



**Escola de Camins**  
Escola Tècnica Superior d'Enginyeria de Camins, Canals i Ports  
UPC BARCELONATECH

COMPARISON OF DATA-DRIVEN MODELS FOR AVALANCHE  
SUSCEPTIBILITY ASSESSMENT IN ANDORRA

Final thesis developed by:

Georgia Manou

Supervisors:

Vicente Cesar De Medina Iglesias

Allen Bateman Pinzón

Master of Science in *Flood Risk Management*



Barcelona, September 2023

## ABSTRACT

Snow avalanches are defined as a mass of snow that rapidly flows down a sloping surface, such as a hill or a mountain. Depending on their velocity and volume they can have highly destructive force and pose major threat to people, property, infrastructure, and ecosystems. Andorra is a small, mountainous country located in the Axial Zone of the central Pyrenees mountain range in southwestern Europe, enclosed by France and Spain. It is exposed to several natural hazards including snow avalanches. The latter occur in the country several times per year and the popularity of winter tourism in the area results in high risk levels. Due to the frequency of the events, Andorra maintains a long avalanche inventory since the 1980's. Although the authorities have developed advanced regulations in terms of managing risk, have put in place a series of monitoring systems, as well as protection measures for avalanches, currently, there is not a forecasting tool available, and the issued warnings are based on weather observations and experience.

The aim of this master thesis has been to evaluate the suitability of data-driven models to assess susceptibility and map initiation areas of snow avalanches in Andorra. This aim was formulated as a classification task and the machine learning algorithms that were brought to the test included the Decision Tree, Random Forest, Adaptive Boosting, Gradient Boosting, Extreme Gradient Boosting, Logistic Regression, Support Vector Classifier, and Neural Network. The models received as input topographical information of the study area, derived by GIS analysis of a 5x5 m DEM and a land cover map, interpolated weather station data, snow cover information from Landsat 8 and the susceptibility to shallow landslides acquired by Shalstab. The obtained results were satisfactory, with achieved accuracies of 80% or higher for all the models, even when terrain features were the only predictors used. In a second model development cycle, when the feature space expanded to include weather, snow cover and Shalstab's susceptibility in order to test the contribution of the additional input to the performance, the results were even better in most cases. More specifically, the boosting models were able to reach accuracies of 93-95%.

Lastly, although not the best performing model, the Decision Tree, with an overall accuracy of 82%, was selected for susceptibility mapping due to its simplicity, and the interpretability of the results that it allows. The produced map depicts areas susceptible to avalanching with a probability of 88% and it enables the identification of exposed locations, and, especially, parts of the road network that require close monitoring or allocation of resources for protection and maintenance.

In conclusion, the current study demonstrated that machine learning models have good potential for predicting snow avalanches, and further research needs to be focused on this direction.

Boosting models, in particular, can be a robust forecasting tool for risk managers, ski resort operators and the civil protection in the study area. Also, the significance of the terrain for avalanche initiation was confirmed. At the same time, the importance of exploring the feature space when building data-driven models was highlighted, with the revealed, major contribution of snow cover and Shalstab's susceptibility to the performance. Based on the results, adopting and adjusting physically-based models developed for shallow landslide susceptibility seems promising for describing snow avalanche formation, while the available remote sensing products prove to be powerful sources of information in areas with a sparse observation network.

Keywords: snow avalanches, initiation, susceptibility mapping, data-driven models, machine learning, forecasting, Shalstab, snow cover

## ACKNOWLEDGEMENTS

A special thanks is owed to the supervisors of this thesis, Vicente and Allen, that have been collaborators and mentors throughout the whole semester. Additionally, to Sergi Riba from Nivorisk and Guillem from the Meteorological Service of Andorra for their insights.

Thank you, Vicente, for trusting in me and for enabling me, first of all, with your topic proposal to live and experience the beautiful city of Barcelona. It was an amazing opportunity to meet new people, learn a bit more about Catalunya, enjoy its nature and be in a country that feels like home away from home. Thank you for your interest, your support and encouragement to me for exploring new fields of study, being inventive, and learning Spanish. You have provided the tools and the space to learn whatever was of interest to me, and I am glad to say that you have been one of the best teachers I ever had.

Allen, thank you for your ideas, critical comments and empathy. You have been an inspiration as a scientist, an engineer, and someone that, as far as I can tell, leads life according to their dreams and values. I am happy that you like rivers so much and you have taught me some of the most interesting things about them.

My friend Andres, I could not be happier that you have been literally next to me from the very first until the last day of this journey. It has been very special to me having you close and sharing life with you. Thank you for all the support and help on both academic and personal level.

Cristiane, without our brainstorming sessions team projects would have been a lot more dull. You have been there for me in every way possible, sharing with me your time, your knowledge, your music, your culture, and many more. Thank you for your help during my first steps with programming and always being available with tips, advice and nice words throughout this semester.

I have been lucky to have family and friends that have been pillars of constant support, safety and happiness in my life, and these two years were no exception. Writing these last lines of the document I feel grateful for the experiences I was exposed to and the empowerment I received as a side effect of this master's, the people I have met and their stories. It has been a long journey that opened my eyes a bit wider and made me realize my luck and potential. More than an opportunity to specialize and gain knowledge, a study program like this is a privilege and cradle of freedom and equality, which should not be taken for granted.

## CONTENTS

List of Figures.....	vii
List of Tables .....	xi
List of abbreviations.....	xii
1. Introduction.....	1
1.1. Background .....	1
1.2. Problem statement and research questions .....	4
1.3. Study area .....	5
1.4. Motivation and innovation of the research .....	9
1.5. Formulation of the research aim and objectives.....	10
1.6. Practical value .....	11
1.7. Thesis structure .....	11
2. Literature Review.....	13
2.1. Key concepts, theories, and studies.....	13
2.2. State of the art: Snow avalanche initiation.....	15
2.3. Gaps in the existing knowledge .....	17
2.4. Studies on snow avalanches in Andorra .....	17
3. Research methodology.....	19
3.1. Methodological framework.....	19
3.2. Phase 1: Database construction .....	21
3.3. Phase 2: Development of data-driven models.....	31
3.3.1. Data preprocessing .....	31
3.3.2. Model selection .....	33
3.3.3. Training-validation-test split vs. Cross validation .....	38
3.3.4. “Training-validation-optimization” cycle .....	39
3.3.5. Evaluation metrics .....	40

3.3.6. Deployment of the models.....	42
3.4. Phase 3: Model testing and susceptibility mapping.....	42
4. Case study: Andorra .....	43
4.1. Phase 1 .....	43
4.1.1. Data availability .....	43
4.1.2. Data analysis .....	48
4.2. Phase 2.....	63
4.2.1. Database preprocessing .....	63
4.2.2. Model building and deployment.....	67
5. Results and discussion.....	80
5.1. Testing and evaluation of the models .....	80
5.2. Discussion .....	81
5.3. Susceptibility mapping.....	82
6. Conclusions and recommendations .....	85
6.1. Conclusions.....	85
6.2. Recommendations.....	87
References .....	89

## LIST OF FIGURES

Figure 1. Types of snow avalanches (Glossary – EAWS, n.d.-a).....	3
Figure 2. Location of Andorra.....	6
Figure 3. Graph of long-time-average monthly precipitation, minimum, average, and maximum temperature in Andorra. ....	7
Figure 4. Avalanche protection measures installed in Andorran mountains. ....	9
Figure 5. Methodological framework of the study.....	19
Figure 6. Phase 1 of the methodology: Processes followed for the construction of the database. ....	21
Figure 7. Combinations of planform and profile curvature (Curvature Function—ArcMap   Documentation, n.d.). .....	25
Figure 8. Sun orientation in Arinsal, Andorra, on the 1st of December 2022 (map on the left) and on the 30th of April 2023 (map on the right) (Arinsal: Sun Direction, n.d.).....	26
Figure 9. Structure in the shape of a tree, based on which the so-called tree-based machine learning algorithms are performing classification tasks.....	34
Figure 10. Typical structure of an MLP with 2 input nodes, 4 hidden nodes and 1 output node.	38
Figure 11. Representation of the different sample folds in a 5-fold cross validation.....	39
Figure 12. The cycle of "training-validation-optimization" followed by model deployment and testing when the desirable accuracy has been achieved.....	40
Figure 13. The confusion matrix. ....	41
Figure 14. Map of historical avalanches detected by direct testimony (purple color) or photointerpretation (yellow color). .....	44
Figure 15. Map of the weather stations that were considered in the study.....	45
Figure 16. Digital Elevation Model of Andorra. ....	47
Figure 17. Land Cover map of Andorra.....	47
Figure 18. Map of forest types distinguished as coniferous or deciduous. ....	48

Figure 19. Initiation and non-initiation avalanche zones depicted with orange circles of different diameters.....	49
Figure 20. Map of slope angles.....	50
Figure 21. Map of slope aspects.....	50
Figure 22. Map of surface roughness characterized by rocky terrain and scree-laden slopes.....	51
Figure 23. Map of snow cover in Andorra depicted as NDSI values.....	52
Figure 24. Map of critical recharge with regard to shallow landslide susceptibility acquired as output of the Shalstab model.....	53
Figure 25. Example of invalid measurements of rainfall in the Borda Vidal station.....	54
Figure 26. Example of outlier detection of the minimum temperature in the station of Soldeu using the histogram.....	54
Figure 27. Example of outlier detection of maximum temperature in the station of Pal using the boxplot.....	54
Figure 28. The double mass curve method applied to the winter stations' observations. The y-axis represents the cumulative precipitation of the individual station (in mm) and the x-axis denotes the cumulative precipitation of all the stations (in mm).....	56
Figure 29. The double mass curve method applied to the winter stations' observations. The y-axis represents the cumulative old snow of the individual station (in cm) and the x-axis denotes the cumulative old snow of all the stations (in cm).....	56
Figure 30. Scatterplot of precipitation in the La Comella station. The red horizontal line represents the 95% threshold for identifying the extremes.....	57
Figure 31. Map of the prevalent wind conditions (speed and directions) represented as wind roses in different parts of Andorra.....	58
Figure 32. Conditional distribution of snow drift intensity given wind direction in the station of Arcalis.....	59
Figure 33. Correlation of maximum temperature and range of daily temperatures in the station of Pal.....	59
Figure 34. Interpolated minimum temperature.....	61
Figure 35. Interpolated rainfall.....	61

Figure 36. Heatmap depicting correlations between the independent features.....	64
Figure 37. Histogram representation of the distribution of elevations in relation to susceptibility to avalanche.....	65
Figure 38. Histogram representation of the binary feature of east-oriented slopes in relation to susceptibility to avalanche.....	65
Figure 39. Graph of mutual information of the features with the target class.....	66
Figure 40. Count plot of the different wind directions in the database.....	66
Figure 41. Histogram representation of the binary feature of south wind in relation to susceptibility to avalanche.....	67
Figure 42. Initial fit of the Decision Tree.....	67
Figure 43. Accuracy of both the training and validation sets in relation to the alpha pruning parameter.....	68
Figure 44. Confusion matrix of the pruned DT with $a = 0.014$ on the validation set.....	69
Figure 45. Plot of the pruned DT with $a = 0.014$ .....	69
Figure 46. Confusion matrix of the pruned DT with $\alpha = 0.017$ on the validation set.....	70
Figure 47. Optimization of hyperparameters of the Random Forest using the "RandomizedSearchCV" class of sklearn.....	71
Figure 48. Training of the optimized Random Forest. ....	71
Figure 49. Confusion matrix of the optimized Random Forest on the validation set.....	71
Figure 50. Bar plot depicting the importance of features for the Random Forest classifier. ....	72
Figure 51. Training of the optimized AdaBoost classifier. ....	72
Figure 52. Confusion matrix of the optimized AdaBoost model on the validation set. ....	73
Figure 53. Optimization of hyperparameters of the Gradient Boost algorithm using "GridSearchCV".....	73
Figure 54. Training of the optimized Gradient Boost model.....	74
Figure 55. Confusion matrix of the optimized Gradient Boost on the validation set.....	74
Figure 56. Graph of increasing accuracy with increasing number of estimators for the Gradient Boost model.....	75

Figure 57. Feature importance graph of the Gradient Boost model.....	75
Figure 58. Training of the XGBoost model with the optimized hyperparameters.....	76
Figure 59. Confusion matrix of the optimized XGBoost on the validation set.....	76
Figure 60. Training of the Logistic Regression model using the optimized hyperparameters.....	77
Figure 61. Confusion matrix of the optimized Logistic Regression on the validation set.....	77
Figure 62. Training of the SVC with the optimized hyperparameters.....	78
Figure 63. Confusion matrix of the SVC before optimization.....	78
Figure 64. Confusion matrix of the SVC after optimization.....	78
Figure 65. Confusion matrix of the feedforward neural network.....	79
Figure 66. Plot of the optimized DT when fitted to the entire database. ....	83
Figure 67. Decision rules of the optimized DT fitted to the entire database. ....	83

## LIST OF TABLES

Table 1. Reclassification categories of the slope aspects.....	25
Table 2. Data availability and record length of the weather stations considered in the analysis... ..	46
Table 3. Representative values selected for each weather variable to be used for their interpolation. .....	60
Table 4. Zonal statistics used for the creation of the database. ....	62
Table 5. Contingency table of wind direction and potential snow transport.....	67
Table 6. Test error of both versions of models. The first version was developed using only topographical features and the second using the entire database. ....	80

## LIST OF ABBREVIATIONS

AUC	Area Under Curve
CENMA	Centre d'Estudis de la Neu i de la Muntanya d'Andorra
COEX	Conservació i Explotació de Carreteres
DEM	Digital Elevation Model
DT	Decision Tree
EAWS	European Avalanche Warning Services
FPR	False Positive Rate
GBDTs	Gradient Boosted Decision Trees
GLM	Generalized Linear Model
IDW	Inverse Distance Weighting
MLP	Multi-Layer Perceptron
NASA	National Aeronautics and Space Administration
NDSI	Normalized Difference Snow Index
OLI	Operational Land Imager
RBF	Radial Basis Function
ROC	Receiver Operating Characteristic
SIGMA	Sistema d'Informació Geogràfica Mediambiental d'Andorra
SVC	Support Vector Classifier
SVM	Support Vector Machine
TPR	True Positive Rate
USGS	United States Geological Survey
VS	Voellmy-Salm

## 1. INTRODUCTION

---

### 1.1. Background

Snow avalanches are rapid flows of snow masses moving down the slope of a mountain, and they constitute a natural hazard that prevails in numerous mountainous regions throughout the world, posing major threat to human life, property, infrastructure, and ecosystems (Hebertson & Jenkins, 2003; Košová et al., 2022; Kumar et al., 2019; Schweizer et al., 2003, 2016; Součková et al., 2022). Avalanche hazard refers to the probability of an avalanche event to take place in a reference period and its potential destructive consequences for people, the environment, and several industries and sectors like recreation, transportation, and energy (Stethem et al., 2003). Avalanche risk is related to the number of people and length of time in an exposed area (D. McClung & Schaerer, 1993).

Techel et al. (2016) reported a drastic increase in avalanche casualties in the European Alps at the second half of the twentieth century, when winter tourism in the region experienced an unprecedented growth, which is indicative of the role that exposure plays in the formation of avalanche risk. If we consider the development of infrastructure in terms of ski resorts and road networks that are needed to support the incoming visitors in popular, winter destinations, it is highlighted how avalanche vulnerability increases in those areas (Munteanu et al., 2013).

The following terms and definitions are essential for understanding the main concepts described in this section and are borrowed from D. McClung & Schaerer (1993):

- “An *avalanche path* is a fixed locality within which avalanches move.”.
- “An *avalanche area* is a location with one or more avalanche paths.”.
- “The *starting zone (or zone of origin)* is the location where the unstable snow failed and began to move.”. The term initiation or release area/zone is another synonym of starting zone.
- “The zone where snow debris collects is called *runout zone*.”. The term deposition area/zone can be used interchangeably to runout zone.
- “The *track (or zone of transition)* is the slope below the starting zone that connects it with the runout zone.”.

Snow avalanches are differentiated based on certain criteria, such as the release mechanism (loose-snow avalanche / slab avalanche), the form of movement (dense flow avalanche / powder avalanche), and the liquid water content in the snow (dry-snow avalanche / wet-snow avalanche)

(*Glossary – EAWS*, n.d.-a). Moreover, gliding snow avalanches are often presented as a distinct type, where the whole snowpack slides on a flat surface, most commonly on a film of water on the ground, causing casualties very rarely (*Glossary – EAWS*, n.d.-a; *Recognize and Distinguish Types of Avalanche | LAB SNOW*, n.d.). Figure 1 depicts the types of avalanches mentioned above.

Another distinction is often made with respect to the triggering mechanism, which, in principle, is near-surface, usually rapid loading of the snowpack, either natural or artificial. A natural trigger can be, for instance, precipitation (snow or rain) or an abrupt, positive temperature gradient, while winter athletes/recreationists, hikers, explosives, etc. can provoke human-induced releases (Stethem et al., 2003). Almost all human-triggered avalanches are slab avalanches. In addition, those are hardly ever wet-snow avalanches (Schweizer, 2002). On the contrary, glide snow avalanches cannot be triggered artificially (*Recognize and Distinguish Types of Avalanche | LAB SNOW*, n.d.).

The two main avalanche types are loose-snow and slab avalanches (Kumar et al., 2019; D. McClung & Schaefer, 1993; Schweizer, 2002). They can occur in both dry and wet snow conditions (Stethem et al., 2003), and it is possible to release either spontaneously or because of human presence or interference. Loose-snow avalanches have a point-shaped fracture, a tongue-shaped or pear-shaped spreading pattern (*Recognize and Distinguish Types of Avalanche | LAB SNOW*, n.d.; Stethem et al., 2003), and very few fatalities are caused by them (Schweizer, 2002). On the other hand, slab avalanches are the most usual and dangerous, with 98% of all accidents being caused by this type of avalanches. They are characterized by a cohesive slab of snow overlying a weak layer, and they have a linear fracture, forming a characteristic crown after their release (*Recognize and Distinguish Types of Avalanche | LAB SNOW*, n.d.; Schweizer et al., 2003; Stethem et al., 2003), with different density and hardness, which are key snowpack properties associated with fracture initiation and propagation propensity (Van Herwijnen & Jamieson, 2007).



SLAB AVALANCHE



LOOSE SNOW AVALANCHE



FLOW AVALANCHE



POWDER AVALANCHE



WET SNOW AVALANCHE



GLIDE-SNOW AVALANCHE

Figure 1. Types of snow avalanches (Glossary – EAWS, n.d.-a).

What causes an avalanche to release is the interplay of various coupled factors that are taking place over several orders of scale, from the microscopic (grain sizes, shapes, and bonds), to the macroscopic (e.g., a weak layer), and, finally, extending to the scale of a mountain slope (Schweizer et al., 2003). The contributory factors can be split into three main categories, i.e., terrain, meteorological conditions, snowpack conditions, and an additional parameter related to the triggering mechanism (Bühler et al., 2013; D. McClung & Schaerer, 1993; Schweizer et al., 2003).

According to Boyd et al. (2009), in the 30-year period of 1978-2007 the recorded fatalities due to avalanches in Canada had been 329. Based on data provided by the Swiss Institute for Snow and Avalanche Research, every year since the winter of 1936/1937 the average number of avalanche

victims in Switzerland has been 24 (*Long-Term Statistics - SLF*, n.d.), while the European Avalanche Warning Services (EAWS) report in their website that every year the loss of lives due to snow avalanches averages 100 people (*Fatalities – EAWS*, n.d.). A rough estimate of 150 winter sport enthusiasts (e.g., skiers, snowboarders, mountaineers) are killed every year in Europe and North America, making snow avalanches the biggest threat to winter, backcountry recreationists (Schweizer, 2002). The winter season 2014/2015 132 fatalities due to snow avalanches happened only in the European Alps, which “served as a cruel reminder of the destructive power of this ubiquitous natural hazard”, as Gaume et al. (2017) described it.

The above-mentioned statistics showcase the severity of the effects of avalanches. For this reason, the common goal of all EAWS is to provide the public and the local stakeholders with accurate warnings in the form of avalanche bulletins (*Fatalities – EAWS*, n.d.; Schweizer et al., 2016). The same is the principle worldwide, where numerous local and regional avalanche centers provide avalanche forecasts in the form of bulletins (*Forecasts - BRASS Avalanche*, n.d.). The avalanche bulletins by the EAWS have the same structure and content, with the most important information being the danger level and the avalanche prone locations based on aspects and elevations (*Information Pyramid – EAWS*, n.d.).

Aside from being useful to the general public, the knowledge of areas that can potentially be susceptible to avalanching enables hazard mapping of a certain site, which is important for land-use planners in mountainous areas with high population density, large amounts of incoming tourists and strong economic interests; additionally, with significant losses in case of interruption of businesses (e.g., ski resorts) and/or damage of building infrastructure and main road axes that provide access to the respective areas (Barbolini et al., 2011; Ghinoi & Chung, 2005).

## 1.2. Problem statement and research questions

Effective avalanche forecasting means ability to define areas susceptible to current or future snow instability in space and time (Ghinoi & Chung, 2005; D. M. McClung, 2000), which has proven to be a difficult task due to our difficulty to understand the processes involved in avalanche release (Schweizer et al., 2016). Although, nowadays, avalanche centers can provide reliable warnings based on occurrence probabilities, the spatiotemporal forecasting of avalanche activity has not yet been achieved on an operational level (Schweizer et al., 2003). What hinders this effort is the complexity arising from the nature of snow itself, which is a highly porous and compressible material that exists close to its melting point and presents a highly variable structure and texture

(Schweizer et al., 2003, 2016). These snow properties are closely interrelated to the mechanical behavior of snow, making it highly rate-dependent and unable to maintain its structural integrity when fracture occurs (Schweizer et al., 2016).

Avalanches from the perspective of flow dynamics have been calculated using the Voellmy-Salm (VS) model (Salm et al., 1990) for many years now (Christen et al., 2002). Since the 1990's several numerical models have been developed (one-dimensional, two-dimensional, even three-dimensional), which are able to represent the phenomenon of avalanche propagation quite well in some cases (Christen et al., 2002, 2012; Granig et al., 2016). Most of these models, like, for example, RAMMS (SLF WSL, 2022), which is probably the most popular among researchers and practitioners, are implementing the VS model, and, in some instances, might have incorporated some extra equations (Christen et al., 2010). Although the VS model is simple to use because it requires only few parameters, and for simple applications it has been proved quite accurate (Christen et al., 2002, 2010), a factor that can majorly influence its results is the avalanche start zone that it requires as an input (Dillon & Hammonds, 2021).

Thereof, defining potential avalanche release areas (alternatively, susceptibility assessment) is of major importance for avalanche hazard estimation and risk management. Here, it should be noted that the terms susceptibility and hazard are borrowed from the vocabulary of geohazard mapping, where "Susceptibility refers to the probability of occurrence of an event within a selected type in a given location, whereas hazard refers to the probability of occurrence of an event within a selected type and magnitude in a given location within a reference period." (Wubalem, 2022).

The current thesis topic attempts to mainly answer the following research question:

*Do data-driven models provide a reliable option for avalanche susceptibility assessment and mapping? If yes, which model performs best on this scope?*

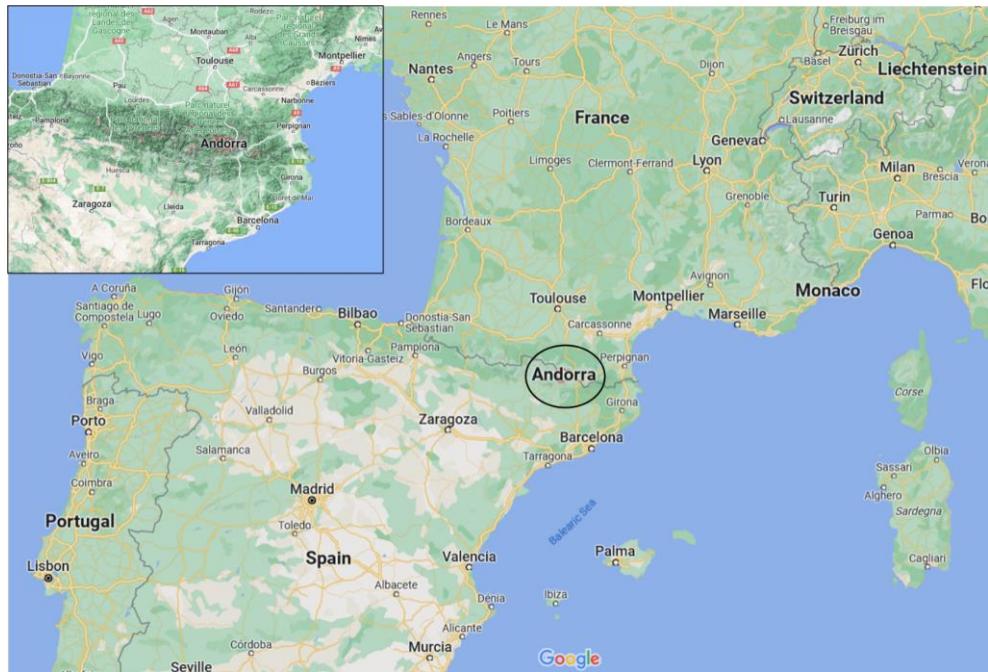
As a secondary quest, the answer to the following question is also explored:

*Which are the most important contributing factors for avalanche initiation in Andorra?*

### 1.3. Study area

Andorra, officially named Principality of Andorra, is a small, mountainous, landlocked country, located in southwestern Europe, in the central Pyrenees, between Spain and France (Figure 2). Its lowest point is at 830 m elevation and its highest at 2,942 m, while it has a total area of 468 km<sup>2</sup>

(Gilaberte Búrdalo, 2018; Margalef et al., 2018). The climate in Andorra is subdivided into three regions: in the northern slope the so-called Atlantic climate is prevailing, in the central part the climate is characterized as Transitional, and in the southern slope as Mediterranean.



