 16, 155–165, 2020.

Spusta, V. and Kociánová, M.: Avalanche cadastre in the Czech part of the Krkonoše Mts. (Giant Mts.) during winter seasons 1961/62 – 1997/98, Opera Corcontica, 35, 3–205, 1998.

Spusta, V., Spusta, V. J., and Kociánová, M.: Lavinový katastr a zimní situace na hřebenu české části Krkonoš v období 1998 / 99 – 2002 / 03. Avalanche Cadastre and Winter Condition in Summit Area of the Giant Mts . (Czech part) during 1998 / 1999 – 2002 / 2003, Opera Corcontica, 40, 5–86, 2003.

Spusta, V., Spusta, V. J., and Kociánová, M.: Lavinový katastr české části Krkonoš v zimním období 2003 / 04 až 2005 / 06 Avalanche cadaster of the Czech part of the Giant Mountains in winter season 2003 / 04 – 2005 / 06, pp. 81–93, 2006.

Spusta, V., Spusta, V. J., and Kociánová, M.: Lavinový katastr a zimní situace v české části Krkonoš Avalanche cadastre and winter conditions in the czech part of the Krkonoše Mts . in the winter seasons, Opera Corcontica, 56, 21–110, 2020.

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, 2018.

Strapazzon, G., Schweizer, J., Chiambretti, I., Brodmann Maeder, M., Brugger, H., and Zafren, K.: Effects of Climate Change on Avalanche Accidents and Survival, Frontiers in Physiology, 12, 450, https://doi.org/10.3389/fphys.2021.639433, 2021.

Stull, R.: Wet-bulb temperature from relative humidity and air temperature, Journal of applied meteorology and climatology, 50, 2267–2269, 2011.

Štursa, J., Jeník, J., and Váňa, J.: The alpine forest-limit in the Giant Mts (Central Europe) and Abisko Mts (subarctic Sweden), Opera Corcontica, 47, 129–164, 2010.

Techel, F., Zweifel, B., and Winkler, K.: Analysis of avalanche risk factors in backcountry terrain based on usage frequency and accident data in Switzerland, Natural Hazards and Earth System Sciences, 15, 1985–1997, 2015.

Techel, F., Jarry, F., Kronthaler, G., Mitterer, S., Nairz, P., Pavsek, M., Valt, M., and Darms, G.: Avalanche fatalities in the European Alps: Long-term trends and statistics, Geographica Helvetica, 71, 147–159, https://doi.org/10.5194/gh-71-147-2016, 2016.

Therneau, T. and Atkinson, B.: rpart: Recursive Partitioning and Regression Trees, https://CRAN.R-project.org/package=rpart, r package version 4.1-15, 2019.

Tolasz, R. e. a.: Atlas podnebí Česka–Climate atlas of Czechia, International series of monographs on physics, Českỳ hydrometeorologickỳ ústav–Univerzita Palackého v Olomouci, 12007.

Vrba, M. and Spusta, V.: Avalanche Survey and Map of Krkonoše Mountains, Opera Corcontica, 12, 65–90, 1975.

Vrba, M. and Spusta, V.: The avalanche cadastre of the Krkonoše Mountains, Opera Concordica, 28, 47–58, 1991.