*Figure 2. Location of Andorra.*

Figure 3 presents the range of the long-time-average, monthly temperature and the long-time-average, monthly precipitation depth. The latter seems to be relatively high throughout the year. The minimum temperature climbs a bit above 10°C in July and August and drops a few degrees Celsius below zero during the winter months of December through March. Nevertheless, daily temperatures drop even lower in winter, especially in higher elevations. The winter season, defined here as the months that the avalanche monitoring services in the country are on duty, lasts roughly from the beginning of November to the beginning of May. The highest parts of the country might remain covered with snow for more than 5 months (Margalef et al., 2018). There is not an avalanche season or avalanche type typically found in Andorra, since the snowfall conditions and avalanche activity are highly dependent on the amount and direction of the fronts that cross the country each year (Esteban et al., 2002, 2005).

Apart from the frequent avalanche hazard existing in the country (Riba Porras et al., 2018), vulnerability is also quite high. Although its population is just 80,000, the population density of around 170 inhabitants/km<sup>2</sup> in 2017 is considered particularly high (Gilaberte Búrdalo, 2018;

Margalef et al., 2018). In addition to that, more than 8 million winter tourists visit the country and its ski resorts each year (Gilaberte Búrdalo, 2018). In 2019, tourism contributed more than 60% to the country's GDP (*Development and Importance of Tourism for Andorra*, n.d.). Hence, a large amount of not only people but also buildings, infrastructure and kilometers of road network that support the ski industry and enable transportation and other activities are exposed to snow avalanches, posing potential large threats to the local businesses and the national economy (Margalef et al., 2018; Pons-Pons et al., 2012).

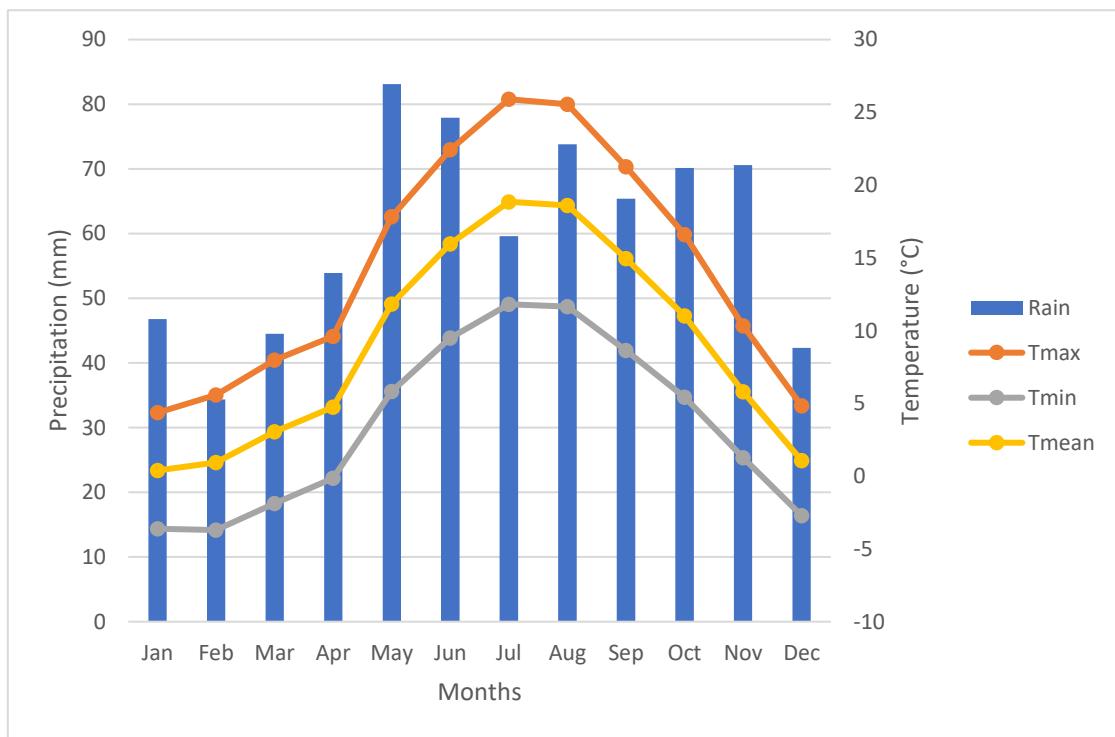


Figure 3. Graph of long-time-average monthly precipitation, minimum, average, and maximum temperature in Andorra.

Since 1971 there have been 18 fatalities due to snow avalanches in Andorra and many damages to roads, properties, and infrastructure (Margalef et al., 2018). The country's most extreme event, the pure powder, dry-snow avalanche that released in the winter of 1996 from the Les Fonts peak, propagated down the slope for 1,200 m, crossed the Arinsal valley, and flowed further up the opposite slope for about 200 m. Major role to the release and the flow regime played the strong winds from the north and north-west that resulted in overaccumulation of drifted snow in the south-east and south-west oriented slopes of the mountain in the release zone. Fortunately, no victims or injuries occurred, due to efficient evacuation, but 322 people suffered property damages and losses (Furdada et al., 2020a).

Regarding avalanche forecasting, every morning at 8:30 a bulletin is published by the Andorran Meteorological Service (in collaboration with Météo France and EAWS) considering the snow cover conditions of the day before (measured at 06:00 a.m. UTC the day of the bulletin's release), its evolution throughout the night and the available daily forecast. The bulletins provide information on the avalanche danger level in all parts of the country, usually from December until May, depending on the snow situation, and indicate the elevations and aspects that are most susceptible to avalanche release. Whenever necessary, the Civil Protection transmits warnings as well (Margalef et al., 2018). A recently launched platform, open to the public for information dissemination about the weather and avalanche risk, is the <https://allaus.ad/>. There, the viewer can find safety tips, maps with terrain characterization of all areas of Andorra where winter mountain activities can take place, and links to weather and avalanche forecasts. Also, there is the option of receiving (downloadable) information about suggested routes, including the avalanche danger of a given day.

Additionally, Nivorisk, a private company that is the current contractor of the Andorran Government for avalanche control, is responsible for monitoring and registration of avalanche activity during winter time and for issuing related reports twice a month. The procedure is that every time there is avalanche activity, members of the company go out on a helicopter mission, shoot photographic reports, and afterwards digitize them. In the past, the cadastre was carried out through surveys, old photographs, etc. The company is operating in coordination with the Andorran Service for Conservation and Exploitation of Roads (COEX) for closing down/re-opening roads or evacuating buildings whenever necessary.

In the winter they are on duty 24 hours a day, 7 days a week from November 1 to May 1, and their responsibilities include continuously monitoring the snowpack, its stability and structure evolvement by performing stability tests in the field and getting stratigraphic profiles. Often times, they perform artificial triggering with explosives under controlled conditions, so as to prevent a potential spontaneous release. They also carry out numerical modeling and hazard mapping, based on which they suggest urban planning solutions and design protection systems (Figure 4). In the summer season they focus more on maintenance and repairs of protection measures, roads, and buildings.



*Figure 4. Avalanche protection measures installed in Andorran mountains.*

#### 1.4. Motivation and innovation of the research

The source of uncertainty associated with the prediction of snow avalanches has already been underlined when presenting the various contributory factors that interplay in their formation, and the complexity of the physical and mechanical properties of the snow and their interdependencies. Furthermore, it has also been mentioned previously, that the wide range of scales (temporal and spatial) involved in the instability of the snowpack increases the complexity of avalanche formation and makes it difficult to answer crucial questions like “where”, “how often” and “under which circumstances do snow avalanches occur?” (Schweizer et al., 2003; Stoffel et al., 1998). From the above, it can also be concluded that the development of a physically-based model in order to study avalanches is complicated, requires a lot of site-specific, good quality data and local expertise (Bellaire et al., 2017; Gaume et al., 2019).

The most stable of all the elements related to the phenomenon of snow avalanches are terrain features, which do not change as rapidly as weather or snow cover characteristics might do (Bühler et al., 2013). Thus, one of the main motivations of this research has been to take advantage of these less variable components, which can be derived from a Digital Elevation Model (DEM), in order to narrow down the potential avalanche start zones in the study area. Furthermore, although the Pyrenees is a hotspot of avalanche activity in Europe, Andorra has barely been the study

interest of researchers in the field. The fact that it keeps a long avalanche inventory that covers the whole country, has a fairly dense weather station network, and observes several weather-related, and snow-related, in particular, variables, make the country a suitable case study to apply the methodology presented in this thesis. Additionally, the attempt to develop a methodology for assessing avalanche susceptibility in a previously unstudied, at least from the perspective of research, area emerges as an innovative endeavor.

Moreover, the use of machine learning tools has been limited for avalanche susceptibility assessment and mapping. In the few instances that machine learning models have been put to the test for such reason, the predictors are, usually, restricted to either topographical features or weather variables. Thereby, another innovation that the current research is introducing lies in the inclusion of not only topography, but also meteorological features, data derived from remote sensing, as well as the output of a qualitative shallow landslide model, namely Shalstab (Montgomery & Dietrich, 1994), as inputs to the models to test their contribution, if any, to the performance.

More specifically, regarding Shalstab and the idea of incorporating it in this study, it needs to be mentioned that avalanches share a lot of similarities with landslides. According to Hincks et al. (2011), they are almost synonym words, while slab avalanches, in particular, are considered a specific type of shallow landslides (Guillet et al., 2023). Both phenomena belong to the larger category of mass movements that result in slope failure, depending on various conditioning factors that they have in common, like slope angle, aspect, elevation, curvature, etc. (Akay, 2021; Choubin et al., 2020). Additionally, both types of events showcase gravity-driven flows, which can cause slope erosion along their track, as well as displacing of forest and vegetation due to their large impact pressures derived from runout velocities (Geitner et al., 2021; Guillet et al., 2023; Hincks et al., 2011). In fact, the two phenomena are often examined together when it comes to simulation methods, like in the paper of Mast et al. (2014). Also, they are usually treated in a similar manner with respect to mitigation measures (Van Zadelhoff et al., 2022). All the above-mentioned commonalities provide an important background and motive to test a model developed for landslides as predictor for avalanche initiation and adds to the innovation of the current study.

### 1.5. Formulation of the research aim and objectives

The aim of this research, emerging as a response to the main research question formulated in section 1.2, is to:

*Investigate the suitability of data-driven models for avalanche susceptibility assessment and mapping.*

In order to facilitate this aim, but also to answer the secondary research question, the following objectives are set:

- a. *Develop and test different machine learning models.*
- b. *Compare the developed models, find the best performing ones and evaluate if their performance is satisfactory for avalanche susceptibility assessment.*
- c. *Evaluate the differences in performance of the models when using different sets of predictors.*
- d. *Find the most important contributing factors to snow avalanche susceptibility in Andorra.*
- e. *Produce a snow avalanche susceptibility map of Andorra.*

## 1.6. Practical value

In recent years, an extra effort has been carried out by the responsible authorities in Andorra to identify which of the available snow dynamics models are able to reproduce recorded avalanche tracks and deposition areas, by receiving as input the weather conditions at the time, and through calibration based on avalanche inventory information (Riba Porras et al., 2018).

The knowledge of which areas are prone to release an avalanche combined with the fracture depth, determines quite certainly the volume of an avalanche (Barbolini et al., 2011; Maggioni et al., 2002). In turn, the initial volume influences the development/propagation of the avalanche, namely its runout distance and velocity, which are crucial for risk assessment, spatial planning and designing avalanche protection measures and mitigation strategies (Barbolini et al., 2011; Bühler et al., 2013; Maggioni et al., 2002).

Thereof, the product of the current study, which is the indication of potential avalanche initiation zones, can be a valuable input for the avalanche flow models developed by the monitoring services in the country, enabling better simulations, and, consequently, better-informed decisions regarding protection measures and spatial planning.

## 1.7. Thesis structure

The outline of this thesis is divided into five chapters (excluding the introductory one), and below there is a small description of the contents of each one of them.

In Chapter 2 an overview of the literature relevant to snow avalanches and, more specifically, to release mechanisms and theories is provided. Also, gaps in the existing knowledge, prior application of data-driven models to study avalanches and avalanche-related publications focused on Andorra are briefly mentioned.

In Chapter 3 the methodological framework and its main components, namely Phase 1, Phase 2, and Phase 3, developed for the realization of this project are presented. More specifically, initially, the steps followed to construct the database for feeding the data-driven models are explained (Phase 1). Then, the theoretical background of the selected machine learning algorithms, as well as the process of training-validation-optimization are described (Phase 2). Lastly, the final results (testing and comparison of the models) are analyzed, and the adoption of a method for susceptibility mapping is explained (Phase 3). In this chapter, only general descriptions of the methods and models are provided.

In Chapter 4 the specifics of the application of the methodology presented in Chapter 3 to the selected study area are demonstrated. The structure of this chapter is similar to the previous one, except for the results of Phase 3 that are showcased in the next chapter.

In Chapter 5 an evaluation of the performance of the data-driven models on “out-of-sample” data is provided. Based on these performances a comparison between the deployed models is given, and, also, the produced susceptibility map is presented.

In Chapter 6 the main findings of the study are summarized, some conclusions are drawn on whether the aims and objectives were met, and the research questions are attempted to be answered. Finally, improvements and ideas for future extension of this study are recommended.

## 2. LITERATURE REVIEW

---

### 2.1. Key concepts, theories, and studies

As already mentioned, the most fatal avalanches are slab avalanches, due to the fact that 9 times out of 10 either the victim or someone in their group trigger the avalanche (Schweizer, 2002; Schweizer et al., 2016). Also, the majority of these avalanches are dry-snow, since wet, slab avalanches usually cannot be triggered by rapid, near-surface loading similar to that of a skier (Schweizer, 2002). Thus, dry, slab avalanches are the most investigated type. On the other hand, loose-snow avalanches do not cause that many casualties, but when they do, 50% of these are wet-snow avalanches (Schweizer, 2002).

The first incidence during the release process of dry-snow, slab avalanches is failure initiation. In other words, localized damage happens at the microscale (i.e., at the level of snow grains and their bonds), which forms an initial crack in a weak layer inside the snowpack. Subsequently, this crack further expands at the macroscale (i.e., at the scale of the weak layer). Next, the crack propagates rapidly along the weak layer, but this time, the failure propagation is more generalized and occurs at the slope scale. Finally, if the slope is steep enough, the slab slides down. This sequence of failure steps that results in the release of the slab is, in general, accepted by the majority of researchers in the field (e.g., (McClung, 1979; Schweizer et al., 2003, 2016; van Herwijnen & Heierli, 2009; Van Herwijnen & Jamieson, 2007)).

For many years, the dominant and simplest approach to describe dry-snow, slab avalanches used to not distinguish between the different processes described above. It compared the strength of the material to the accumulated shear stress at a certain time, location, and depth in the snow cover (somewhere in the weak layer), and recommended a stability index (strength over stress) for slope failure assessment (D. McClung & Schaefer, 1993). Values of this index lower or close to 1 indicated unstable conditions (Schweizer et al., 2003, 2016). However, as Schweizer et al. (2003) emphasized, strength and stress within the snowpack are scale-dependent, and, thus, the crack propagation process should be considered separately.

A more modern approach is proposing differentiating the two initial stages, failure initiation and crack propagation. According to Schweizer et al. (2016), the first stage of a dry, slab avalanche release is, indeed, best described by a “strength over stress” approach, while the second stage, which includes the crack propagation process, is better assessed following an energy criterion.

According to the latter, if the amount of energy released by the fracture is greater than the energy consumption, only then fracture propagation will occur. A critical crack length is defined, derived from the balance of the previously mentioned energies. When the crack grows equal or bigger to this critical length, extensive, shear crack propagation is initiated, which leads to slope instability. In this case, the failure criterion is the critical crack length (Schweizer et al., 2016; Van Herwijnen & Jamieson, 2007).

In 2008, another theory emerged, complementing the one described in the previous paragraph by considering slab release due to mixed-mode (mode I/II) anticracks, propagating under both shear and compression, as indicated by fracture mechanics (Gleason et al., 2008; Heierli et al., 2008). This theory highlighted that, although weak layers present less resistance in shear than in compression due to their often highly anisotropic microstructure, completely neglecting the compression component of the gravitational force in snow is the reason why the above-mentioned shear crack models fall short in explaining phenomena like “whumpfs” (i.e., whumping sounds resulting from the collapse of a weak layer, indicating clear avalanche alerts), which propagate in flat terrain and remotely triggered avalanches (Gleason et al., 2008; *Glossary – EAWS*, n.d.-b; Heierli et al., 2008; Schweizer et al., 2016). Basically, this theory emphasizes how uncertain the stress distribution at the scale of the bonds is and underlines that at the microscale any failure mode seems possible, including a vertical collapse or displacement (Schweizer & Jamieson, 2008).

Aiming to unravel the complex interrelations of weather phenomena and the release of snow avalanches, mostly dry, slab avalanches, many studies have been published. For instance, McClung (1996) investigated how temperature affects fracture and slab stability, while Schweizer & Camponovo (2002) tested experimentally the temperature dependence of the effective elastic properties of snow. Other studies focused on the effects of surface warming on the snowpack, and, more specifically, on properties like stiffness and propensity to fracture (Reuter & Schweizer, 2012; Wilson et al., 1999). Bakermans & Jamieson (2009) took the above research question a step further and proposed a method that used field data of daytime warming in order to develop a semi-empirical warning model. Other studies have addressed snow drift, the impact of winds on avalanche activity, and have proposed sensors for warning (Fohn, 1980; Lehning et al., 2002; Meister, 1989; Michaux et al., 2002). Schweizer et al. (2003) provided a comprehensive summary of all the factors and their specifics that contribute to avalanche formation, not only temperature and wind, but also new snow, terrain features and many more.

Based on the literature review performed under this study, loose-snow avalanches have not been researched as much. And the same applies for wet-snow avalanches, either loose-snow or slab. The study of Baggi & Schweizer (2009) showed that snow depth, precipitation and air temperature had the highest correlation with wet-snow avalanche activity of both types in the Eastern Swiss Alps.

Regarding the flow regime of avalanches, the resulting erosion and impact pressures, velocities in the track zone, run-out distances, and deposition areas, several studies and modeling attempts can be found in literature, for both dense-snow and powder avalanches (Bartelt et al., 1999; J. Chen et al., 2022; Christen et al., 2007, 2010; De Haas et al., 2019; Dillon & Hammonds, 2021; Hopfinger, 1983; Hopfinger & Tochon-Danguy, 1977; Issler, 1998; Košová et al., 2022; Li et al., 2021; Naaim et al., 1996, 2004; Naaim-Bouvet et al., 2004; Sanz-Ramos et al., 2020; Sovilla et al., 2008, 2015; Zarrini, 2016).

## 2.2. State of the art: Snow avalanche initiation

For the purpose of snow avalanche release prediction and susceptibility mapping, which is the focus of this thesis, there is a variety of techniques that have been developed and several studies ranging from conceptual methods and GIS-based approaches to mathematical models and numerical algorithms, analytical methods, and, lastly, machine learning to a smaller extent. A few characteristic examples are presented in the following paragraphs.

In the paper of Maggioni et al., 2002), the idea of delineating potential release areas (PRA) based on specific terrain features derived from a DEM was presented for the first time. For the development of the methodology, the authors used a historical record of avalanche activity in Davos, Switzerland. A bit more than a decade later, Bühler et al. (2013) refined the method by using an improved algorithm that included high spatial resolution optical remote sensing imagery to gather information on forest land and very rugged rock surfaces, so as to completely exclude them from potential avalanche release areas in Manali, India. Again, the algorithm was tested with a reference dataset from the Davos region. On a second attempt, Bühler et al. (2018) coupled the PRA method with an object-based image analysis algorithm which enabled the estimation of the fracture depth and avalanche volume. This information was then inputted to the RAMMS model and hazard mapping was performed for two scenarios with different return periods of avalanche occurrence in Davos.

Many more research papers have used GIS tools and DEM information to map areas susceptible to avalanches. As an example, Ghinoi & Chung (2005) combined the concept of PRA with Fuzzy Set theory algorithms and validated their results using a 20-year record of avalanche occurrence and weather observations in the Italian Dolomites. Bergua et al. (2018) presented a detailed snow avalanche susceptibility map of the eastern slope of the Aramo range, Spain, which they acquired using tools like digitized topographical maps, a DEM of the area, remote sensing imagery and orthophotographs, but, also, interviews, newspapers and field data. Barbolini et al. (2011) used GIS tools, computational routines, and statistical analysis to define locations potentially susceptible to snow avalanche release and flow. Other studies followed the alternative of susceptibility mapping using multicriteria analysis methods supported by geographical information tools (Durlević et al., 2022; Selçuk, 2013).

Addressing the phenomenon of dry-snow, slab avalanches from the perspective of fracture mechanics (failure initiation and crack propagation), a plethora of studies have been published. Reiweger et al. (2015) proposed an altered Mohr-Coulomb model to describe the double mode (shear and compression) weak layer failure behavior that was mentioned in the previous section, while finite element methods (Gaume, Chambon, et al., 2015; Schweizer et al., 2014; Sigrist & Schweizer, 2007), discrete element approaches (Gaume, Van Herwijnen, et al., 2015; Gaume, van Herwijnen, et al., 2017), analytical expressions (Heierli et al., 2008; Rosendahl & Weißgraeber, 2020), geostatistical techniques (Reuter et al., 2016), and mechanically-based statistical/probabilistic models (Gaume et al., 2013, 2014) are some examples of alternative methods found in the literature. From the perspective of triggering mechanisms and how they influence the release, the probabilistic analysis of slope and slab depth performed by McClung (2013) exhibits a lot of interest, while, with emphasis on skier-triggered, slab avalanches, it is worth mentioning the studies of Gaume et al. (2016), Gaume & Reuter (2017), Heierli et al. (2011), and Monti et al. (2016).

Lastly, machine learning methods seem to have attracted some but not a lot of attention for avalanche initiation prediction. There have been a few studies focused on predicting avalanche occurrence mainly using classification and regression trees (Davis et al., 1999; Hendrikx et al., 2005, 2014; Peitzsch et al., 2012). Also, logistic regression models have been used for exploring relationships between weather variables and avalanche days, while Gauthier et al. (2017) used logistic regression as a forecasting tool for avalanches on a daily basis in Québec, Canada. Kumar et al. (2019) applied a frequency ratio model for avalanche susceptibility mapping in the Lahaul

region in the Western Himalaya, and Choubin et al. (2019) produced avalanche susceptibility maps for two mountain regions in Iran after comparing a Support Vector Machine, a Random Forest, a Naïve Bayes model, and a Generalized Additive model.

### 2.3. Gaps in the existing knowledge

As already mentioned, our ability to quantitatively predict snow avalanche activity is impeded by our imperfect understanding associated with avalanche release mechanisms. A main reason for this is that the processes involved in the occurrence of the phenomenon range from snow deformation and metamorphism at the microscale to fracture mechanics at the slope scale, under the influence of often overlapped weather conditions (wind, precipitation, temperature variations, etc.) (Schweizer et al., 2016). The fact that avalanches are rare events and not possible to be reproduced, access to the field can often be dangerous, and laboratory experiments are hard to perform due to the fragile character of natural snow are some additional factors that further hamper our grasp of the physics behind the initiation of the phenomenon (Schweizer et al., 2003). Schweizer et al. (2003) concluded their comprehensive report on snow avalanche formation by highlighting several gaps in the existing knowledge on the topic, which remain until today. For example, “Which mechanical properties of which slab/weak layers describe the propensity for fracture propagation?”, “What are the spatial scales of variability relevant to slab release as influenced by topography (aspect, inclination, distance to ridge, and ground cover), snow type, and meteorological conditions during/after deposition?”, “How is snow stability affected (quantitatively) by surface warming?”, and “How does failure initiation and fracture propagation occur for wet slabs?”, are just a few of them.

### 2.4. Studies on snow avalanches in Andorra

Andorra is regularly faced with avalanches but has not been sufficiently researched for such kind of phenomenon. As it is mentioned by Oller et al. (2010), Janeras Casanova & Furdada Bellavista (2002) were the first to conduct an avalanche risk assessment by modelling the event of L’Armiana de Canillo and assessing the vulnerability of the affected area. More recently, Riba Porras et al. (2018) showcased that although useful to provide insights about the impacts of an avalanche, current one-dimensional and two-dimensional flow models need careful examination and acquisition of quality, historical snow and weather data, cadastral maps, DEMs, and stratigraphic profiles of the snowpack for the scope of model calibration.

Apodaka et al. (2018) and Esteban et al. (2005) stressed how different circulation types that cross the country contribute to the variability of the spatial distribution of heavy snowfalls even at locations really close to one another. Furdada et al. (2020b) provided a holistic summary and documentation of the terrain characteristics, weather and snowpack conditions that resulted in the biggest avalanche event in the country's history, the one that took place in Les Fonts d'Arinsal on the 8<sup>th</sup> of February 1996. Finally, a less related to avalanche forecasting but a, nevertheless, innovative study was presented by Francisco et al. (2018), who attempted to analyze mobility dynamics of backcountry skiers in the avalanche susceptible region of Ordino, using mobile phone call history data.

### 3. RESEARCH METHODOLOGY

#### 3.1. Methodological framework

The flowchart presented in Figure 5 illustrates the workflow that was followed during this study. It also constitutes the methodological framework developed to address the research questions and meet the aim and objectives of the thesis.

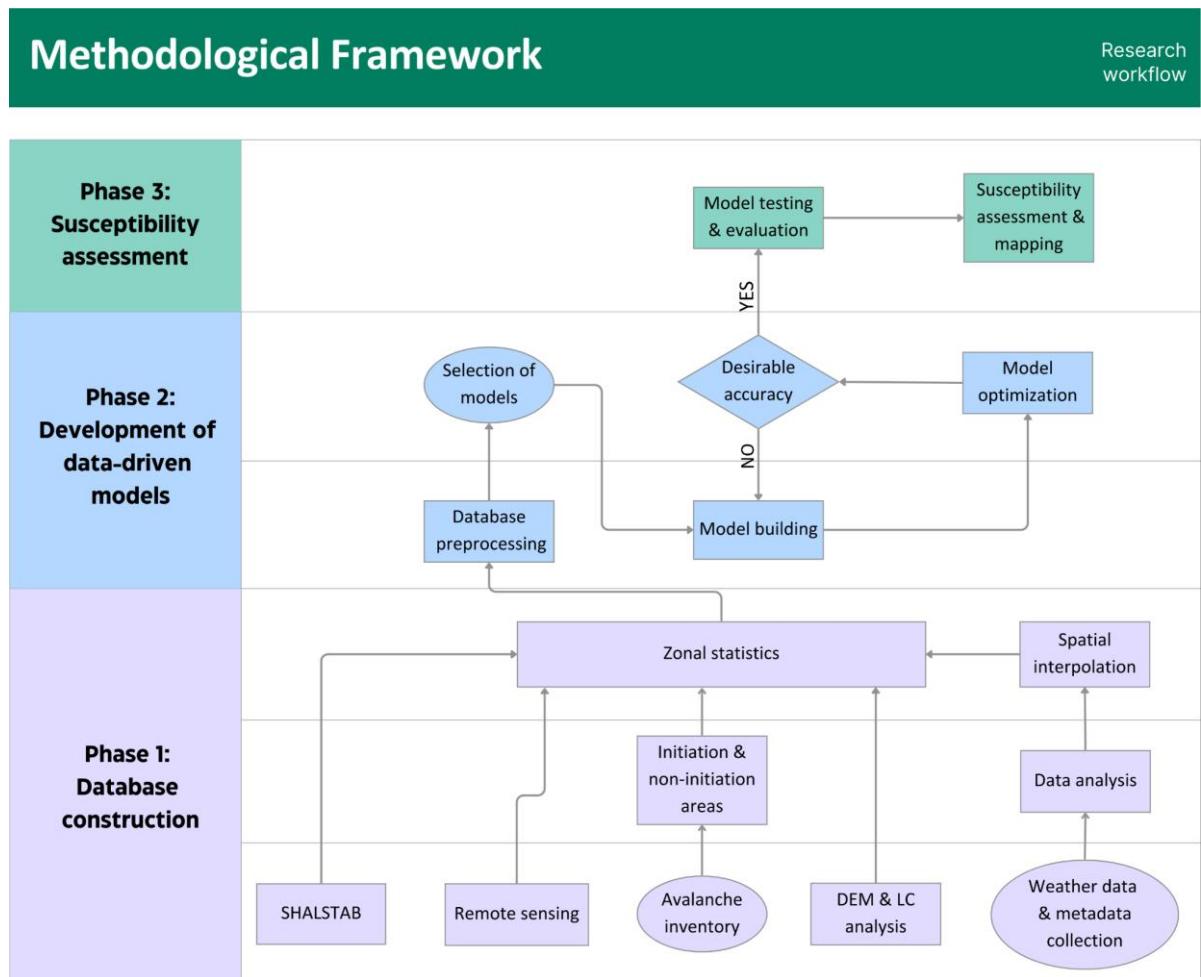


Figure 5. Methodological framework of the study.

The initial phase (Phase 1) comprises two parts and consists of tasks that contribute to gaining some insights regarding the input data, as well as, to the construction of a database to feed the data-driven models. More specifically, initially, the avalanche inventory is collected, analyzed, and processed accordingly, so that it serves the needs of the study, and the same is done for the weather station data. For the avalanche inventory this means selecting the events that are to be used and

developing a tool to acquire the initiation areas of the selected avalanche scars. For the weather data unit transformations, cleaning, checking for missing values and visualizations are the first actions to be taken. Consequently, statistical analysis of all the weather variables is performed, as well as, some dimensionality increase, which means that some extra variables are calculated from the provided raw data. All analysis and handling of the timeseries data is performed using the Python programming language (Van Rossum, 2023). The source code can be found at <https://github.com/ginamanou/avalanche-susceptibility.git>. In addition, snow cover information is acquired from Landsat 8 imagery (U.S. Geological Survey, 2019), using Google Earth Engine (Gorelick et al., 2017), and the Shalstab model (Dietrich & Montgomery, 1998) is run, in order to use its output as a predictor for the models.

The second part of Phase 1 includes the selection of representative values for all the weather variables and their spatial interpolation from point/station data to the whole Andorra. Subsequently, some DEM and land cover analysis is performed in order to acquire additional terrain features of the study area. Lastly, a number of shapes equal to the number of the selected historical avalanches are generated, in random locations of the grid that are not occupied with an avalanche scar. These shapes are considered to represent areas that haven't so far and do not generally initiate avalanches. When all data are converted to a gridded format, their zonal sampling using all polygon areas, both susceptible and not susceptible to avalanche release, follows. All geospatial analysis is performed using QGIS (QGIS Development Team, n.d.) and ArcGIS (Redlands, 2011) complementary to each other based on the needs of each task. After Phase 1 is completed, a database of samples that resulted or not to an avalanche is created.

The second phase of the study (Phase 2) is dedicated to the development of several data-driven models, and, more specifically, comprises the preprocessing of the database created in Phase 1, the selection of different machine learning algorithms, and the repeated cycle of training, validation, and optimization of the models in Python. The selected performance metrics are also presented in this phase. The third and last phase (Phase 3) includes the testing of the models on “unseen” data, evaluation of their performance and their generalization capability, and, finally, selection of the best-performing method for avalanche susceptibility assessment and susceptibility mapping.

In the following sections of this chapter, the steps and methods followed in Phase 1 and Phase 2 are presented with more detail.

### 3.2. Phase 1: Database construction

As depicted in Figure 6, this first phase of the methodology is composed of different processes, which are differentiated with different colors. Starting from the top of the graph boxes with blue color represent a series of actions with regard to the avalanche inventory, the green boxes represent the handling and exploration of the weather data, the yellow elongated box is dedicated to remote sensing and the pink box below it is about the Shalstab model. The grey boxes at the bottom are related to the terrain analysis, while with the brown boxes the land cover is depicted from which the ground roughness is acquired. Lastly, the forest land is presented with the dark-green-colored box. Going through all these processes step by step results in the construction of a database to feed the data-driven models. In the next paragraphs each process is described thoroughly.

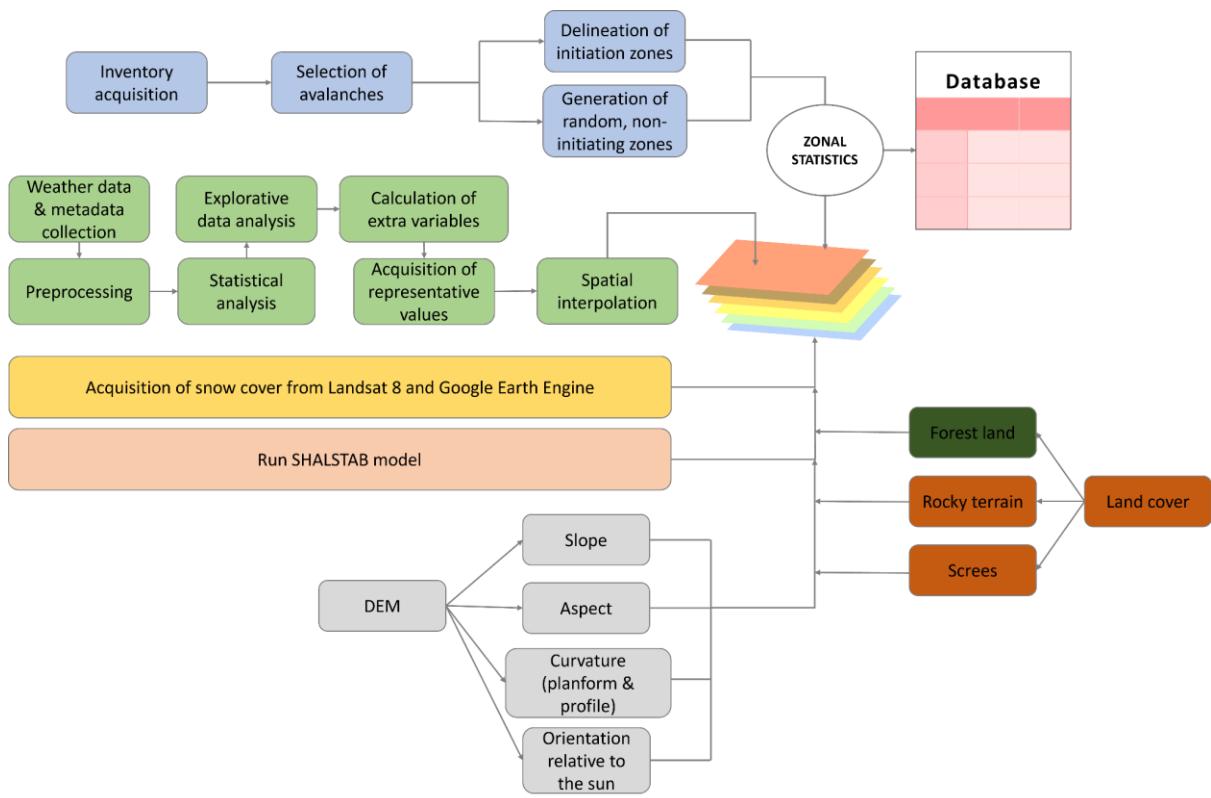


Figure 6. Phase 1 of the methodology: Processes followed for the construction of the database.

#### Avalanche inventory

When the inventory is at hand, information about the observation method and date, differentiation of avalanche type and registration protocol are necessary for the selection of the events which are to be used. The selected avalanche scars should not overlap with each other and they should be

distinguished as a single event. When the final batch of events is collected and presented in an ESRI shapefile format, the delineation of the initiation areas is performed using QGIS in combination with Python. The developed methodology for this comprises the following actions:

- a. derivation of the contour lines from the DEM using QGIS's "Raster -> Extraction -> Contour" processing tool,
- b. intersection of the contour lines layer with the avalanches layer using the tool "Vector -> Geoprocessing Tools -> Intersection",
- c. acquisition of the centroids (points) of the contour lines that fall inside the avalanche scars using the expressions "x(centroid(\$geometry))" and "y(centroid(\$geometry))" in the Field Calculator for latitude and longitude, respectively,
- d. exporting the contours layer and importing it into Python,
- e. development of a code that finds for each avalanche the centroid that is located around (greater than) the 80% quantile of elevations, representing, approximately, the highest part of the scar,
- f. importing a .csv file with all the acquired points back into QGIS,
- g. calculating in the Filed Calculator the area of each avalanche using the "\$area" function, dividing it by 4 and calculating a radius so as if this area was a circle, and, lastly,
- h. buffering each point imported with the .csv file using the radius acquired in the previous step and the SAGA (Conrad et al., 2015) tool "Vector general -> Shapes buffer" (Dong et al., 2003).

After this process is completed, the final product is a vector layer containing the initiation areas of each avalanche initially selected from the inventory, delineated as circles of  $\frac{1}{4}$  of the area of the whole scar. It has already been mentioned that every avalanche is split into three parts: a starting zone, a track, and a deposition area. However, there is not a standard presumption or consensus regarding the size of each part. As D. McClung & Schaefer (1993) put it, "Usually the lower limit of avalanche origin is ill-defined, and sometimes, a guess has to be made when trying to define the starting zone.". For this study, the  $\frac{1}{4}$  of each avalanche is guessed to be representing its initiation area, based on suggestions from local experts in Nivorisk.

Subsequently, another vector layer needs to be created containing zones that do not initiate or haven't initiated avalanches in the past (i.e., areas that are not registered in the inventory). The number of these zones has to be equal to that of actual initiation areas acquired in the previous step, so that the final database is balanced. The sequence of actions to be performed in this step constitutes a very simple approach in order to acquire points that do not overlap with any historical avalanche and is described below:

- a. First, generate a large number of random points in the study area by applying the tool “Vector -> Research Tools -> Random Points in Polygons” in QGIS, then,
- b. select only those points located in areas without registered avalanches. This constraint can be met by first dissolving (“Vector -> Geoprocessing Tools -> Dissolve”) all the avalanches and then buffering (“Vector -> Geoprocessing Tools -> Buffer”) using a distance of 100 m. The points that do not initiate avalanches can then be obtained by taking the difference (“Vector -> Geoprocessing Tools -> Difference”) of the random points layer and the intersection of the random points with the previously created buffer area.
- c. After the points have been acquired, use the “Shapes buffer” tool again and consider random radius for each point. The range of the radii is between 10 m and 70 m. These two limits are selected based on statistics acquired from the historical avalanches. The smallest initiation area was found  $\sim 300 \text{ m}^2$  and the median  $\sim 15,000 \text{ m}^2$ , corresponding approximately to the selected radii values for the non-initiating zones.

### Terrain analysis

The significance of different topographical features for the occurrence of snow avalanches has already been underlined. The steps to be followed for the acquisition of relevant topographical characteristics of the study area are outlined below:

- a. The 5x5 m DEM, providing terrain elevation information, is acquired from the website of the Geographical Information Systems (SIGMA) of the Government of Andorra (*Model Digital d'Elevacions*, n.d.).
- b. Using the DEM as input, the gridded slope, defined as the angle measured in degrees between the surface of the terrain at each cell and a hypothetical horizontal plane, is derived using the built-in tool “Raster terrain analysis -> Slope” of QGIS from the Processing Toolbox.
- c. Similarly, the aspect, defined as the direction towards which the slope faces and measured clockwise in degrees from the north, is derived using the built-in tool “Raster terrain analysis -> Aspect”. Then it gets reclassified using the Raster Calculator. Table 1 presents the breaks for the reclassification, and the orientation labels assigned to the reclassified values.
- d. The curvatures, planform and profile, are acquired using the “Curvature” tool available in the Spatial Analyst of ArcGIS. The output of this tool is the second derivative of the input surface (here the DEM) (*Desktop Help 10.0 - How Curvature Works*, n.d.). The profile curvature is parallel to the direction of the maximum slope, while the planform curvature is perpendicular to it. When the curvature values are negative, convexity is indicated, when they are positive, they

reveal concavity, and when they are zero, they represent a cell with a linear surface. Figure 7 depicts the different combinations of the two types of curvature.

- e. The 25x25 m Land Cover map of Andorra, based on orthophotographs of 2012, is also acquired from SIG (*Mapa de Cobertes Del Sòl d'Andorra* (2012), n.d.). From this, only specific land cover categories are selected, namely areas of dense forest, rocky terrain and screes, and three separate polygons are created. The reason why these classes were distinguished is because surface condition is strongly related to avalanche formation. New snow is likely to bond well with rough surfaces (e.g., rocks), meaning decreased probability of avalanche release (D. McClung & Schaerer, 1993). On the other hand, snow avalanches are common on scree-laden slopes (Blikra & Nemec, 1998; Sanders et al., 2014). The rocky terrain and scree polygons are converted into raster format using the tool “Raster -> Conversion -> Rasterize (Vector to Raster)” of QGIS.
- f. Next, the map of forest types of the study area is downloaded from the database of SIG (*Mapa Forestal Del Principat d'Andorra*, n.d.). Using this map as a basis, another one is produced by selecting only the coniferous tree species. Only dense coniferous forests, where trees do not lose their leaves during winter, are capable of intercepting enough snow and/or prevent snow transport by wind and, thus, hinder avalanche release (Bühler et al., 2013; D. McClung & Schaerer, 1993). The final product of this step is the combined map of dense, coniferous forest, acquired through the intersection of the two vector layers and converted into a raster using the “Rasterize” tool again.
- g. Finally, the exposure of slopes to the sun is related to the incoming radiation, which influences several properties of the snowpack. In general, the snowpack in sunny slopes is warmer and more stable than in shady slopes (D. McClung & Schaerer, 1993). In order to account for this factor and considering that Andorra is located in the northern hemisphere, where during mid-winter the sun doesn't reach the north slopes of ridges, a binary raster is derived using the Raster Calculator, with orientation distinguished between “north” (labels “North”, “North-West”, and “North-East” from Table 1) and “not north” (all the other labels). Figure 8 provides a comparison of the sun orientation in Arinsal, Andorra, on the 1<sup>st</sup> of December 2022 and on the 30<sup>th</sup> of April 2023. Similar angles of sun orientation are observed all around Andorra, and, thus, Figure 8 serves as the base on which the differentiation sunny vs. shady slopes is made.

Table 1. Reclassification categories of the slope aspects.

Degrees	Orientation	Value
-9999	Flat	0
0-22.5	North	1
22.5-67.5	North-East	2
67.5-112.5	East	3
112.5-157.5	South-East	4
157.5-202.5	South	5
202.5-247.5	South-West	6
247.5-292.5	West	7
292.5-337.5	North-West	8
337.5-360	North	1

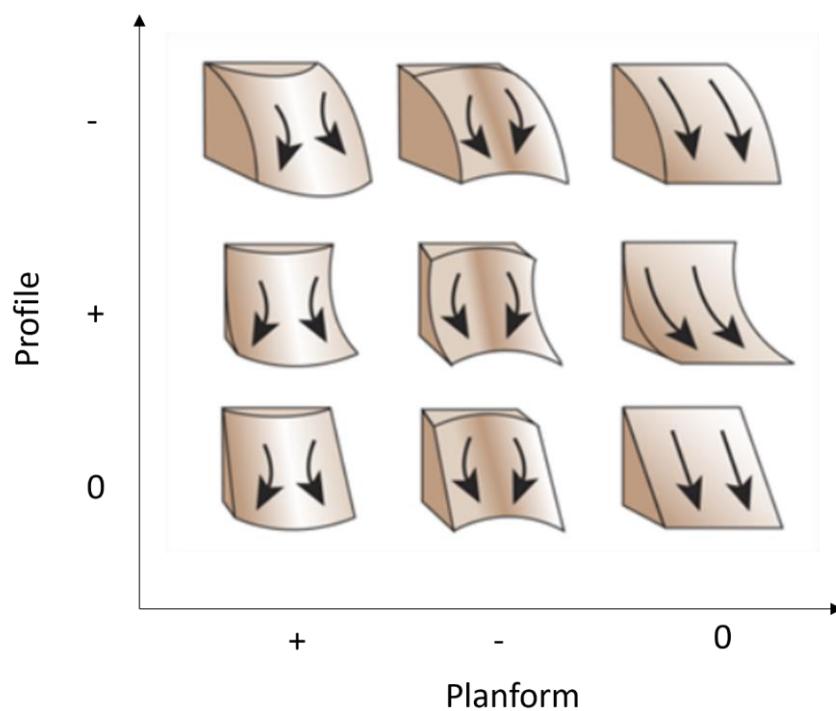


Figure 7. Combinations of planform and profile curvature (Curvature Function—ArcMap | Documentation, n.d.).

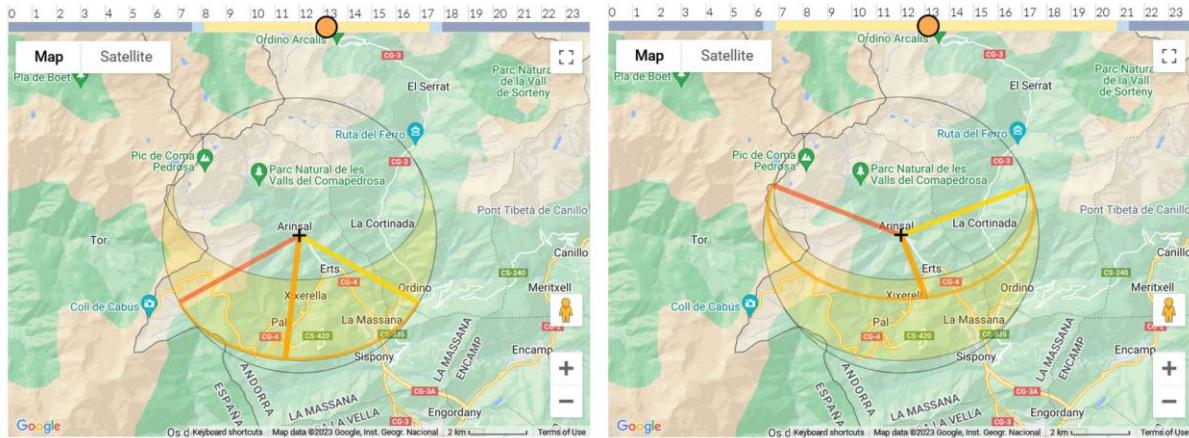


Figure 8. Sun orientation in Arinsal, Andorra, on the 1st of December 2022 (map on the left) and on the 30th of April 2023 (map on the right) (Arinsal: Sun Direction, n.d.).

### Remote sensing

For the acquisition of locations in the study area that are usually snow covered, the alternatives of the Modis (Hall & Riggs, 2021) and the Landsat 8 Collection 2 (Masek et al., 2020) remote sensing products, provided by the National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS), respectively, were considered. The Modis dataset, available since the year 2000, is a daily global collection of images of composite snow cover and albedo projected to a 500x500 m grid. On the other hand, the Landsat 8 Collection 2 Tier 1 provides a surface reflectance-derived normalized difference snow index (NDSI) at a 30 m spatial resolution (Vermote et al., 2016). Tier 1 collection, available since 2013, includes images from the Landsat 8 Operational Land Imager (OLI) instrument with the highest available data quality. Landsat is preferred over Modis because of the higher resolution of the images, which is considered as an important factor with regard to susceptibility mapping.

The NDSI is the normalized difference between spectral bands green (Band 3) and the shortwave infrared (Band 6), ranging from -1 to 1 and delivered as a single band product. The closer the NDSI pixel value is to 1, the higher is the probability of snow cover. In the current methodology, the final image is acquired as a GeoTIFF file using Google Earth Engine.

### Shalstab

Shalstab is a simple, deterministic model used for qualitative assessment of shallow landslides. It combines a slope stability model and a hydrological model. Regarding slope stability, Shalstab is

based on an infinite slope approach derived from the Mohr-Coulomb failure criterion as described in Equation (1):

$$\tau = C + (\sigma - u) * \tan\varphi \quad (1)$$

where  $\tau$  is the critical shear tension caused by the weight of the soil,  $C$  is the cohesion,  $\sigma$  is the normal stress,  $u$  is the pore pressure opposing  $\sigma$ , and  $\varphi$  is the internal friction angle of the soil mass at the failure plane.

One of the simplifications of the model is to consider the cohesion equal to zero. To counterbalance this omission, the developers suggest considering a higher value of the internal friction angle. Hence, with the assumption of zero cohesion and after some trigonometrical processing, Equation (1) is converted into the form presented below:

$$\frac{h}{z} = \frac{\rho_s}{\rho_w} * \left(1 - \frac{\tan\theta}{\tan\varphi}\right) \quad (2)$$

in which  $z$  is the soil depth,  $h$  is the water level above the failure plane,  $\rho_s$  and  $\rho_w$  are the soil and water bulk density, respectively,  $\theta$  is the slope, and, again,  $\varphi$  represents the internal friction angle.

In effect,  $\frac{h}{z}$  is the relative wetness or degree of saturation of the soil at instability.

In addition, Shalstab uses a simple, steady-state, shallow subsurface, flow model as described by Equation (3):

$$\frac{h}{z} = \frac{q}{T} * \frac{a}{(b * \sin\theta)} \quad (3)$$

where  $q$  is the magnitude of the precipitation event,  $T$  is the transmissivity of the subsurface (i.e., the ability of the soil to transfer the water downslope),  $\theta$  is the slope,  $a$  is the runoff drainage area of each cell of the grid and  $b$  is the size/width of the cells.

According to Equation (3), the pattern of  $\frac{h}{z}$  is determined by a hydrological ratio ( $\frac{q}{T}$ ) and a topographic ratio ( $\frac{a}{(b * \sin\theta)}$ ). The larger the hydrological ratio the more likely the ground is to saturate, and, as a result, the greater the number of sites on a hillslope that will become unstable

(where the  $\frac{h}{z}$  specified by (3) exceeds that given by (2)). The topographic ratio describes how topography affects runoff. To borrow the explanation from the model's manual, "The effect of topographic convergence on concentrating runoff and elevating pore pressures is captured in the ratio  $\frac{a}{b}$ , which shows that the larger the drainage area relative to the cell width, the higher the  $\frac{h}{z}$ ". Lastly, the steeper the slopes (greater  $\sin\theta$ ), the faster the subsurface flow and, thus, the lower the  $\frac{h}{z}$  ( $\sin\theta$  is in the denominator of (3)).

Equation (4) describes the coupled hydrological and slope stability model, by combining Equations (2) and (3) and solving for the hydrological ratio:

$$\frac{q}{T} = \frac{\rho_s}{\rho_w} * \left(1 - \frac{\tan\theta}{\tan\varphi}\right) * \frac{b}{a} * \sin\theta \quad (4)$$

The developers suggest considering  $\rho_s$  and  $\varphi$  to be constant throughout the grid, and, in this way, making the model parameter-free. Consequently, considering that Shalstab is a qualitative model, by comparing  $\frac{q}{T}$  values, conclusions about cells that are more prompt to instability than others can be drawn.

In summary, the input variables that are necessary in order to run the model are the slope, the internal friction angle, and the ratio  $\frac{b}{a}$  (removing the constant factor  $\frac{\rho_s}{\rho_w}$ ). It should be noted that the assumptions made in this step by considering snow instead of soil are bulk and simplistic. However, it has already been stated that the goal is not to assess the direct transferability of the model for avalanche susceptibility assessment, but rather use its output in order to evaluate if it could be a good predictor for the data-driven models.

Shalstab can be run in QGIS directly in the Raster Calculator or by using the respective SAGA tool (*Tool SHALSTAB / SAGA-GIS Tool Library Documentation (v3.0.0)*, n.d.). In this case, the option of SAGA is preferred. The slope raster is already available from the terrain analysis described above,  $b$  is equal to the grid cell size, here 5 m,  $\varphi$  for snow is considered 40°, and the drainage area  $a$  is acquired using the "Hydrological Analysis -> DInf Flow Accumulation" tool available from the "WhiteboxTools" plugin in QGIS (*Whitebox Geospatial*, n.d.). As explained by Dietrich & Montgomery (1998), the selection of the algorithm for the calculation of the drainage area is particularly important, in order to avoid "huge artifacts, in which the drainage area to a

point on the landscape may depend strongly on where it is relative to the orientation of the grid system.” The selected tool generates a flow accumulation grid using the D-infinity algorithm (Tarboron, 1997), which is suitable to evade such artifacts. Note that the DEM is required to be filled for sinks/depressions before being inserted as input to the “DInf” tool. This is achieved using another tool from the ‘WhiteboxTools’ plugin, namely “Hydrological Analysis -> FillDepressions”.

The output of SAGA’s Shalstab tool is the critical recharge in two formats: a) a continuous raster with unconditionally stable cells blanked (no data), unconditionally unstable cells set to zero, and cells with positive values, and b) a classified grid with classes of decreasing instability probability ranging from 1 to 7 (with zero representing either unconditionally unstable or unconditionally stable cells). The critical recharge signifies the recharge at failure, and, therefore, if instability occurs with lower critical recharge, the respective cell is considered more unstable compared to others with higher values of critical recharge, because it means that less rainfall is needed to set the soil into movement downslope.

### Weather data

As discussed in Chapter 2, the weather variables that interplay for the release of snow avalanches are a lot, and discovering strong causal relationships between avalanches and the meteorological conditions during or close to the events is very important for successful avalanche forecasting (Gauthier et al., 2017). In this case, however, the inventory does not include the dates of the events and, thus, the acquisition of the weather during the avalanche days was not possible. Due to this constraint, the developed methodology aims at approaching the weather-avalanche relationship by using historical weather information in order to obtain the “local climate” or most prevalent weather conditions in the locations around the weather stations, and, afterwards, interpolate this information in the whole study area. To achieve this, additional information is needed. For instance, the location (coordinates, altitude) and type (e.g., meteorological, ski resorts etc.) of each station, the operating service and the time period of operation (e.g., winter vs. all-year stations and since when), the collection methods and instruments used, the units of the data, and their quality are important to know.

In this case, the available raw data include precipitation in the form of rain, minimum and maximum temperature, wind speed, wind direction, snow drift, old and new snow. The total snow depth is characterized as old snow, while freshly fallen snow is considered as new snow. Initially,

the raw data are checked and cleaned against invalid entrances (e.g., symbols like “?” or “/” that do not represent a measurement), negative values when they do not apply (e.g., for rainfall), and missing values. The duration of each station’s record and if it is the same for all the variables that it measures are also important information to be attained from the data. In addition, consistency between the records of interrelated variables needs to be checked. For example, there cannot be an observation of specific wind direction in a day when wind speed has been recorded as zero. At the end of this step, the data are visualized with different kinds of plots (e.g., boxplots, histograms, scatterplots for the continuous data, and wind roses), in order to get a first idea of the trends but also other inconsistencies that might be present, like outliers. For outlier detection the “Z-score” method (Saleem et al., 2021) is applied, and the detected values are removed.

The next step incorporates a test of homogeneity of the datasets. A timeseries is homogeneous if the variations are caused by (local) weather/climate only. Non-climatic, systematic changes that cause inhomogeneity in atmospheric variables’ records include changes in station location, changes of the measuring devices or their calibration, and micro-climate disturbances like houses, streets, and vegetation. Detected variations of such causes need to be corrected (González-Rouco et al., 2001; Mohammed & Scholz, 2023; Peterson et al., 1998). Local weather, systematic changes with either anthropogenic causes (e.g., irrigation, hydroelectric power plants, dams, land use changes etc.) or natural causes (e.g., disasters, climate change etc.) should not be corrected. The method used for this test is the “Double-Mass curve” (Searcy & Hardison, 1960). A double-mass curve is defined as “An arithmetic plot of the accumulated values of observations of two variables that are paired in time and thought to be related. As long as the relationship remains constant, the double-mass curve will appear as a straight line; a deviation denotes the timing of a change.” (*Double-Mass Curve - Glossary of Meteorology*, n.d.). Subsequently, further explorative data analysis is performed, including hypothesis testing and correlations between the variables.

From the available data some extra variables are calculated. The scope of this dimensionality increase has an explorative scope and intends to extract additional information from the existing weather predictors and test if the performance of the data-driven models is improved. As a guide for the selection of these extra variables both literature recommendations and intuition can be used. Extreme rainfall is estimated using the “Peak-Over-Threshold” method (Acero et al., 2011; Lang et al., 1999). In snow avalanche forecasting, not only daily, but also 48 and 72 hours are considered as critical time periods for accumulated snow (Conway et al., 2008; Gobiet et al., 2016; Hendrikx et al., 2004, 2014; Lehning et al., 1998). Therefore, a moving sum of new snow with a

window of 2 and 3 days, respectively, is calculated. The same calculation is performed for rainfall. Also, an average of the monthly rainfall and new snow are considered as extra inputs. Finally, the average temperature and the daily range of temperatures (absolute difference between maximum and minimum) are computed.

Representative values from the timeseries of each weather variable in each station are selected in order to be used for the interpolation. Once the representative values are selected, the point data is imported into QGIS and spatial interpolation using “Inverse Distance Weighting (IDW)” (Shepard, 1968) is performed (“IDW interpolation” tool in QGIS). Other interpolation methods like the simple “Thiessen/Voronoi polygons” (Okabe et al., 2009) and the more sophisticated, geostatistical method of “Kriging” (Oliver & Webster, 1990) with elevation as external drift were considered as well. Nevertheless, both were rejected because of the low network density. Creating smooth interpolation surfaces with IDW was considered as the option with the least impact on the final susceptibility map.

### Construction of the database

After the completion of the processes described in the previous paragraphs, all the different types of data (topographical features, weather data, snow cover, Shalstab susceptibility/critical recharge) are, finally, presented in a raster format. Also, a vector layer including initiating and non-initiating avalanche zones has been produced. Subsequently, using this vector layer as input and the QGIS tool “Raster analysis -> Zonal statistics” all rasters are sampled.

## 3.3. Phase 2: Development of data-driven models

### 3.3.1. Data preprocessing

The next phase of the methodology is the preprocessing of the database, comprising an explorative data analysis, feature engineering, feature selection and other techniques like feature scaling (e.g., normalization or standardization). Here, standardization is applied as part of the data preparation for some of the models which require such transformation of the input data. Standardization or “Z-score normalization” scales the features based on the standard normal distribution, where the mean is equal to 0 and the standard deviation equal to 1. It is important that standardization is performed for each one of the train, validation, and test subsets individually, in order to avoid “data leakage”, which, in simple words, is when information about the training set influences the validation and test sets. Additionally, it is important to standardize the validation and test set using

the mean and standard deviation of the training set and not calculate it separately for all the datasets. The reason is that a true test of the built model is how it performs on “out-of-sample” data using the parameters to learn from the training set. Feature engineering and feature selection are the processes of using field knowledge and tools borrowed from math and statistics in order to transform and select the most relevant data when building a predictive model.

During the explorative data analysis trends and relationships between the data are investigated and the data are prepared so that their format complies with the requirements of each model. More specifically, in some cases, appropriate labeling of some features is necessary (e.g., in this case, assigning cardinal labels to wind direction and aspect). Some additional dimensionality increase might also be performed in this step. Here, leeward and windward slopes are identified by comparing the wind direction with the aspect orientation. This extra feature could be an indicator of the possibility of snow transport and deposition on the leeward sides of ridges. Furthermore, during feature engineering, one-hot encoding of the categorical features is necessary, because the library used to implement the machine learning models does not support categorical data. One-hot encoding is a process by which categorical data are converted into numerical data. The “OneHotEncoder” class of the scikit-learn 1.3.0 (or sklearn in short) package (Pedregosa et al., 2011), available in Python, is used in order to create multiple binary features for each label of the initial categorical data.

The next step is feature selection, which is a process of identifying the features that are more useful to the models’ predictive performance. There are several techniques to achieve this, including the removal of data that have low variance (e.g., constant data) or variables that are highly correlated with others. Constant data are not good predictors because there is nothing in them that can explain the prediction, while including highly correlated variables into the models is likely to not provide any extra value to the performance and instead, make the models slower and cost on optimization time. In addition to that, certain models, like the logistic regression, are vulnerable to multicollinearity effects. The term multicollinearity describes the high correlation of independent variables and is a problem for logistic models because small changes in one variable can cause change in another, producing fluctuating results, unstable coefficient estimation and even overfitting.

Moreover, statistical methods are common ways of optimizing the selection of the data to be used as input to the machine learning models. Often times, machine learning algorithms are used out of the box, without any optimization, just to grasp an idea of the most important features (when

the algorithm provides this option of feature importance). For this study, a combination of methods is consulted for the final feature selection. The Pearson correlation coefficient (Freedman et al., 2007), included in the pandas library (McKinney, 2010) in Python and defined as the covariance of two variables divided by the product of their standard deviations, is calculated to check for multicollinearity. Also, both the chi-square (Pearson, 1900) and the mutual information method (Kraskov et al., 2004), which are included in scikit-learn, are employed here. The chi-square method, which identifies the features that are most likely to be independent of class, is used to rank the categorical features' relevance with regard to the target classes. The mutual information method measures the relationship between the independent variables and the dependent variable, with zero values denoting independence and higher values meaning higher dependency. Additionally, the output of feature importance of the “ExtraTreesClassifier” from sklearn is considered. The algorithm is explained in more detail in the next section.

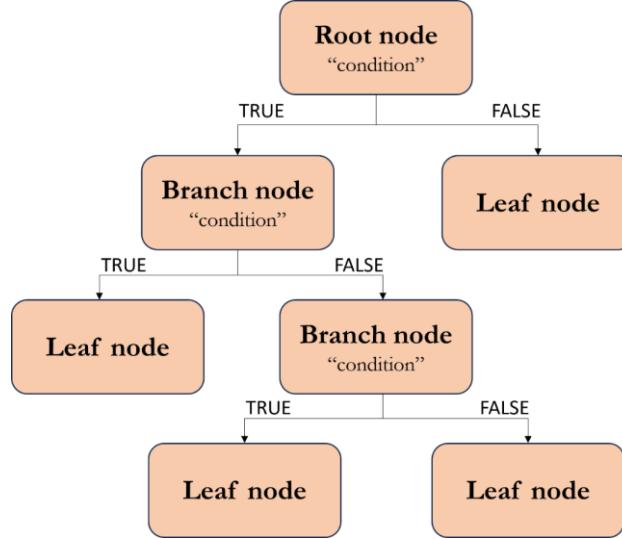
Finally, two different versions of the database are created by selecting different combinations of features each time. Each version of the database is then used to develop two versions of the models, so as to scout the predictive strength of certain predictors and meet the objectives c and d. In particular, one version consists of only topographical features, and the second version is the whole database, as described previously, including topographical data, weather data, snow cover from remote sensing and the output of Shalstab.

### 3.3.2. Model selection

The problem addressed in this thesis is, in effect, a classification with two possible outcomes, namely true or false, 1 or 0, or, more specifically, area susceptible to avalanche initiation or not. Binary classification is a sub-category of supervised learning, where the output is a discrete value, category or class. For the scope of the current study, different types of machine learning models are brought to the test, with the purpose of comparing their predictive strength and identifying the most suitable one for avalanche susceptibility assessment and mapping in the study area.

The first category of models to be tested includes those that are built with a tree-like structure (Figure 9), based on which they decide the target class. This kind of structure starts with a “root node”, which splits into either branches or leaves. Branches are called the nodes that split further, while leaves represent terminal nodes. In the second category belong Logistic Regression and Support Vector Machines (SVMs). These two models are grouped together because, under certain conditions, they share some similarities. The last category includes a feedforward neural network,

which is characterized as a “black-box” model, because it is impossible to interpret the results in terms of how the learnt function is acquired. In the following paragraphs, the selected models/algorithms and the libraries used for their implementation are described more extensively.



*Figure 9. Structure in the shape of a tree, based on which the so-called tree-based machine learning algorithms are performing classification tasks.*

### Decision Tree (DT)

In the current methodology, the DT is implemented using the `sklearn` library. It is based on the CART algorithm which is very similar to the C4.5 algorithm (Quinlan, 1993). According to `sklearn`'s documentation, the difference is that “CART constructs binary trees using the feature and threshold that yield the largest information gain at each node.”. Due to this characteristic, the algorithm is called greedy. The mathematical formulation of the algorithm consists of recursive splits of the feature space so that the same labels are grouped together. Each potential split is evaluated using an impurity or loss function. The two available options of impurity functions in `sklearn`'s “`DecisionTreeClassifier`” are the “gini index” (Equation (5)) and the “cross-entropy” (Equation (6)), which is equivalent to minimizing the log-loss between true and predicted labels.

$$\text{Gini index} = \sum_k p_{mk} * (1 - p_{mk}) \quad (5)$$

$$\text{Entropy} = - \sum_k p_{mk} * \log p_{mk} \quad (6)$$

where  $p_{mk}$  represents the proportion or probability of class k observations in node m.

DT learners tend to create deep, over-complex trees that do not demonstrate good generalization properties (overfitting). To overcome this, the current methodology suggests the Cost-Complexity pruning algorithm (Breiman et al., 1984) with a single complexity parameter,  $\alpha \geq 0$ . The larger the value of  $\alpha$ , the more pruning is applied to the tree. However, a fine balance needs to be maintained between avoiding overfitting through pruning and the desired performance.

### Random Forest

Random Forest classification (Hastie, 2009) is an ensemble learning technique that, again in a greedy manner, fits a number of decision trees on various, randomly selected sub-samples of the training dataset (with replacement) and averages them in order to improve the predictive accuracy and control overfitting. The method of using sub-samples of the original dataset is called bootstrapping. When the final decision is made based on the aggregate of all the fitted trees, which is the case of the Random Forest, the method is called bagging (Breiman, 1996). Again, sklearn is used for the implementation of the Random Forest algorithm.

### Extra Trees

The Extra Trees estimator, available in sklearn, is an extremely randomized tree classifier, included in the ensemble methods family (Geurts et al., 2006). The main element that distinguishes them from the classic DT is that Extra Trees is not a greedy algorithm. When looking for the best split to separate the samples of a node into two groups, splits are drawn for a number of randomly selected features and the best split among those is chosen. In the current study, the “ExtraTreesClassifier” is only used to assist the selection of features and not to build a predictive model.

### Adaptive Boosting (AdaBoost)

AdaBoost (Freund & Schapire, 1997) is another ensemble estimator similar to the Random Forest, with the difference that the trees built by AdaBoost, also called stumps, are usually the smallest possible, with only one node and two leaves. Because of their size, stumps are not as good at making predictions, and this is why they are often characterized as “weak learners”. AdaBoost fits a classifier/stump on the training dataset and then fits additional copies of the classifier on the same dataset, but it does so by adaptively adjusting to the errors of the incorrectly classified

instances, such that subsequent classifiers focus more on prior weak hypotheses. Thus, the order with which the stumps are created matters. Unlike Random Forests, the trees are not built independently of one another, and, in addition, they are scaled so that each stump contributes differently to the final classification. This scaling factor is called learning rate and can take values between 0 and 1. The algorithm implemented in sklearn is known as AdaBoost-SAMME (Zhu et al., 2009).

### Gradient Boosted Decision Trees (GBDTs)

GBDTs is, again, an ensemble machine learning technique, which approaches function estimation from the perspective of numerical optimization in function space rather than parameter space (Friedman, 2001). Similar to AdaBoost, it builds fixed-sized trees based on the previous tree's errors. However, Gradient Boost fits the regression trees on the residuals between the predicted and the true classes. In Gradient Boost, each tree can be larger than a stump, with the maximum number of leaves ranging, usually, between 8 and 32. Also, GBDTs are scaled, but unlike AdaBoost, the weight by which each tree is scaled is the same. Gradient Boost uses the  $\log(\text{likelihood})$  as a loss function. In Logistic Regression, as explained below, the larger the  $\log(\text{likelihood})$ , the better the prediction. Hence, when  $\log(\text{likelihood})$  is used as a loss function, where smaller values represent better fitting models, it needs to be multiplied by -1. Gradient Boost continues to build trees until the number of requested trees is reached or additional trees fail to improve the fit. The Gradient Boost algorithm is implemented, here, using the sklearn library.

### Extreme Gradient Boosting (XGBoost)

XGBoost is an open-source library which provides a regularizing gradient boosting framework for several programming languages, including Python (T. Chen & Guestrin, 2016). Here, the library's sklearn estimator interface is used. Same as the unextreme Gradient Boost, the XGBoost trees are built on the residuals of actual and predicted values, but instead of using the general regression trees, XGBoost constructs a unique type of regression trees. Additionally, XGBoost uses regularization, which is intended to reduce the sensitivity of its trees to individual observations by pruning and combining them with other observations.

### Logistic Regression

Logistic Regression (James et al., 2013) is similar to Linear Regression, but instead of fitting a line to the data it fits an "S" shaped "logistic function". The curve goes from 0 to 1 and provides the

probability of an outcome to be True or False. In the case of binary classification, a threshold is applied to its numerical output (usually equal to 0.5). The function is estimated using the maximum likelihood. Logistic Regression is a type of a Generalized Linear Model (GLM), which are a generalization of the concepts and abilities of regular Linear Models. The corresponding class of the sklearn package is used to implement it.

### Support Vector Machines (SVMs)

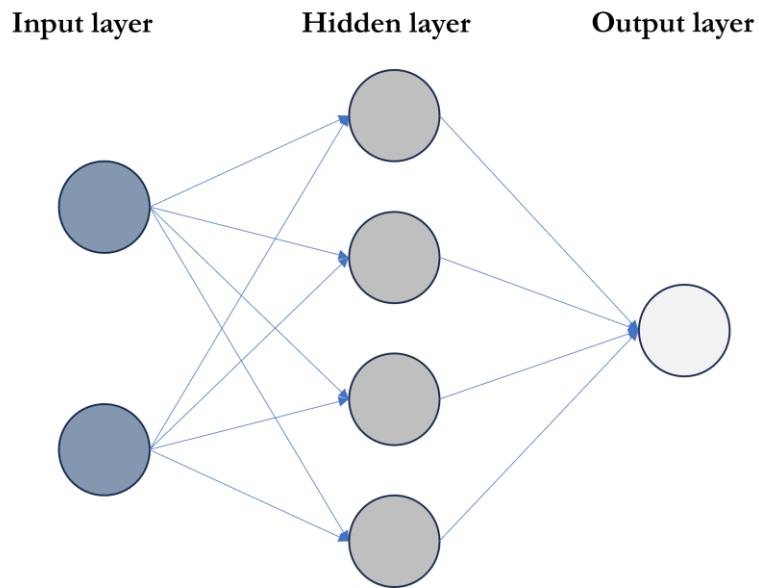
SVMs perform classification tasks by drawing lines or hyperplanes in order to separate the target groups based on patterns in the data (Cortes & Vapnik, 1995). There are many possible hyperplanes, and, thus, the objective of the model is to find the one that has the maximum margin, i.e., the maximum distance between datapoints of both classes. Hyperplanes are a kind of decision boundaries and by maximizing the margin distance, more confidence is ensured regarding the correct classification of future data. The implementation of SVMs in sklearn is based on the libsvm library (Chang & Lin, 2011).

### Artificial Neural Network (ANNs)

Artificial Neural Networks loosely mimic the functioning of neurons in the human brain, and, therefrom, originates their name. The concept behind ANNs is that by adding up many simple functions they can approximate any function that describes the relationship in a dataset, given enough points and values of this function. A commonly used ANN is the multi-layer perceptron (MLP), which consists of a number of interconnected nodes (called neurons) arranged into three types of layers: input, hidden (could be more than one) and output. Figure 10 depicts a simple representation of the structure of an MLP with only one output node. The lines represent weighted connections between the neurons. A node/neuron in a hidden layer receives signals/values from the nodes of the input layer and transforms them into signals which are sent to the output nodes. There, the signals are transformed into outputs. These transformations of input into output signals are realized through transfer functions. The transfer function in the hidden layer can be of any type, but in the output layer it is, usually, a non-linear, bounded function (usually of sigmoidal shape). The most widely used one is the sigmoid (logistic) function (Equation (7)), which is bound between 0 and 1.

$$f(u) = \frac{1}{1 + e^{-a*u}} \quad (7)$$

Here, a Feedforward Neural Network (with a similar structure to an MLP) is built using Keras (Chollet & others, 2015) and Tensorflow (Abadi et al., 2016). Tensorflow is an end-to-end machine learning platform and Keras is its high-level API, which covers every step of the machine learning workflow, from data processing, to hyperparameter tuning and deployment. The term feedforward indicates that the neural network is built following only one direction, from input to output, and there are no feedback loops between the layers. The ability to approximate is highly dependent on the network's complexity, which, in turn, is determined by the number of nodes in the hidden layer(s). The more hidden nodes a network has, the more complex functions it can approximate.



*Figure 10. Typical structure of an MLP with 2 input nodes, 4 hidden nodes and 1 output node.*

### 3.3.3. Training-validation-test split vs. Cross validation

There are two main ways to handle the available data in order to build machine learning models. One is called “Validation set” approach or “Hold-out set” approach and the other “Cross validation” approach. Often, the functionalities of the two methods can be combined. In the first approach, the available dataset is divided into two parts, a training and a validation or hold-out set. The model is fit on the training data and the fitted model is used to predict the classes of the validation set. The resulting validation-set error provides the best possible estimate of the test error or the error of the model on “unseen” data. In classification problems, the error is usually the misclassification rate.

The second approach is very similar to the first. The difference lies in that in cross validation each sample is used both to train and validate the model, but never at the same time. In situations where data is limited, this approach is preferred, because it requires less data to be set aside in each fold. This means that more data are seen by the model during training, which, generally, translates into better models. Figure 11 depicts a representation of a 5-fold cross validation.

Sample					
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Split 1	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Split 2	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Split 3	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Split 4	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Split 5	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5

Figure 11. Representation of the different sample folds in a 5-fold cross validation.

Another important concept to be mentioned is the “Bias-variance tradeoff”. In machine learning, variance is the difference between the fit of the model using the training and the validation set. Bias is the ability of the model to capture the true relationship of the data. In reality, there is a tradeoff between minimizing the bias and minimizing the variance, and an ideal model finds the balance between the two by being simple enough and as complex as necessary.

In this case, the “Hold-out set” approach is applied. But, since the database consists of a satisfactory number of samples (2,300 in total), it is split into three parts (60-20-20), namely one training set of 1,380 samples, one validation set of 460 samples and one test set of 460 samples. The latter is used for the final evaluation of the models and for reporting on the obtained error. However, this approach is susceptible to some variance on the error estimation. Thus, 10-fold cross validation is used to give estimates on the variability of the true error estimation. Also, cross validation is incorporated in the model development during the optimization stage.

### 3.3.4. “Training-validation-optimization” cycle

The cycle of “training-validation-optimization” is a recursive process which, theoretically, results in the best possible model or, at least, a model of satisfactory accuracy (Figure 12). A key aspect during this procedure is that, when the final model is optimized and ready to be deployed, only then can the validation set be utilized. A common mistake is to test a model using the hold-out set in an iterative process. When repeatedly using a holdout set to test a model during development, the hold-out set becomes contaminated, meaning that it no longer gives unbiased estimates of the

true model prediction error. Another example of contamination could be that training and validation data are processed together before model validation.

In the sklearn package, hyperparameters are parameters that are not learnt within the estimators, but, instead, they are passed as arguments to the constructors of the estimator classes (i.e., the different classifiers). The library offers two options of parameter search. One the one hand, the “GridSearchCV” approach searches the hyperparameter space exhaustively. On the other hand, “RandomizedSearchCV” can sample a certain number of candidates. Both approaches search the hyperparameter space for the best cross validation score. The difference between the two is that the first approach provides better results, since it considers all the parameter combinations. Experimentation in the scope of the current study showed that there is only small difference in the acquired performance between the two methods. On the contrary, the optimization time differs significantly, with “RandomizedSearchCV” being, by far, the fastest option. Therefore, both methods are applied here, depending on the case.

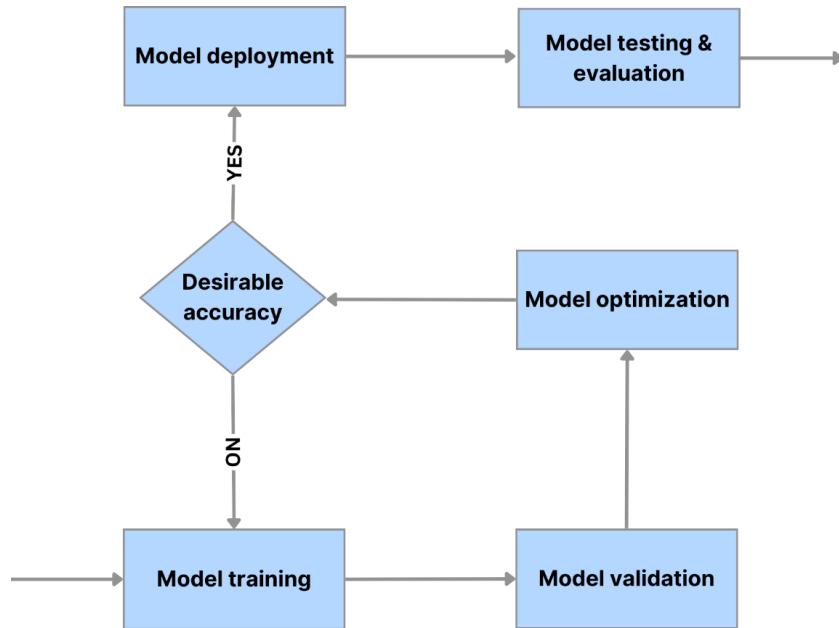


Figure 12. The cycle of "training-validation-optimization" followed by model deployment and testing when the desirable accuracy has been achieved.

### 3.3.5. Evaluation metrics

With regard to the accuracy metrics, the confusion matrix (Figure 13) is considered the most interpretable evaluation tool, because it provides a visual representation of the total amount of events tested and, more specifically, the ones that were correctly predicted, missed or the predictions that were false alarms. Hence, it is the primary tool to be used for evaluation of the

models' performances. Thereafter, there are a few other accuracy metrics that can be derived from the confusion matrix, which provide more concise information of the performance, given in percent.

First of all, the total accuracy score (Equation (8)) is the ratio of correct prediction per total amount of predictions. Recall (Equation (9)), also known as the True Positive Rate (TPR) or Sensitivity, is the percentage of data samples that are correctly attributed to the “positive class” out of the total samples for that class. Precision (Equation (10)) refers to the number of true positives divided by the total number of positive predictions (i.e., the true positives plus the false positives). The F1 score (Equation (11)) is the harmonic mean of the precision and recall, where the relative contribution of each are equal. Lastly, the “Receiver Operating Characteristic” or ROC curve, is a plot that has recall on the y axis and False Positive Rate (FPR) on the x axis. The False Positive Rate is described in Equation (12). The larger the area under the curve (AUC score), the better the model's accuracy. In the current study, all the above-mentioned metrics are advised to assess the performance of the models in different stages of the model development, and during the final evaluation and comparison.

True values	Predicted values	
	Not susceptible	Susceptible
Not susceptible	True Negatives (TN)	False Positives (FP) / False alarms
Susceptible	False Negatives (FN) / Misses	True Positives (TP)

Figure 13. The confusion matrix.

$$\text{Accuracy} = \frac{(TP + TN)}{\text{Total sample}} \quad (8)$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (9)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (10)$$

$$F1 = 2 * \frac{\text{Recall} * \text{Precision}}{(\text{Recall} + \text{Precision})} \quad (11)$$

$$FPR = \frac{\text{FP}}{(\text{TN} + \text{FP})} \quad (12)$$

### 3.3.6. Deployment of the models

The standard procedure suggests that, when the “training-validation-optimization” cycle is completed and the best parameters have been selected, the final model is trained using all the data, training and validation. This step is usually called model deployment, and in this case, it is realized by fitting all the optimized models/classifiers to a total of 1,840 samples (training and validation data added together).

## 3.4. Phase 3: Model testing and susceptibility mapping

On this last phase of the methodology, the created models are tested on “unseen” data in order to report about their performance and enable a comparison between them. This process is only performed once, after the iterative act of model development and the subsequent stage of model deployment have been completed (Figure 12). This way, using the test dataset conclusions can be drawn regarding each model’s performance and predictions can be made regarding avalanche susceptibility in the study area.

Lastly, in order to meet the last objective of this thesis (objective e), an avalanche susceptibility map covering the whole country needs to be produced. Which algorithm to select for this task is a matter of finding a balance between performance and interpretation. Some algorithms perform better than others, and in some cases, significantly better. However, they might be less elaborate on their learning process, making it hard to interpret the results and gain useful insights in order to delineate potential release areas on a map. Here, the selected algorithm for susceptibility mapping is the Decision Tree. The optimized DT is fitted to the whole database of 2,300 samples (training + validation + test), then the decision rules are extracted using the “export\_text” function of the “tree” class of sklearn, and, finally, these rules are passed into the “Raster Calculator” of QGIS, in order to produce the final map, where susceptibility is presented in a binary raster.

## 4. CASE STUDY: ANDORRA

---

### 4.1. Phase 1

In this section, the specifics of the application of the developed methodology are presented, starting with the availability of the input data, followed by their analysis, and concluding with the final product of Phase 1, which is the database.

#### 4.1.1. Data availability

##### Avalanche inventory

The first registries of the avalanche inventory that is maintained by the government of Andorra date back to 1984, which shows the long-lasting effort of the country to map avalanche hazard. Nevertheless, as has already been mentioned, the records do not include the dates of the events and, in the beginning, they were solely based on oral surveys of direct testimonies. Subsequently, the avalanche scars were manually digitized according to the descriptions. In recent years, the responsible authorities for avalanche control have complimented the inventory with identified historical avalanches through photogrammetry. For the scope of this study, both entries of data have been taken into consideration. Entries comprising multiple avalanches in certain zones but not distinguishing them into individual paths were not included in the analysis. Also, in case of identical codes (i.e., the same avalanche detected by both testimony and photogrammetry) but not completely overlapped avalanche scars (e.g., due to differences in observation methods and digitization accuracy), more weight was given to the ones recognized by a witness, and, thus, the ones derived through photointerpretation were considered duplicates and were excluded. At the end, 1,122 avalanche scars were identified (Figure 14).

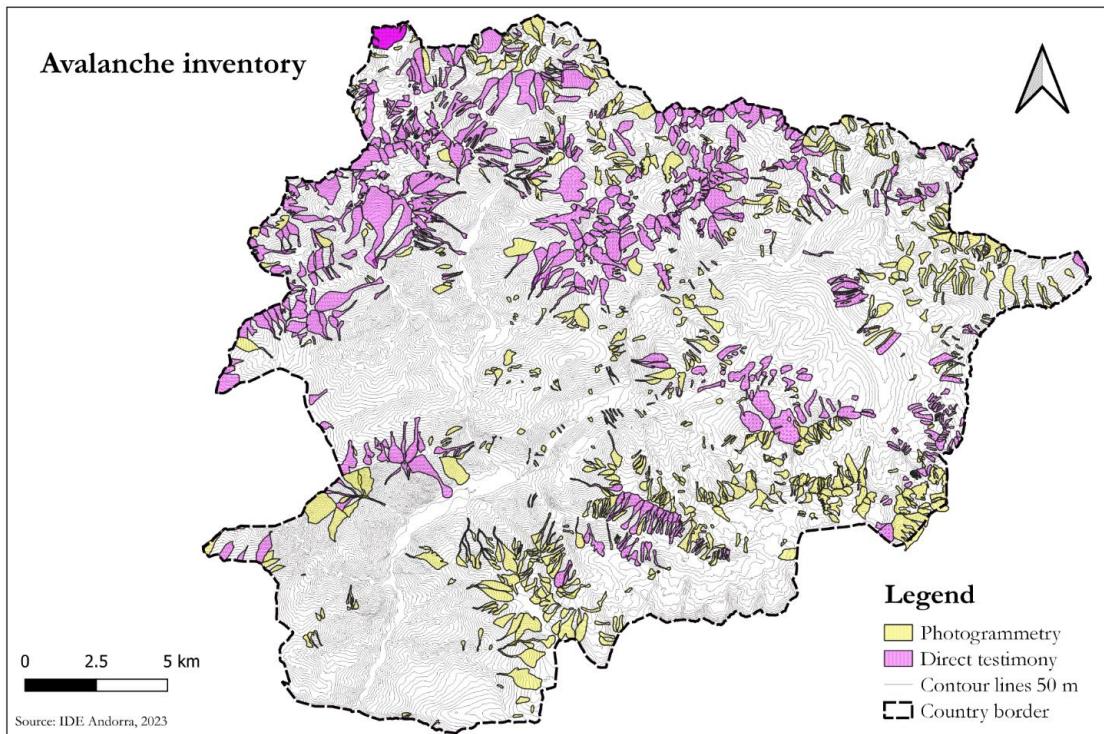


Figure 14. Map of historical avalanches detected by direct testimony (purple color) or photointerpretation (yellow color).

### Weather data

Andorra has a dense network of 33 weather stations, which, usually, measure wind speed, wind direction, snow depth, precipitation, temperature, relative air humidity, and solar radiation. In some cases, they also have sensors to capture atmospheric pressure, soil temperature, and snow temperature. Some of them are automatic hydrological stations, others are operated by the National Hydropower Service or by the Center of Studies of the Snow and the Mountains of Andorra (CENMA), and the majority belongs to the Andorran Meteorological Service, which at the moment, is in the process of checking the quality of all their stations' records. Following their advice, only the stations that have already passed the quality control were considered (Figure 15), resulting in a not so dense network. More specifically, the data for the study were acquired from 13 stations, of which 7 are winter stations and the rest operate all year round.

In Andorra, snow-related data are collected only during the winter months, in stations located in ski resorts. Meteorological stations that operate all-year round do not have snow observations, because, in most cases, their location is in areas where snow conditions are rare. The thickness of

new snow is measured at 8 a.m. and at 1 p.m. on top of a board that is cleaned of new snow every day at 8 a.m. It is not cleaned at 1 p.m. if it has snowed. Thus, if the thickness of new snow decreases between 1 p.m. and 8 a.m. the following day, it means that there has been snow transport. At the ski resorts, the wind observations are instantaneous, measured at about 2 meters height by an observer. At the other stations average but also maximum/gust wind speed and direction records are available, with measurements taken between 3.5 meters and 10 meters from the ground. For the scope of this study, the instantaneous wind from the ski resorts' records is considered to represent the average conditions of each day, and from the rest of the stations the measurements of both average and gust wind are considered in the analysis.

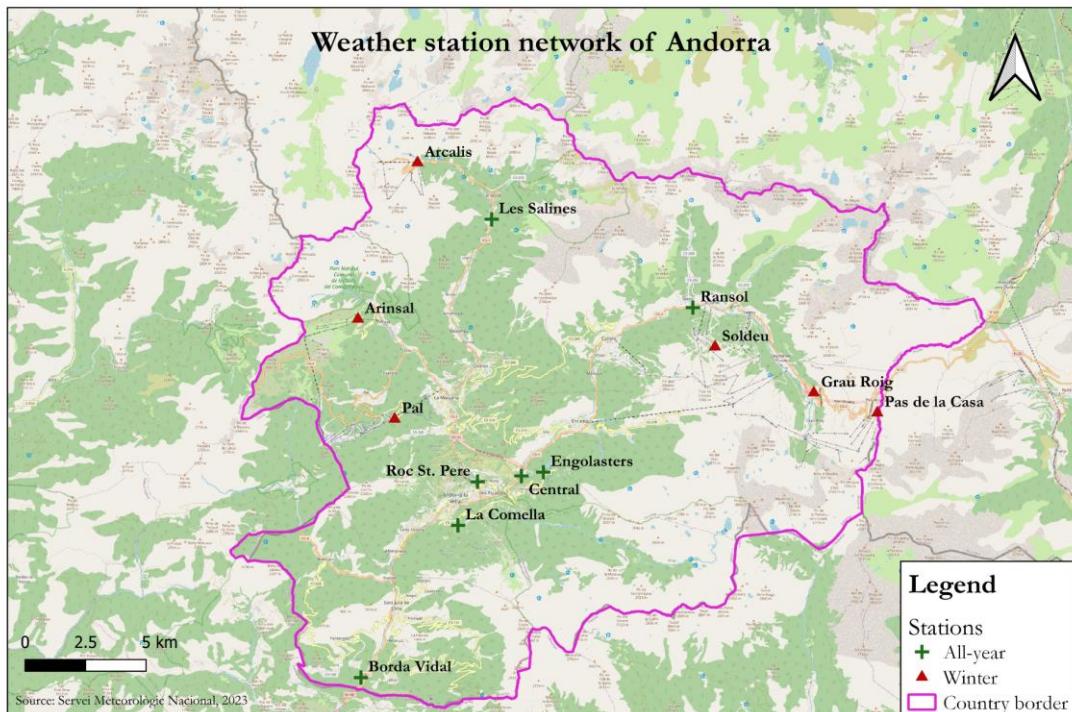


Figure 15. Map of the weather stations that were considered in the study.

The considered stations have daily observations but different record lengths. Nevertheless, since the current methodology does not include timeseries analysis but rather long-term averages of the weather observations, each station was analyzed for the whole period of available records. Table 2 provides a summary of the variables that are observed in each one of them (cells highlighted with light-blue color), as well as the start and end date of their records. Stations in green operate the whole year, whereas stations in pink only during winter.

According to the gathered metadata, the transport of snow (or snow drift), for the most part, must be consistent with the direction of the wind. However, it may not be. As mentioned above, wind direction measurements are instantaneous, while snow transport is an observation considered since the last observation. The surroundings of the observation point are observed and if there have been changes (e.g., bare ridges, congestion, etc.), it is noted. Wind transport of snow means that some areas will gain snow and some will lose it, and usually the observation point loses it. Hence, the new snow, and even the old snow can decrease. Sometimes it can be seen in the data that it snowed in a station (there is a precipitation measurement), but it was also windy and, thus, the new snow was 0, even though the temperature was negative. In the ski resorts snow transport is recorded as codes that describe intensity (e.g., moderate transport) and direction (e.g., which wind direction drifted the snow).

*Table 2. Data availability and record length of the weather stations considered in the analysis.*

Station	Start year	End year	Temperature	Precipitation	Wind speed	Wind direction	Old snow	New snow	Snow drift
Arcalis	1983	2022							
Arinsal	1980	2022							
Grau Roig	1980	2022							
Pal	1981	2022							
Pas de la Casa	1981	2022							
Soldeu	1981	2022							
Ransol	1981	2022							
Borda Vidal	2008	2022							
Central	1934	2022							
Engolasters	1934	2022							
La Comella	2007	2022							
Les Salines	1985	2022							
Roc St. Pere	1992	2022							

### Digital Elevation Model and Land Cover

The DEM and the land cover map based on which the analysis of the topographical features was performed are presented in Figure 16 and Figure 17. In the elevation map it can be seen that the east, north-east, north, and north-west parts of the country are found in high altitudes, up to almost 3,000 m, while the central, south, and south-west parts are located lower, in valleys. In the land cover map, it is depicted that dense forest is the prevalent class, occupying almost 200 km<sup>2</sup> out of the 468 km<sup>2</sup> of the total area of the country. It is followed by meadows, scree-laden areas, thickets in the lower elevations and rocky terrain. In Figure 18, a map of forest types is presented. Coniferous trees constitute the majority of the forest land.

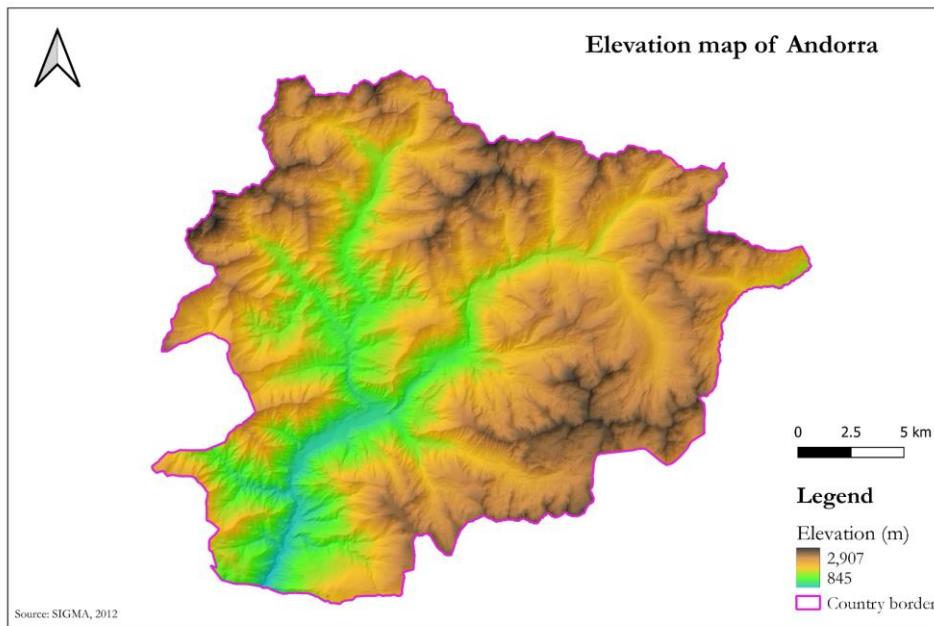


Figure 16. Digital Elevation Model of Andorra.

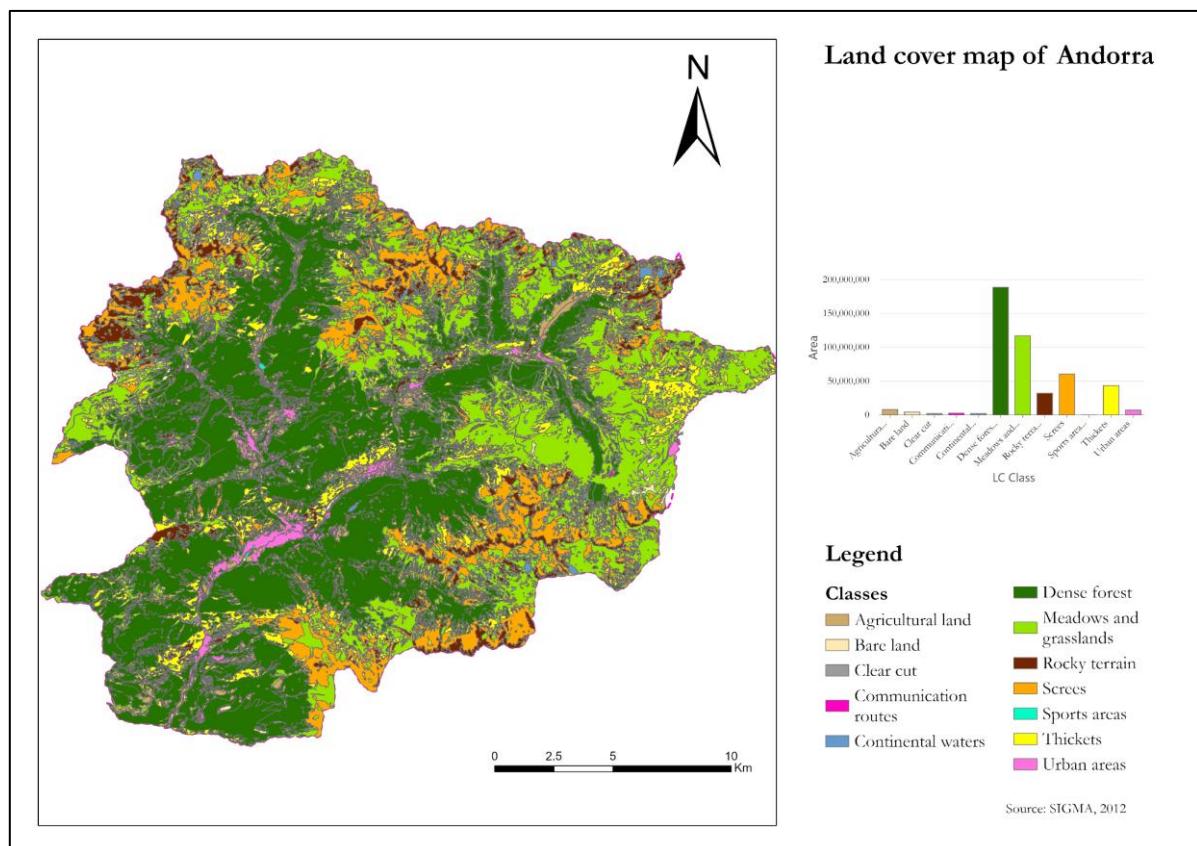
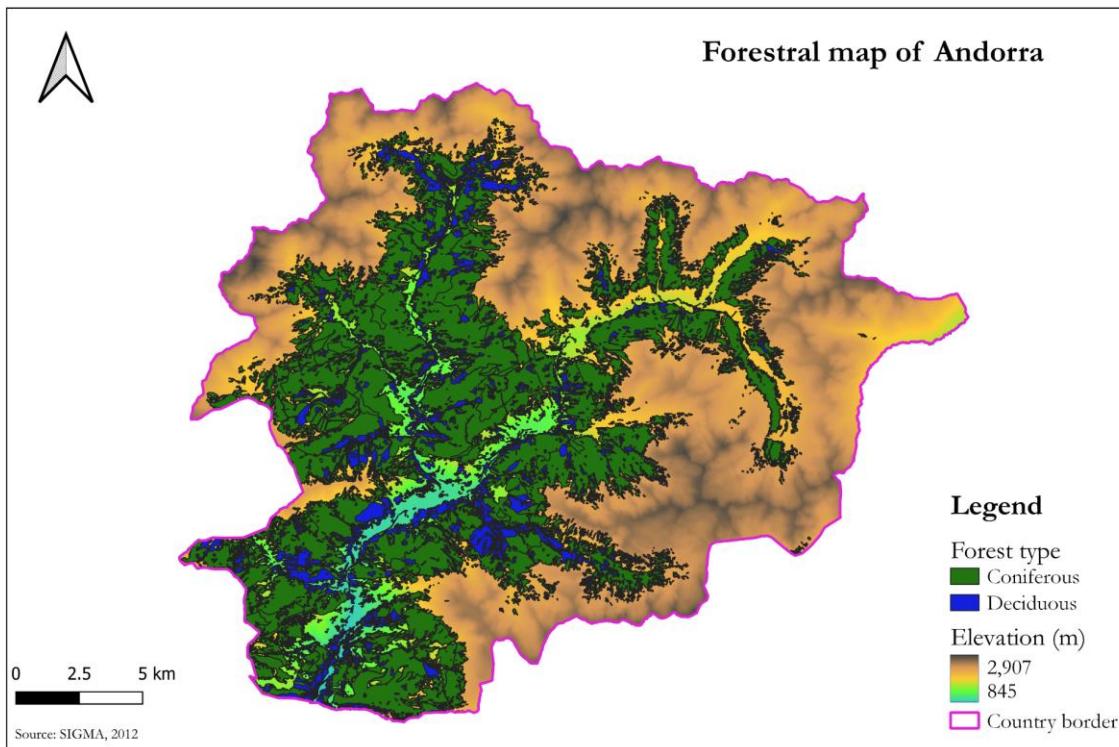


Figure 17. Land Cover map of Andorra.



*Figure 18. Map of forest types distinguished as coniferous or deciduous.*

#### 4.1.2. Data analysis

##### Avalanche inventory

Following the methodology described in section 3.2 a vector layer containing 1,150 initiating and 1,150 non-initiating avalanche zones, in total 2,300 shapes, was created (Figure 19). The non-initiation zones were randomly generated and have surface areas ranging from as small as the smallest initiation area and as large as the median initiation area from the inventory. It needs to be pointed out that the initiation areas, and, thus, the non-initiation zones as well, are more than the 1,122 scars that were initially identified. The reason for this is that for some avalanches, based on their shape and topography of the area where they occurred, it was concluded that they had multiple starting points. Thereof, these extra points were added manually, where it was considered necessary, resulting in 1,150 release areas for 1,122 avalanches.

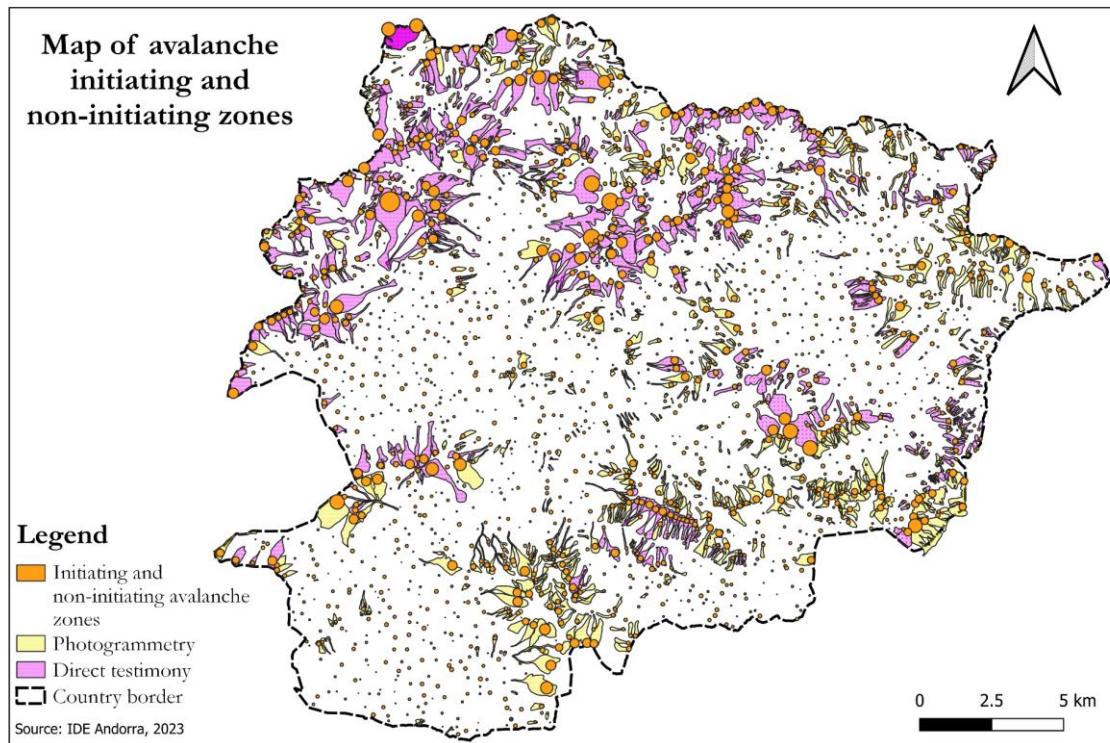


Figure 19. Initiation and non-initiation avalanche zones depicted with orange circles of different diameters.

### Topographical features

Using the DEM and land cover information, the maps of slope angles (Figure 20), aspects (Figure 21) and terrain roughness (Figure 22) were created as explained in the methodology. The curvature map does not provide significant visual insights, and, thus, it is not presented here, although it was used as input to the models. The map of orientations relative to the sun is also omitted from this section because it is solely a biproduct of the aspects map, distinguishing between north, north-east, north-west slope aspects as being, roughly, shady and the rest as being, roughly, sunny. In Figure 22, it can be seen that rocky terrain is mostly encountered in the high peaks, while screes occupy many of the hillslopes at relatively high altitudes.

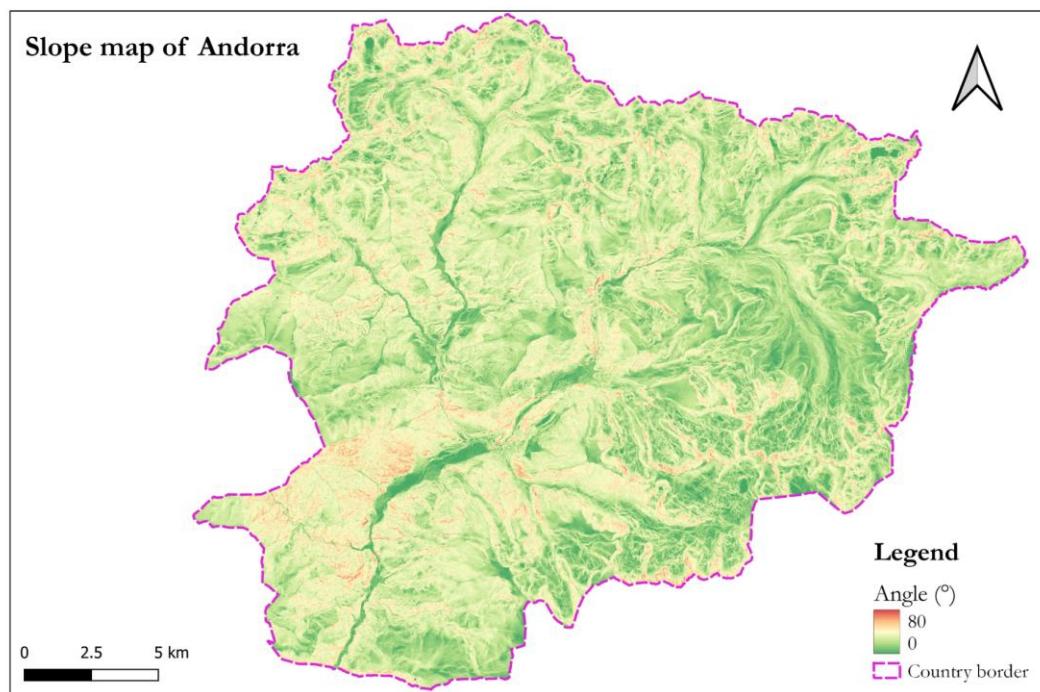


Figure 20. Map of slope angles.

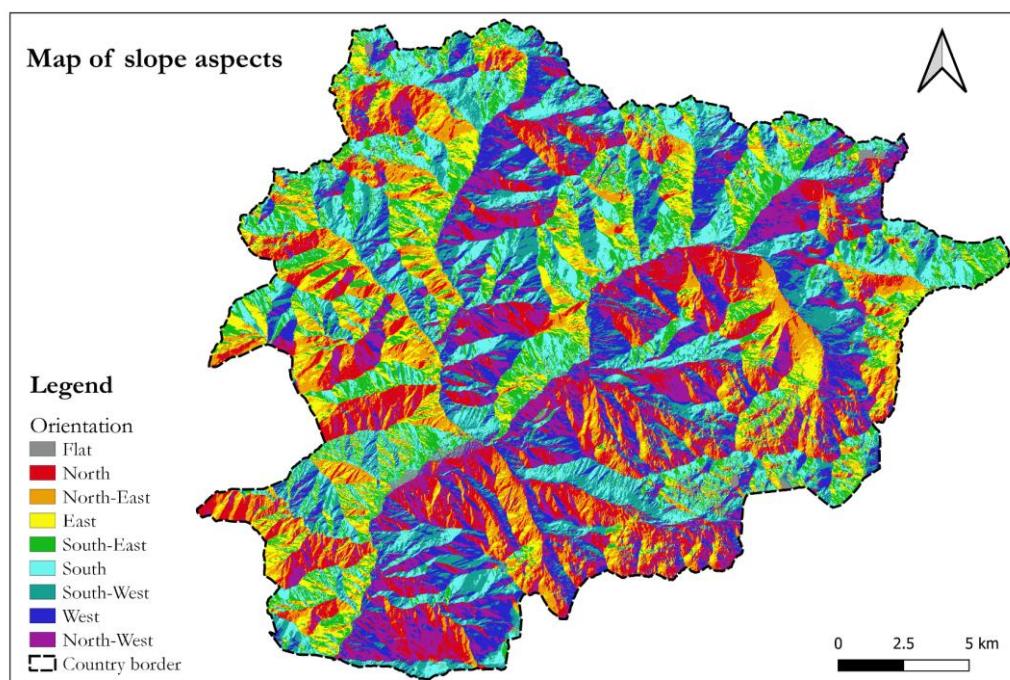
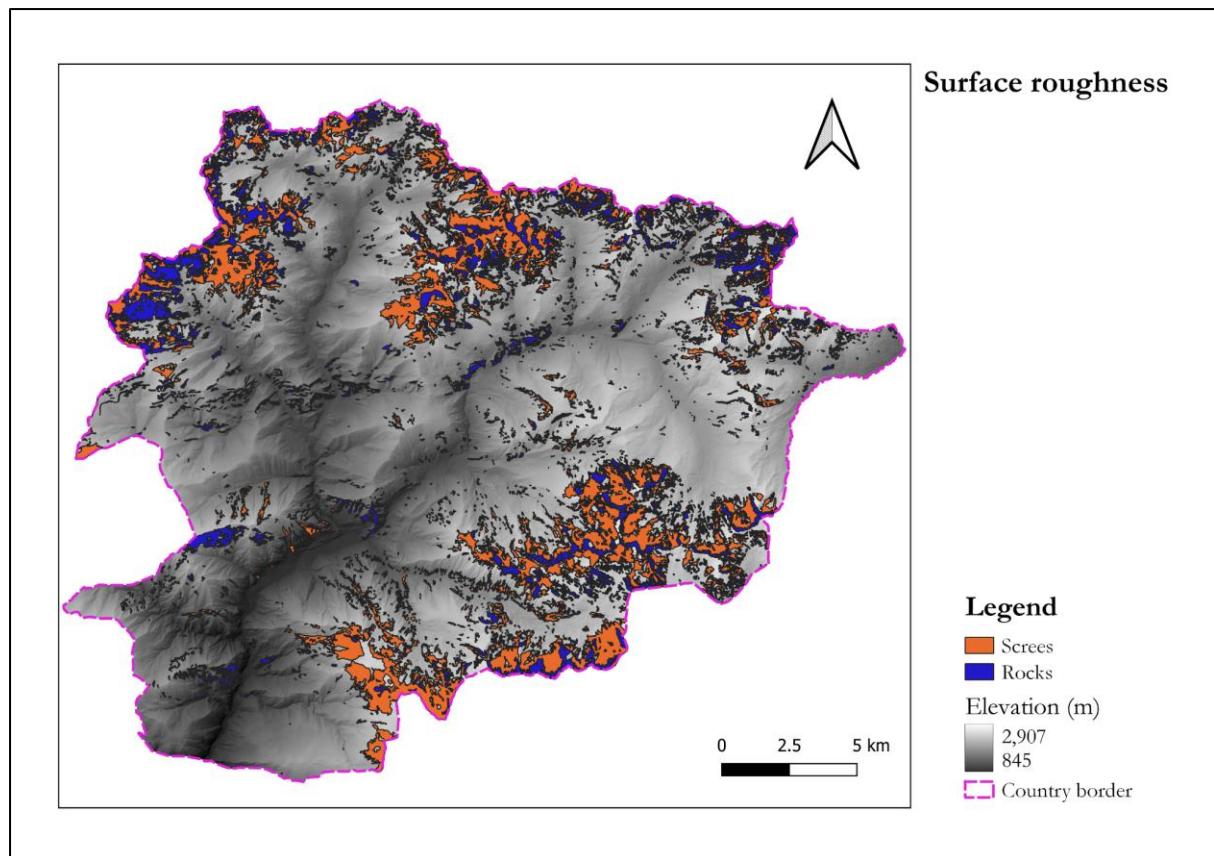


Figure 21. Map of slope aspects.



*Figure 22. Map of surface roughness characterized by rocky terrain and scree-laden slopes.*

#### Landsat NDSI

As described in section 3.2, the NDSI raster is derived using Google Earth Engine. In this case, a threshold value is set equal to 0.4, which filters out the cells that throughout the whole image collection had, on average, an NDSI value lower than the threshold. As a result, in the final raster there are pixels without values (white color cells), which represent areas that are not snow covered, and pixels with NDSI values greater than 0.4 (Figure 23). As mentioned before, higher NDSI values indicate higher probability of snow presence.

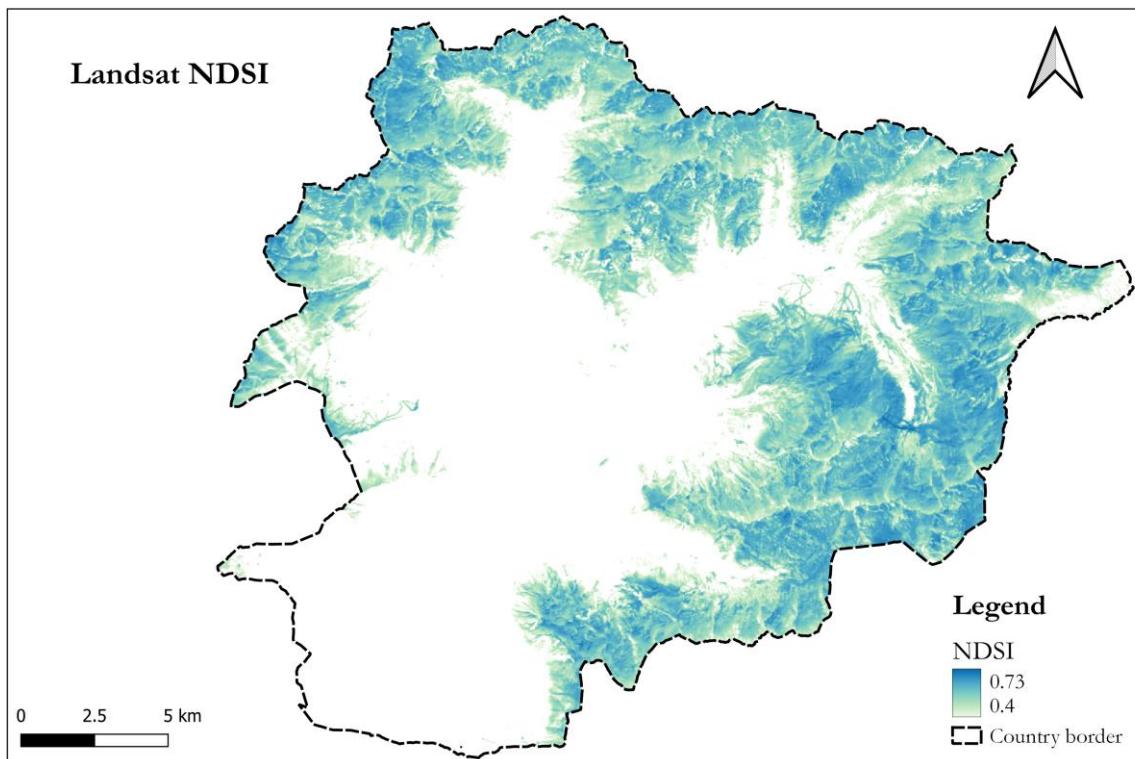
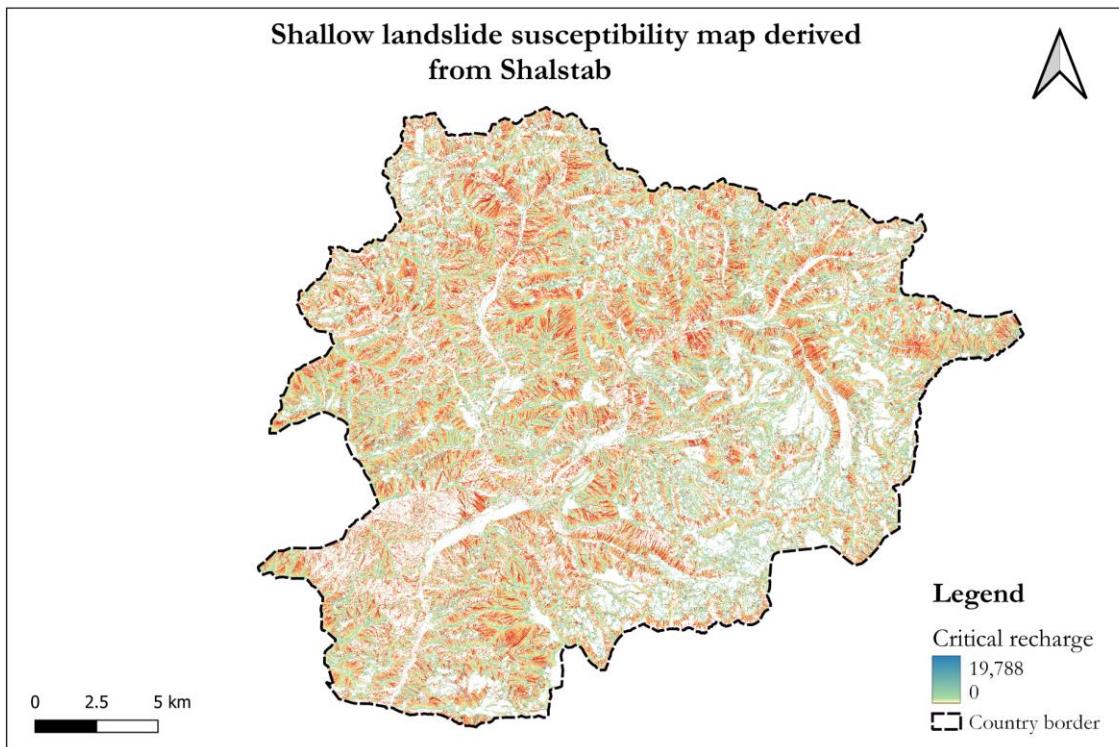


Figure 23. Map of snow cover in Andorra depicted as NDSI values.

#### Shalstab critical recharge

The set up of the Shalstab tool in SAGA requires the slope to be converted into rads (multiplication of the previously acquired slope raster with  $\frac{\pi}{180}$  in the Raster Calculator). The bulk density is considered equal to  $1.6 \frac{g}{cm^3}$  (default value), the hydraulic conductivity (here, considered as transmissivity per snow thickness) equal to  $2.7 \frac{m}{h}$  (default value), and snow instead of soil thickness equal to 1 m (default value). The acquired map is presented in Figure 24.



*Figure 24. Map of critical recharge with regard to shallow landslide susceptibility acquired as output of the Shalstab model.*

#### Weather data

In this section, initially, some indicative results of the performed explorative analysis of the weather data are demonstrated. Finally, the selected representative values extracted from the timeseries of each variable in each station are presented. Part of the analysis was, first of all, the cleaning of the datasets against invalid entrances. Figure 25 showcases an example of the Borda Vidal station with invalid rainfall records (negative values), which were removed after detection. Outlier detection and removal was only performed for the temperature observations of the stations. Figure 26 and Figure 27 depict two such cases of detected outliers that had to be removed. Outlier removal of other datasets, like rainfall or snow, was avoided, due to limited knowledge of the specificities of the local weather. It was preferred not to lose any potential storm events, than perform the outlier test for these datasets.

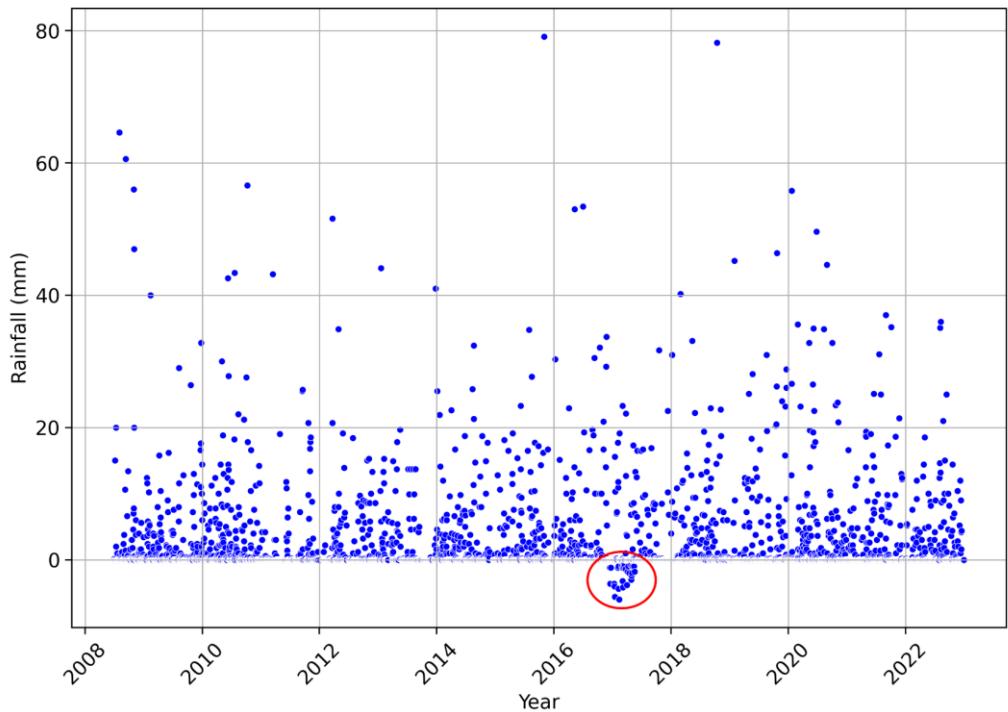


Figure 25. Example of invalid measurements of rainfall in the Borda Vidal station.

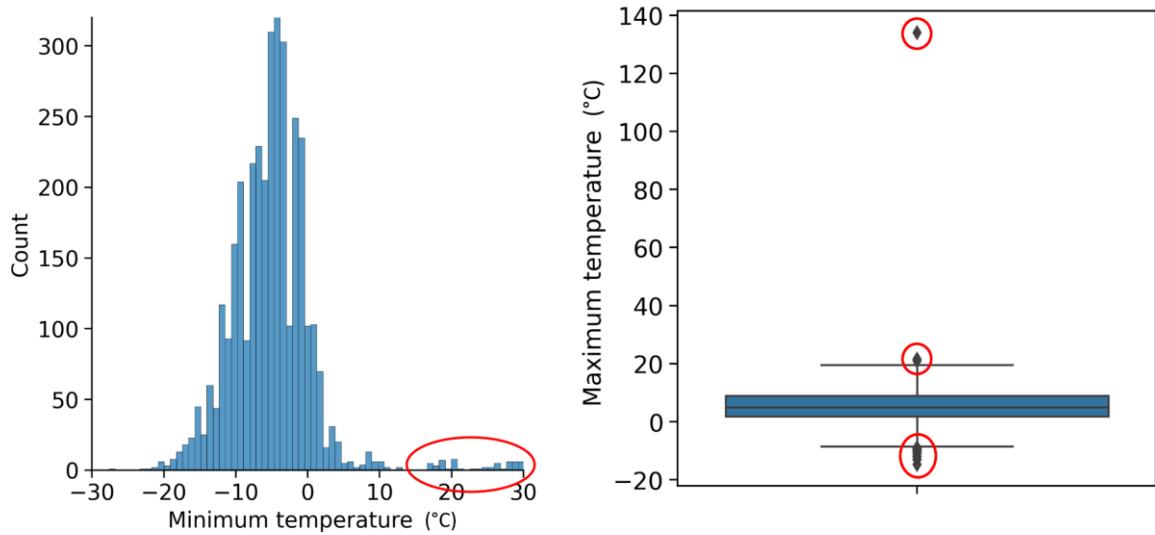


Figure 26. Example of outlier detection of the minimum temperature in the station of Soldeu using the histogram.

Figure 27. Example of outlier detection of maximum temperature in the station of Pal using the boxplot.

Subsequently, although the data have, theoretically, passed a quality check, a homogeneity test was performed, which revealed that inhomogeneity indeed exists in the data (Figure 28 and Figure 29). However, the metadata of the stations have not been recorded throughout the years of their operation, and, as a result, it was impossible to assess which of these heterogeneities (breaks in the double mass curves) are due to non-climatic causes and need correction. Therefore, no relevant actions were taken.

Nevertheless, the double mass curve method revealed another problem with the data. The measurements of rainfall in Soldeu and old snow in Ransol, enclosed in red polygons in Figure 28 and Figure 29, respectively, were found to be many magnitudes of order lower than the rest of the stations. The measurements were considered erroneous, and, thus, these datasets were excluded completely from the analysis. The stations with such problems in their records were mostly the winter stations operated by the ski resorts. In fact, according to information received by the Meteorological Service of Andorra, these data are recorded with not so high standards of quality. Sometimes different instruments are used, the observation site has changed multiple times in search for the locations with most snow, and, often, the observations have been scaled up. All these facts, accompanied with the lack of documentation and metadata of the stations' history, underline the quality issues of the available datasets, which is common problem throughout the Pyrenees, and act as a reminder to use the data skeptically.

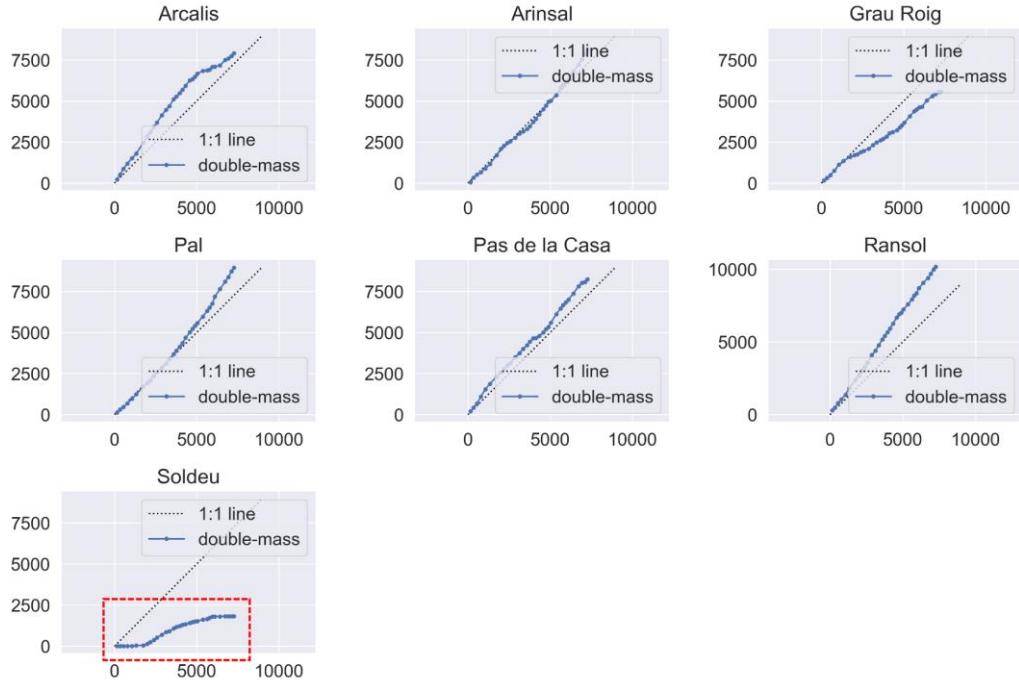


Figure 28. The double mass curve method applied to the winter stations' observations. The y-axis represents the cumulative precipitation of the individual station (in mm) and the x-axis denotes the cumulative precipitation of all the stations (in mm).

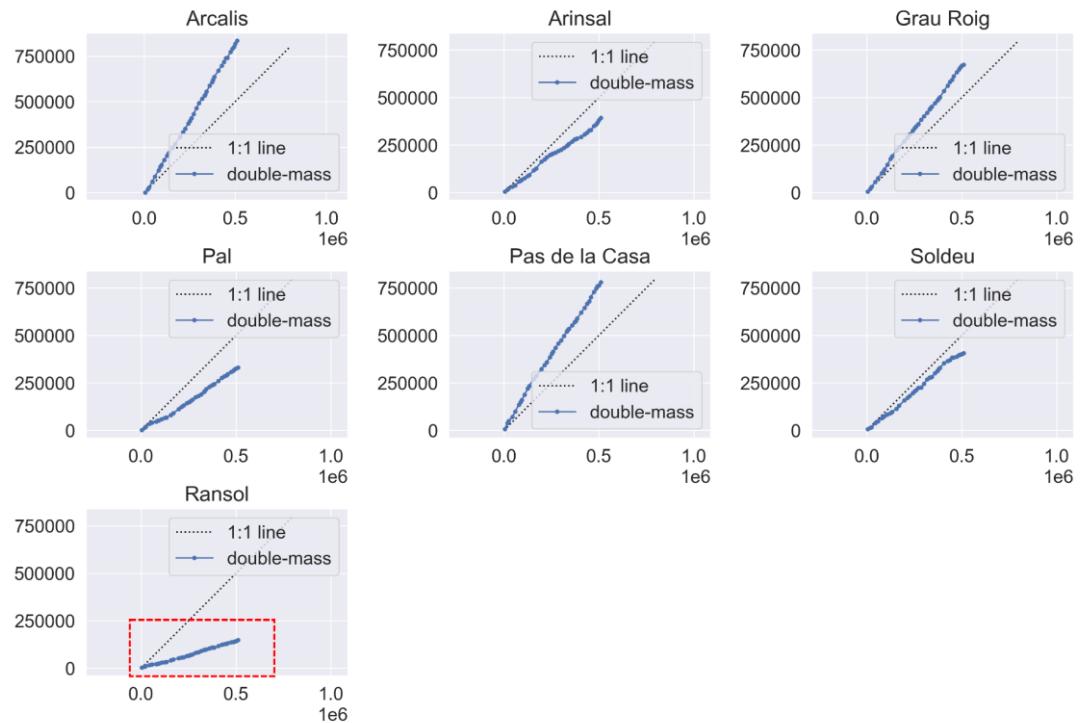


Figure 29. The double mass curve method applied to the winter stations' observations. The y-axis represents the cumulative old snow of the individual station (in cm) and the x-axis denotes the cumulative old snow of all the stations (in cm).

In an effort to identify extremes in the rainfall timeseries, the peak over threshold method was applied and the values over the threshold of 95% were considered as extreme events. Figure 30 provides an example of the application of the method for the La Comella station. Figure 31 depicts the prevalent wind direction and speed in different parts of the country.

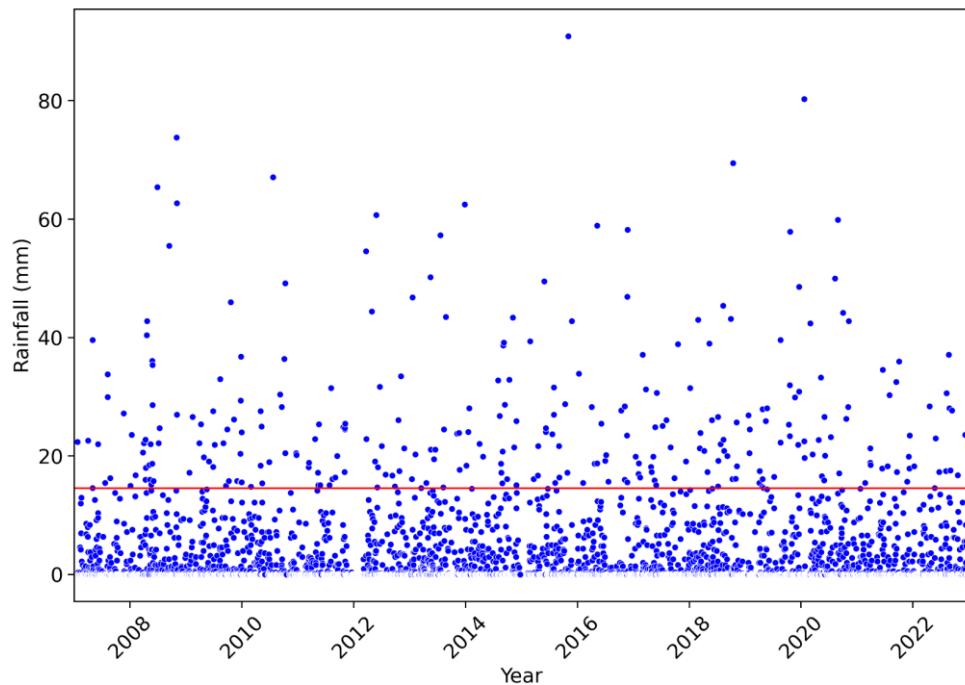


Figure 30. Scatterplot of precipitation in the La Comella station. The red horizontal line represents the 95% threshold for identifying the extremes.

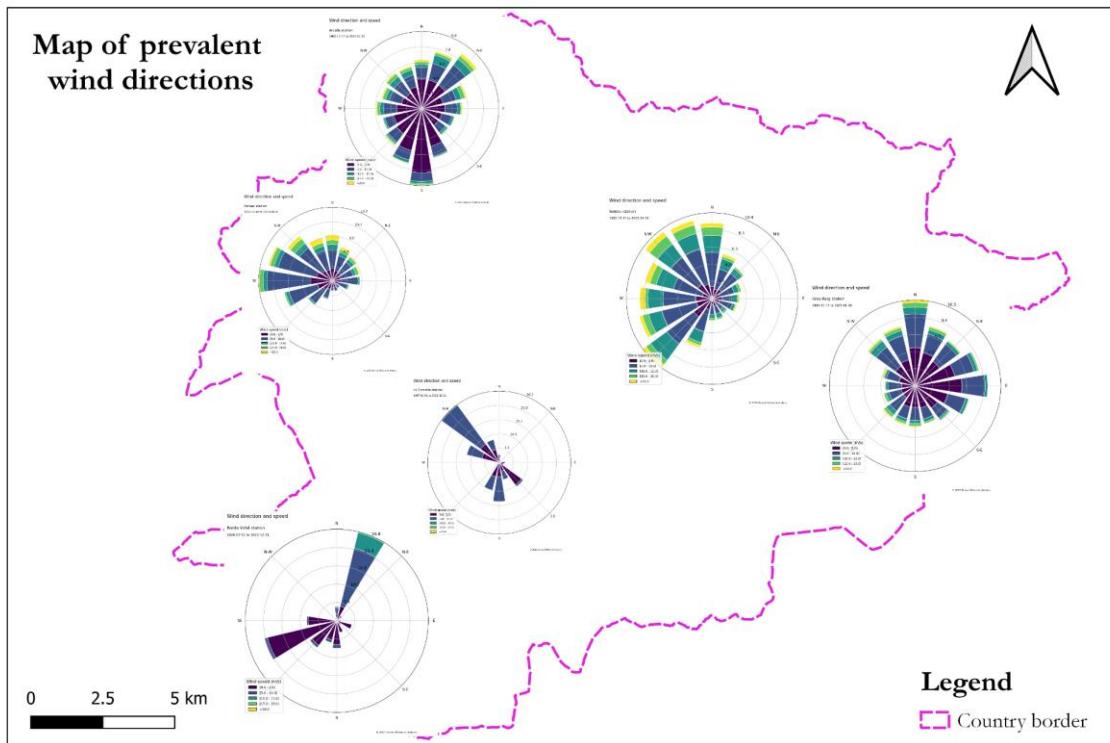


Figure 31. Map of the prevalent wind conditions (speed and directions) represented as wind roses in different parts of Andorra.

Complementary to the wind analysis, the relationship between wind and snowdrift was explored through a Chi-square test for hypothesis testing. As described earlier, wind direction and snow drift direction are expected to be correlated. The Chi-square test confirmed this assumption for all stations where snow data were available, with very low p-values (lower than 0.05) indicating that the relationship between the two variables is statistically significant. Figure 32 shows the conditional distribution of snow drift intensity given wind direction in the station of Arcalis. From the graph it can be concluded that, in this particular station, the most intense snow transport was observed when the wind was blowing from the north, north-west, north-east, east or variable directions.

Lastly, Pearson correlation analysis was performed between all the numerical variables. As an example, Figure 33 clearly showcases, as expected, a correlation between maximum temperature and daily range of temperatures in the station of Pal.

After the above-mentioned analysis was finalized, as indicated in the methodology, representative values were selected for all the variables, in order to be used for their interpolation. Only exception

were the data of snow drift, because the categorical/coded entries available in the records were found hard to be sampled for a representative value. Moreover, the information of snow transport, the way it was attained, is very characteristic to the observation site and, thus, interpolating would incur large errors. Table 3 provides a summary of the values selected as representative for each variable.

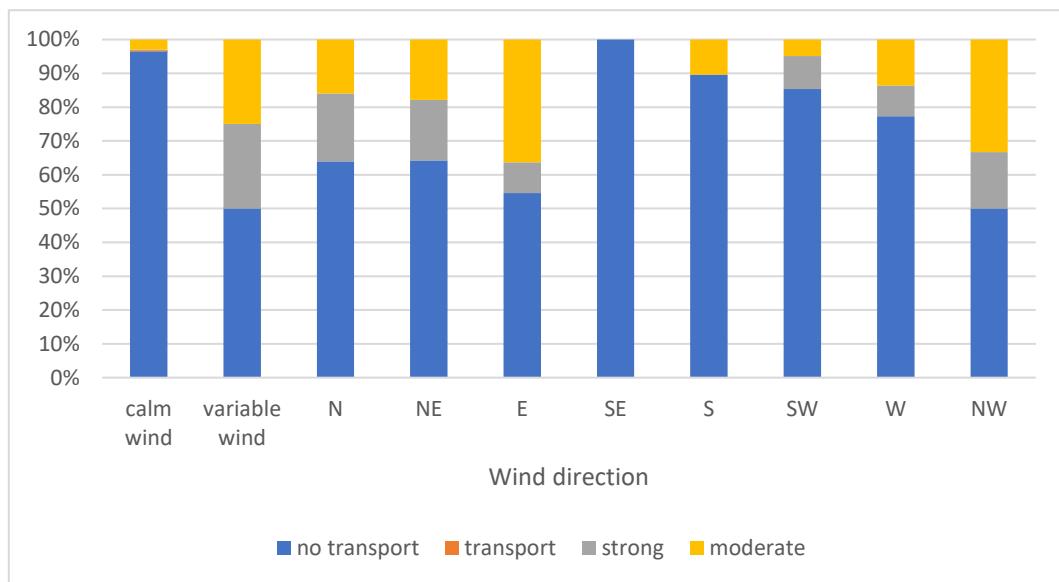


Figure 32. Conditional distribution of snow drift intensity given wind direction in the station of Arcalis.

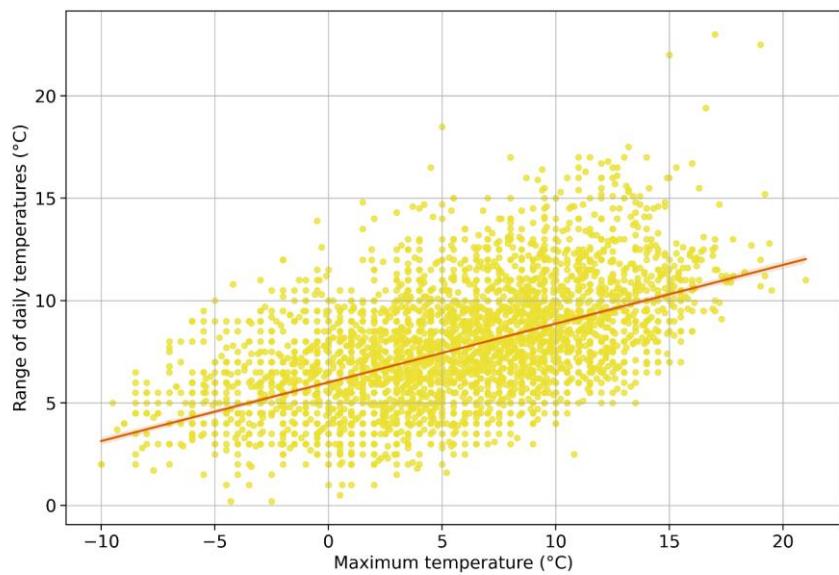


Figure 33. Correlation of maximum temperature and range of daily temperatures in the station of Pal.

*Table 3. Representative values selected for each weather variable to be used for their interpolation.*

Weather variable	Representative value
Rainfall	Average of timeseries
Accum. rainfall (48 h, 72 h)	Average of timeseries
Extreme rainfall	Average of timeseries
Monthly rainfall	Average of timeseries
Temperature (min, max, mean)	Average of timeseries
Temperature range	Average of timeseries
Old snow	Average of timeseries
New snow (24 h)	Average of timeseries
Accum. snow (48 h, 72 h)	Average of timeseries
Monthly snow	Average of timeseries
Wind speed (max, mean)	Maximum/Average in the prevalent wind direction
Wind direction	Most usual (mode)

#### 4.1.3. Database construction

##### Interpolations

As already mentioned, the representative values were interpolated using the IDW method and two of the produced maps are presented below as examples (Figure 34, Figure 35).

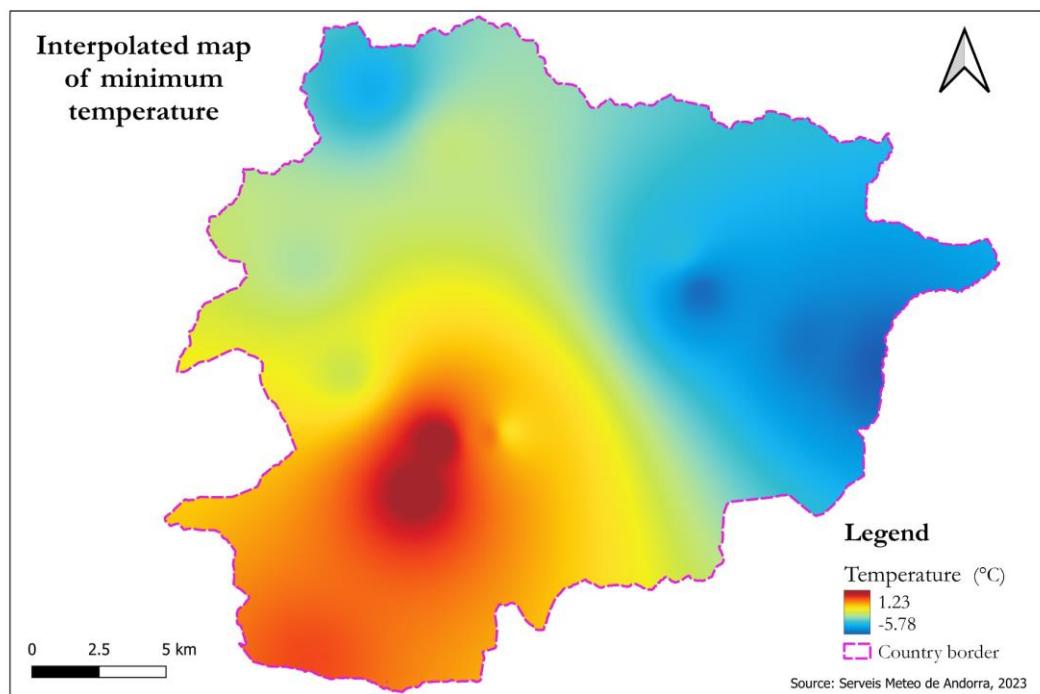


Figure 34. Interpolated minimum temperature.

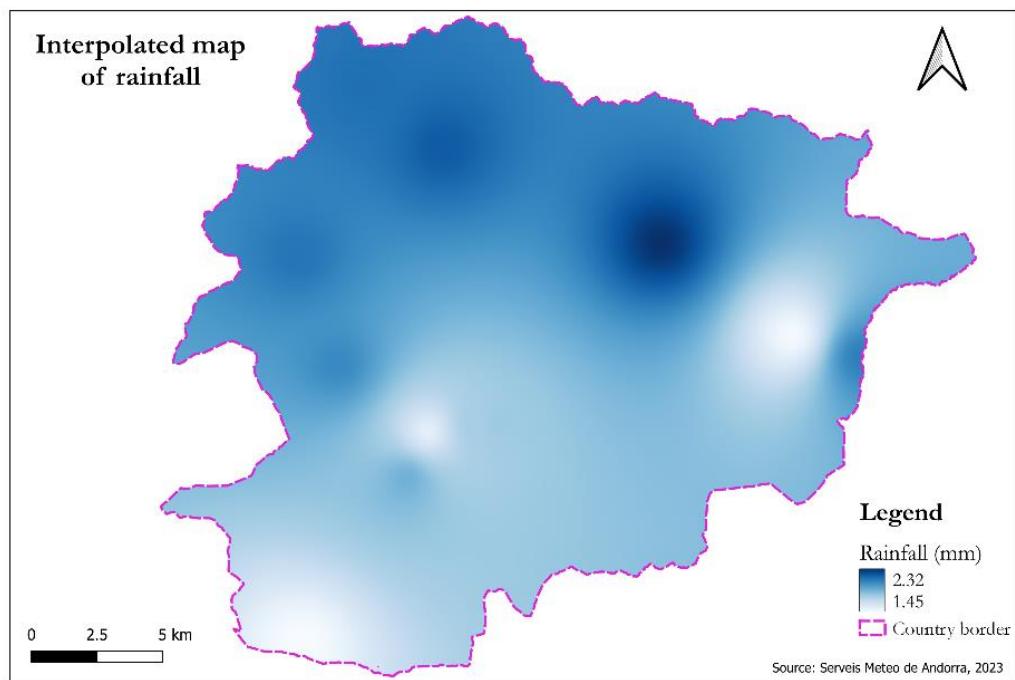


Figure 35. Interpolated rainfall.

## Zonal sampling

Table 4 shows a summary of all the rasters, their data type, and the statistics that were used for their sampling. The colors of the rows correspond to the colors of each process as shown in Figure 6.

*Table 4. Zonal statistics used for the creation of the database.*

Raster layer	Data type	Statistic
Rainfall	Continuous	Mean
Accum. rainfall (48 h, 72 h)	Continuous	Mean
Extreme rainfall	Continuous	Mean
Monthly rainfall	Continuous	Mean
Temperature (min, max, mean)	Continuous	Mean
Temperature range	Continuous	Mean
Old snow	Continuous	Mean
New snow (24 h)	Continuous	Mean
Accum. snow (48 h, 72 h)	Continuous	Mean
Wind speed (max, mean)	Continuous	Mean
Wind direction	Categorical	Majority
Elevation	Continuous	Mean
Slope	Continuous	Maximum
Curvature planform	Continuous	Mean
Curvature profile	Continuous	Mean
Aspect	Categorical	Majority
Orientation	Binary	Majority
Forest	Binary	Majority
Rocky terrain	Binary	Majority
Screes	Binary	Majority
NDSI	Continuous	Mean
Critical recharge	Continuous	Mean

Categorical data can have values from a set of predefined labels (e.g., slope aspects can be north, south etc.). Binary data are data that can either be True or False (1 or 0). For example, a single cell can either be forest land or not; it cannot be both at the same time. Continuous data are numeric and can range independently of scale. For instance, elevations reach values above 2,000, while NDSI can be 1 at maximum. For the categorical and binary data, the majority of labels of the cells that overlap with the initiating/non-initiating zones is considered, and for most of the continuous rasters the average. However, a special note needs to be made for the slope. According to D. McClung & Schaefer (1993), the critical incline for avalanche release refers to the steepest part of a hillslope not the average. This is the reason why the maximum slope instead of the average is extracted for each zone.

The final product is a matrix of 29 columns (including the target) and 2,300 rows. The feature space consists of the first column of the Table 4 plus the elevation.

## 4.2. Phase 2

### 4.2.1. Database preprocessing

The first step was to convert the wind direction and aspect into string/text format, in order to follow the labeling convention of Table 1 (second column). Then, after One-hot encoding was applied, the number of columns of the database was increased, because for both these variables a separate binary column was created for each one of its categorical labels. In addition, since the information of snow drift provided by the Meteorological Service of Andorra was excluded from the study, the alternative of potential snow transport calculation was explored, as explained in the methodology, adding one more column to the database. After this step of feature engineering, the columns of the database were 44 (including the target).

Subsequently, explorative data analysis was performed, revealing correlations between the independent variables of the database, but also relationships of the features with the target. In the correlation heatmap depicted in Figure 36, some expected relationships, dictated by basic physics, were confirmed, but a few more interesting and not so expected interrelations were also spotted. For instance, it is no surprise that the products of two-day rain or three-day snow are highly correlated with daily rain and snow, respectively (areas of concentrated white tiles along the diagonal in the heatmap). Additionally, snow and temperature are inversely correlated, same as elevation and temperature, while the snow thickness and elevation are significantly correlated. The high correlation between the snow cover and the elevation, and the inverse correlation with

temperature was also an expected result. The most interesting result from this figure was the significant inverse correlation between the gust wind speed and temperature.

In explorative data analysis the histogram is a powerful tool for gaining insights about the data. Figure 37 and Figure 38 are two examples of histograms that were plotted for two of the features, based on their target class. On the one side, Figure 37 reveals the impact that increasing elevation has on the susceptibility to avalanche. On the other side, whether the orientation of a slope is towards the east seems to be irrelevant to avalanche release, indicating that this predictor is weak, and, thus, unnecessary for the models.

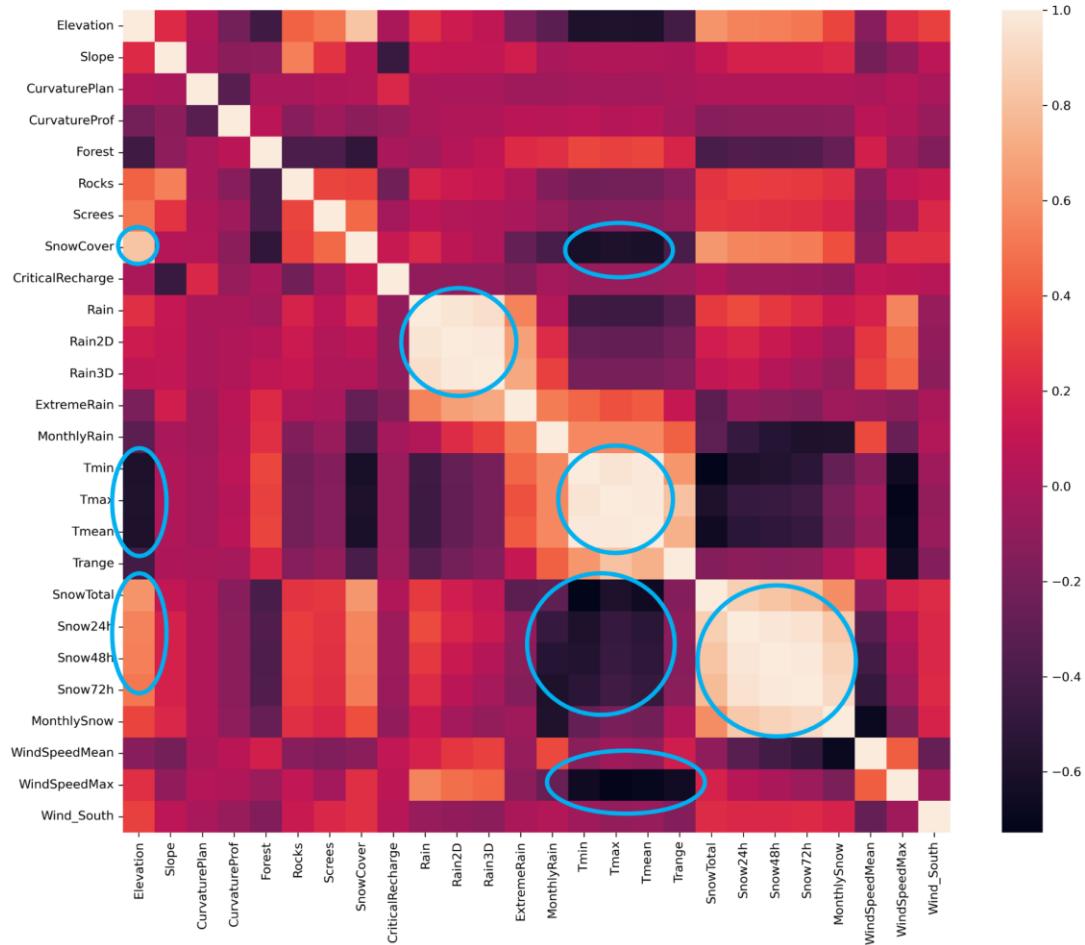
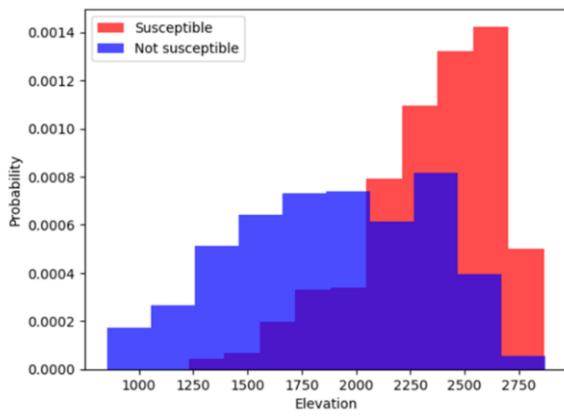
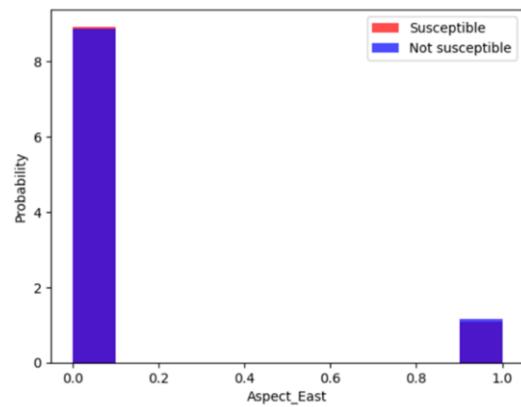


Figure 36. Heatmap depicting correlations between the independent features.



*Figure 37. Histogram representation of the distribution of elevations in relation to susceptibility to avalanche.*



*Figure 38. Histogram representation of the binary feature of east-oriented slopes in relation to susceptibility to avalanche.*

The histogram was the main criterium for selecting the most relevant features. As additional tools, the chi-2 test, the mutual information method (Figure 39) and the feature importance provided by the Extra Trees classifier confirmed that all of the aspect and wind direction labels had very low or almost no relevance to avalanche initiation. However, this result raised concern with regard to how those variables were categorized in the beginning of the preprocessing. Indeed, splitting the data into eight categories caused some labels to be very unrepresented (e.g., wind direction from the North in Figure 40). To test if this was the actual reason of these labels not scoring as high in the feature importance, another categorization was tried using only four labels (North, South, East, West) for both aspect and wind direction. Nevertheless, no major differences were observed. Therefore, the irrelevant features were removed from the database, with the south wind being the only exception, because it showcased a slightly different tendency between classes (Figure 41) and, also, it was the most correlated direction to potential snow transport (Table 5).

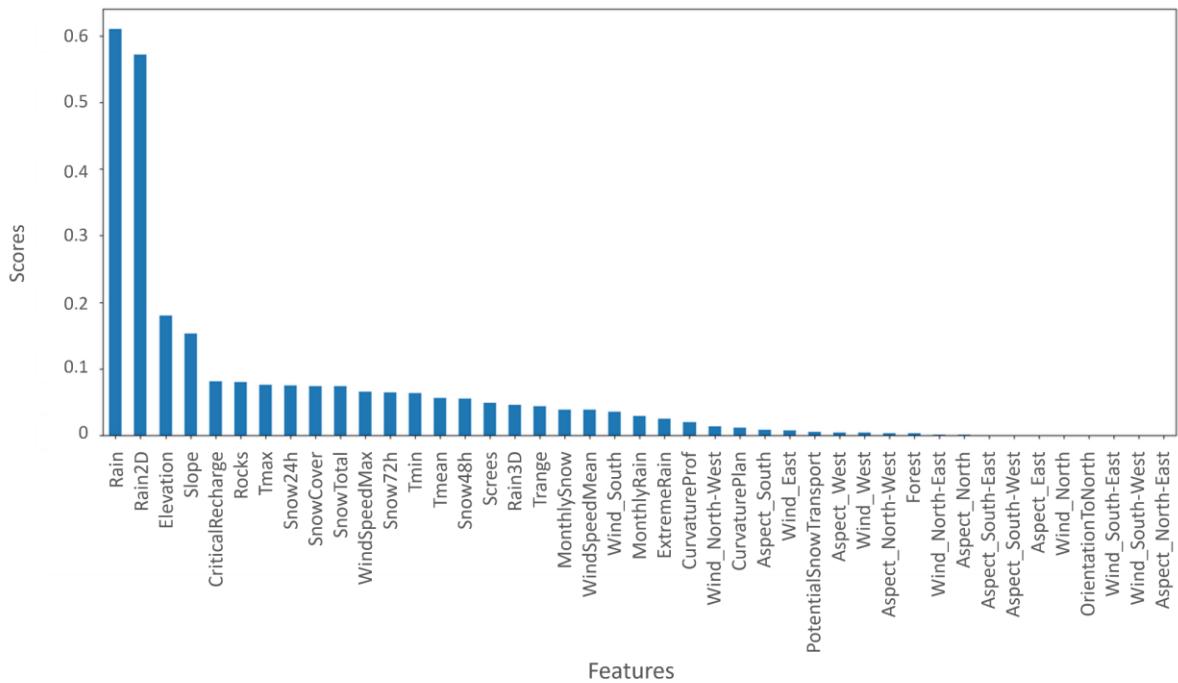


Figure 39. Graph of mutual information of the features with the target class.

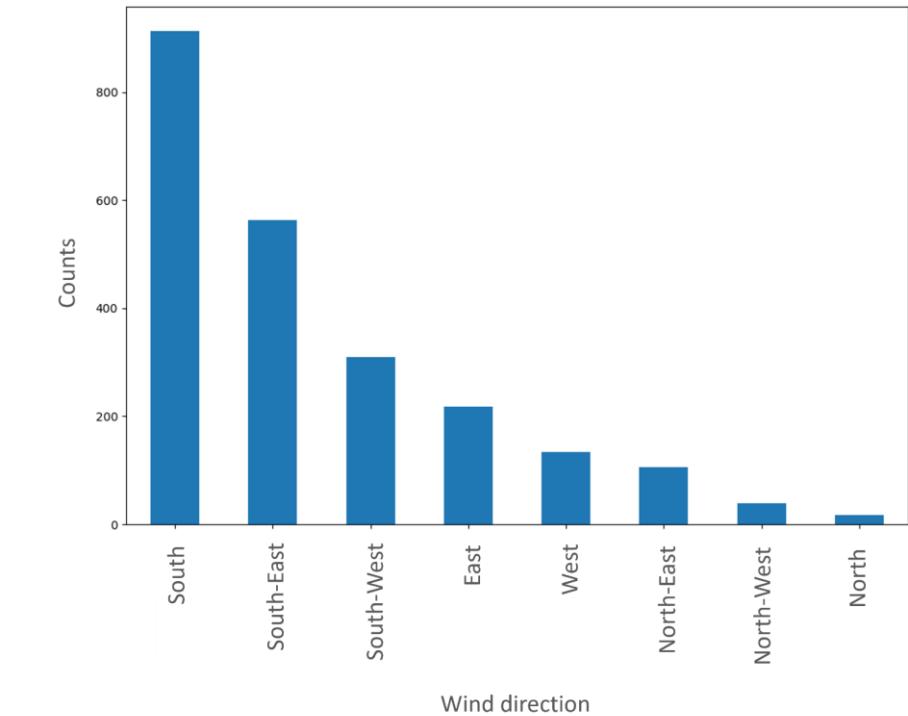


Figure 40. Count plot of the different wind directions in the database.

Table 5. Contingency table of wind direction and potential snow transport.

PotentialSnowTransport	0.0	1.0
WindDirection		
East	151	67
North	11	6
North-East	106	0
North-West	32	7
South	578	335
South-East	381	183
South-West	231	78
West	90	44

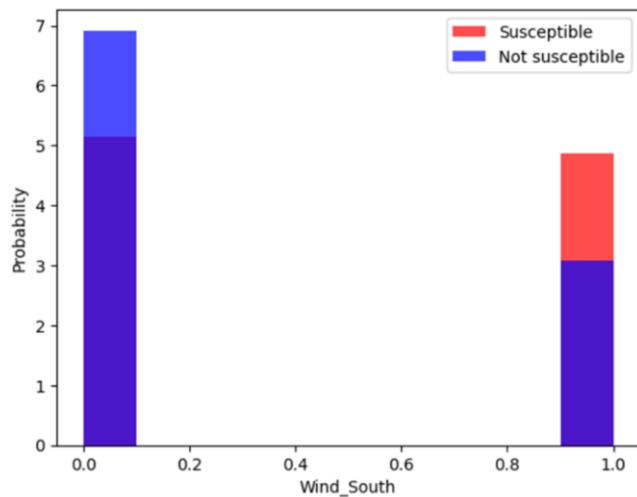


Figure 41. Histogram representation of the binary feature of south wind in relation to susceptibility to avalanche.

#### 4.2.2. Model building and deployment

In this section, some important snapshots from the process of training-validation-optimization are presented, as well as the validation accuracies of the models. As mentioned previously, the selected models were developed twice, based on different versions of the database but following the same procedure, in order to test the significance of certain predictors to the performance. Here, only the results obtained using the entire database are presented. The contribution of the different types of predictors to the performance is going to be demonstrated in the next chapter with the testing of both versions of the models on “unseen” data.

#### Decision Tree

The model was initially fitted to the training data using the “DecisionTreeClassifier” estimator from sklearn and the fit function as shown in Figure 42 below. The “random\_state” parameter was used in all of the model builds for reproducibility of the results by setting it equal to any integer number.

```
clf_dt = DecisionTreeClassifier(random_state=42)
clf_dt = clf_dt.fit(X_train, y_train)
```

Figure 42. Initial fit of the Decision Tree.

The created tree was clearly overfitted, having a very deep structure. As mentioned earlier, overfitting means bad generalization of the model to “out-of-sample” data, and, thus, pruning of the tree was performed by optimizing the parameter  $\alpha$  of the cost-complexity algorithm. Figure 43 demonstrates the relationship between  $\alpha$  and the achieved accuracy of the DT for the training and the validation sets. In other words, it provides a nice visualization of the “bias-variance tradeoff” and how it could be balanced through pruning. As expected, the initial accuracy on the training dataset is perfect, because the model fits exactly to the data. As the pruning proceeds, the training accuracy recedes more and more. The validation accuracy, on the other hand, seems to fluctuate around 85% for small values of  $\alpha$  and to stabilize after  $\alpha \approx 0.02$ . The falling limbs of both graphs when  $\alpha = 0.08$ , denote that when too much pruning is applied, the model fails to capture the actual relationship of the data (high bias) and the accuracy drops vertically for both the training and the validation set.

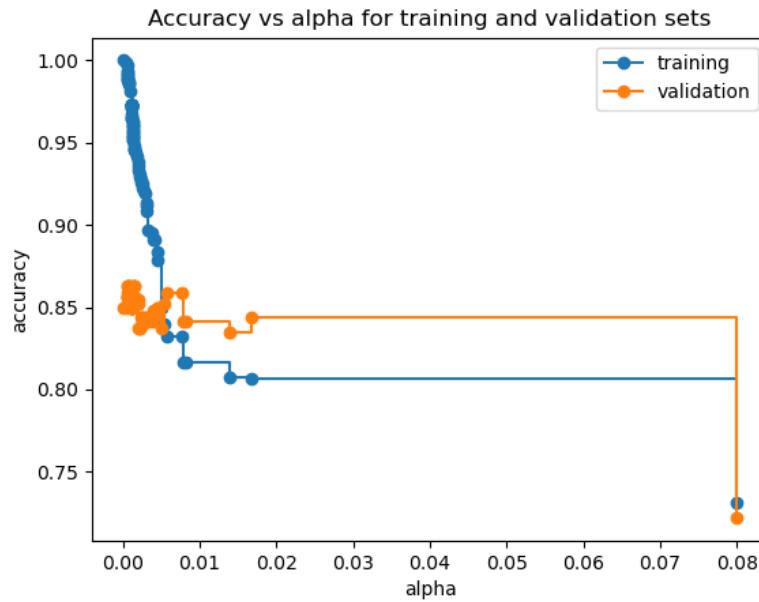


Figure 43. Accuracy of both the training and validation sets in relation to the alpha pruning parameter.

Initially, a value of  $\alpha = 0.014$  was tested. The accuracy of the model on the validation set is depicted in the confusion matrix in Figure 44. The model seems to perform pretty well in terms of misses (22 missed events), but not so ideally considering false alarms (51 events). When observing the pruned tree (Figure 45), it can be seen that some of the leaves have high impurity scores, meaning high uncertainty in the predictions. This is not favorable, and, therefore, a higher value of  $\alpha$  was searched, in order to apply more pruning and get rid of the impure leaves.

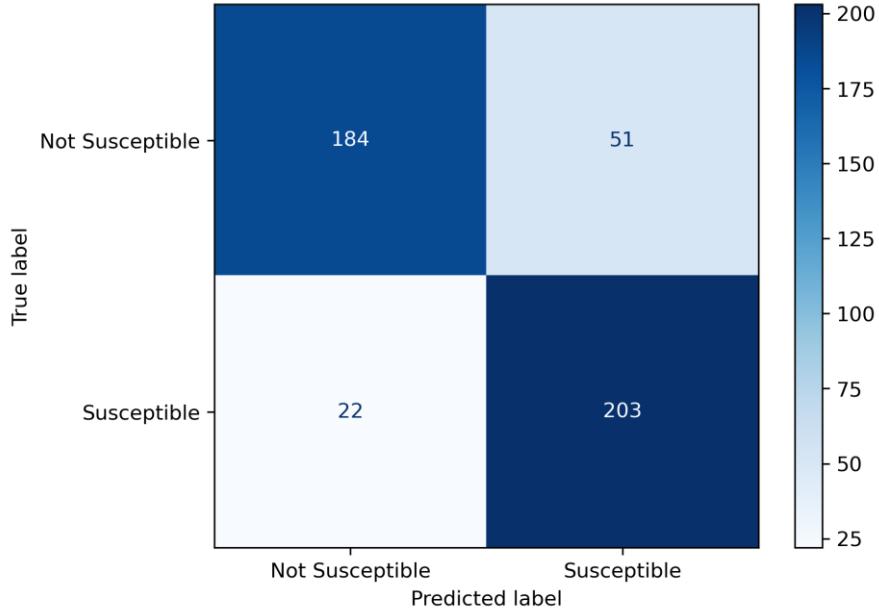


Figure 44. Confusion matrix of the pruned DT with  $a = 0.014$  on the validation set.

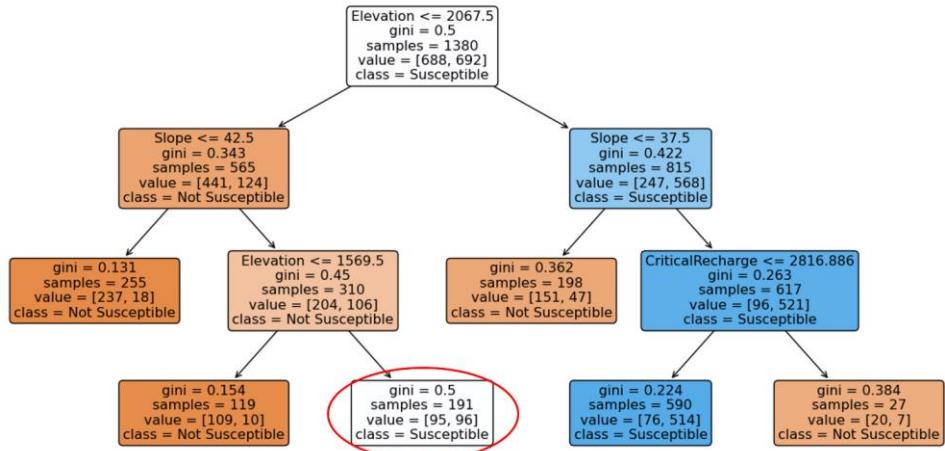
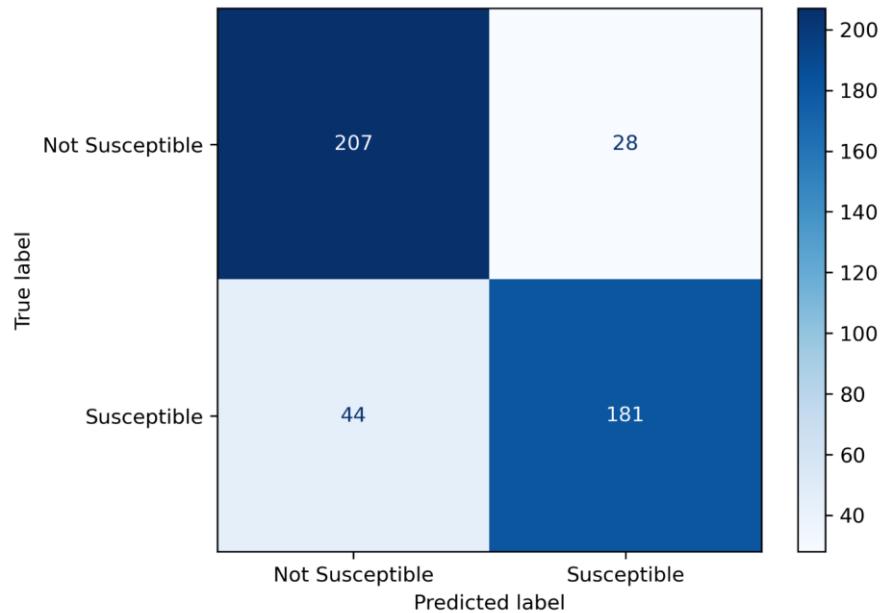


Figure 45. Plot of the pruned DT with  $a = 0.014$ .

When setting  $\alpha = 0.017$ , the produced confusion matrix looked like the one in Figure 46. It can be noticed that more pruning of the tree predicted more misses and less false alarms than before. This is an interesting result and showcases how sensitive the final predictions might be to the optimization procedure. In this case, having less missed events was prioritized over reducing the false alarms, because a missed event of a snow avalanche can cause high damages and even loss of

lives. On the contrary, when in the hands of local experts, that know the study area well, false alarms can be examined with a more critical eye, and areas that, for example, do not seem prone to avalanche release, based on experience but also from on-site tests, can be excluded from the high susceptibility zones during the mapping process.



*Figure 46. Confusion matrix of the pruned DT with  $\alpha = 0.017$  on the validation set.*

Since further optimization did not provide better results,  $\alpha = 0.014$  was finalized as the pruning parameter. The validation accuracy of the DT was 84%, the recall score 90% and the precision score 80%.

### Random Forest

The Random Forest was trained using the “RandomForestClassifier”. Because it is a greedy algorithm and optimization with “GridSearchCV” was very slow, the option of “RandomizedSearchCv” was preferred in this case (Figure 47). Figure 48 depicts the training of the model using the optimized parameters, and Figure 49 the acquired confusion matrix on the validation set.

```

model = RandomForestClassifier(n_jobs = -1)

parameters = {'min_samples_split': sp_randInt(2, 8),
              'criterion':('gini', 'entropy'),
              'n_estimators': sp_randInt(10, 200),
              'max_depth': sp_randInt(5, 8)}

rf_grid_random = RandomizedSearchCV(estimator=model, param_distributions=parameters, cv=10, n_iter=100, n_jobs=-1)

rf_grid_random.fit(x_train, y_train)

print(rf_grid_random.best_estimator_)
print(rf_grid_random.best_score_)
print(rf_grid_random.best_params_)

```

Figure 47. Optimization of hyperparameters of the Random Forest using the “RandomizedSearchCV” class of sklearn.

```

clf_rf = RandomForestClassifier(criterion='entropy', max_depth=7, min_samples_split=4,
                                n_estimators=139, n_jobs=-1, random_state=42)
clf_rf.fit(x_train, y_train)

```

Figure 48. Training of the optimized Random Forest.

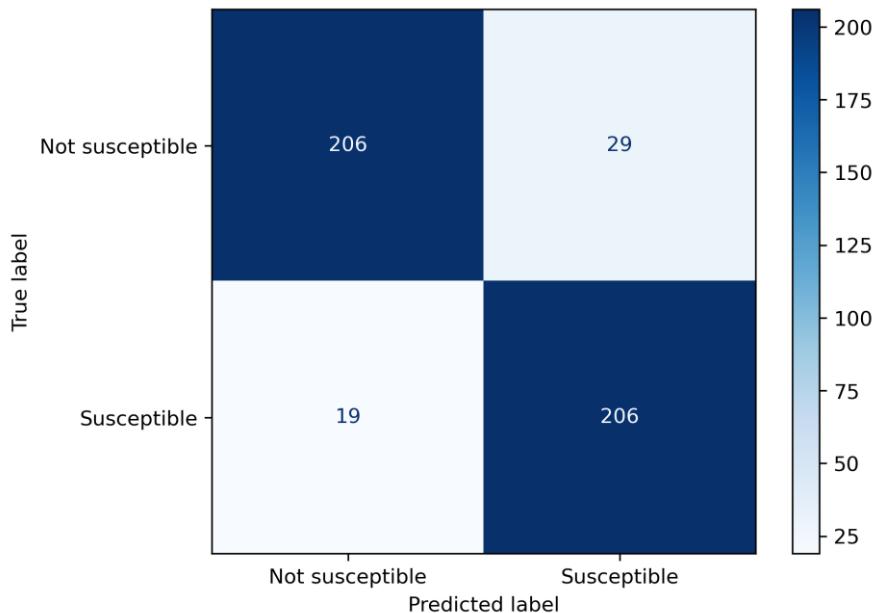


Figure 49. Confusion matrix of the optimized Random Forest on the validation set.

When compared to the DT, the Random Forest provided improved predictions both in terms of misses and false alarms, with an overall accuracy of 90%, recall score 92% and precision score 88%. The Random Forest class in sklearn offers several tools that enable exploring and visualizing the built trees but also the importance of features in terms of improving the performance. The bar plot in Figure 50 depicts a ranking of the features based on their importance/contribution to the

performance of the optimized Random Forest. Slope and elevation appear to be the most important features.

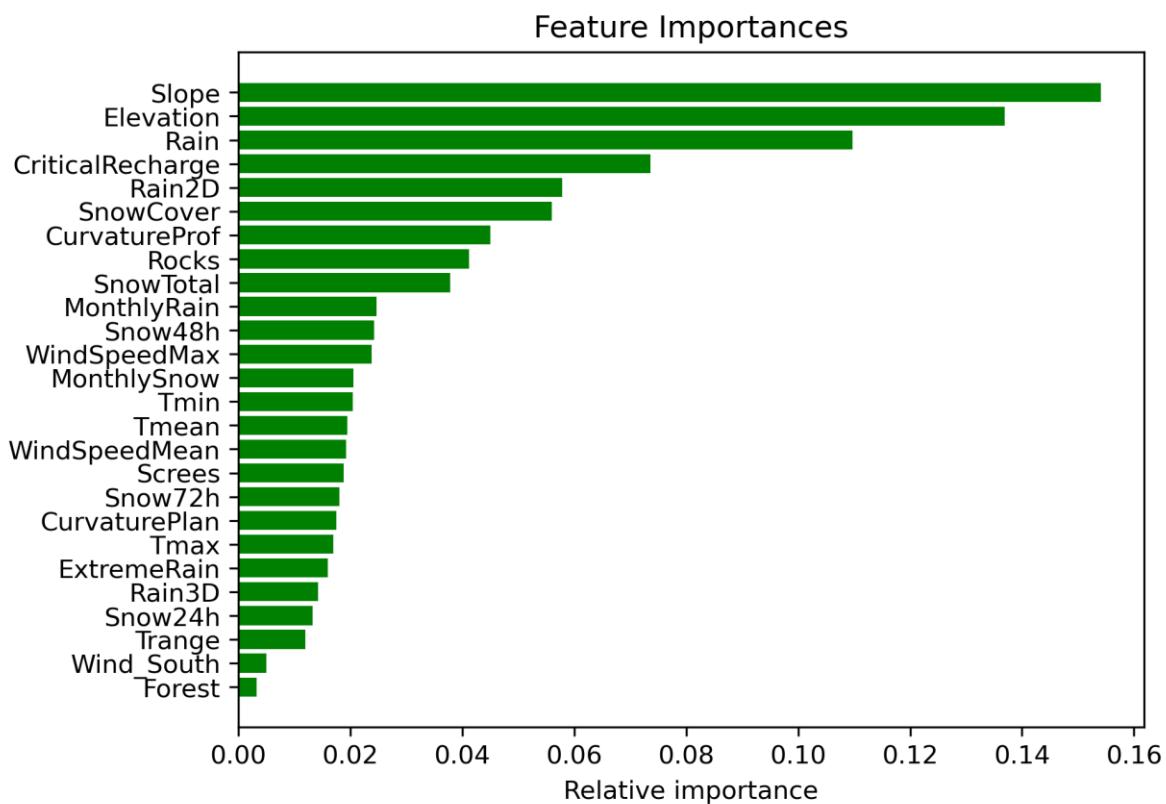


Figure 50. Bar plot depicting the importance of features for the Random Forest classifier.

### AdaBoost

Again, “RandomSearchCV” was applied for optimization, and subsequently, the model was trained as depicted in Figure 51. Based on the confusion matrix acquired when validating the model (Figure 52), the AdaBoost presented even higher accuracy, specifically 98%, by almost diminishing the missed events and false alarms.

```
ada = AdaBoostClassifier(learning_rate=0.5, n_estimators=425, random_state=42)
ada.fit(X_train, y_train)
```

Figure 51. Training of the optimized AdaBoost classifier.

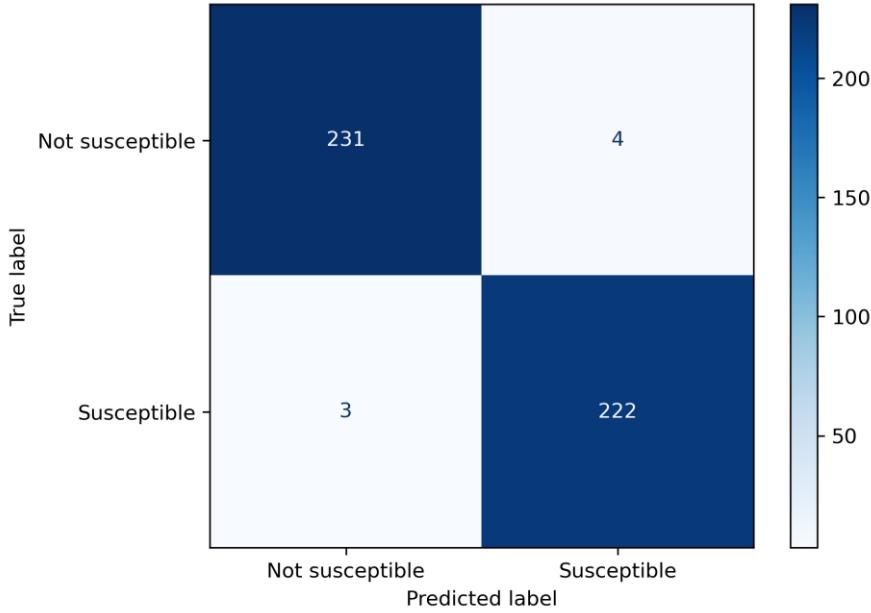


Figure 52. Confusion matrix of the optimized AdaBoost model on the validation set.

### Gradient Boosting

For the optimization of the Gradient Boost, the “GridSearchCV” method was used, performed in two rounds, in order to reduce the run time. In this way, the parameters acquired in the first round were assessed and for those that it was considered necessary to further explore the parameter space, the second round of optimization was performed. Figure 53 provides the search of the first round as an example.

```
# ROUND 1: Optimization of hyperparameters with cross validation
param_grid = {
    'max_depth': [1, 3, 5],
    'learning_rate': [0.1, 0.3, 0.5],
    'subsample': [0.5, 0.75, 1],
    'random_state': [1],
    'n_estimators': [100, 500]
}

optimal_params = GridSearchCV(estimator=GradientBoostingClassifier(), param_grid=param_grid, scoring='roc_auc',
                               verbose=0, n_jobs=-1, cv=10)

optimal_params.fit(X_train, y_train)

print(optimal_params.best_params_)
```

Figure 53. Optimization of hyperparameters of the Gradient Boost algorithm using "GridSearchCV".

The final fit and the confusion matrix on the validation data are shown in Figure 54 and Figure 55, respectively. Similar to the AdaBoost, the Gradient Boost achieved very high accuracy of 98%. An interesting observation is that, although “GridSearchCV” proposed the utilization of 1000

estimators in order to get the best fit, Figure 56 demonstrates that, in reality, the model reached excellent performance by building less than 100 trees. By plotting the feature importances (Figure 57), a ranking similar to the one of the Random Forest was acquired.

```
GBC = GradientBoostingClassifier(max_depth=5, learning_rate=0.1, subsample=0.75, random_state=1, n_estimators=1000)
GBC.fit(X_train, y_train)
```

Figure 54. Training of the optimized Gradient Boost model.

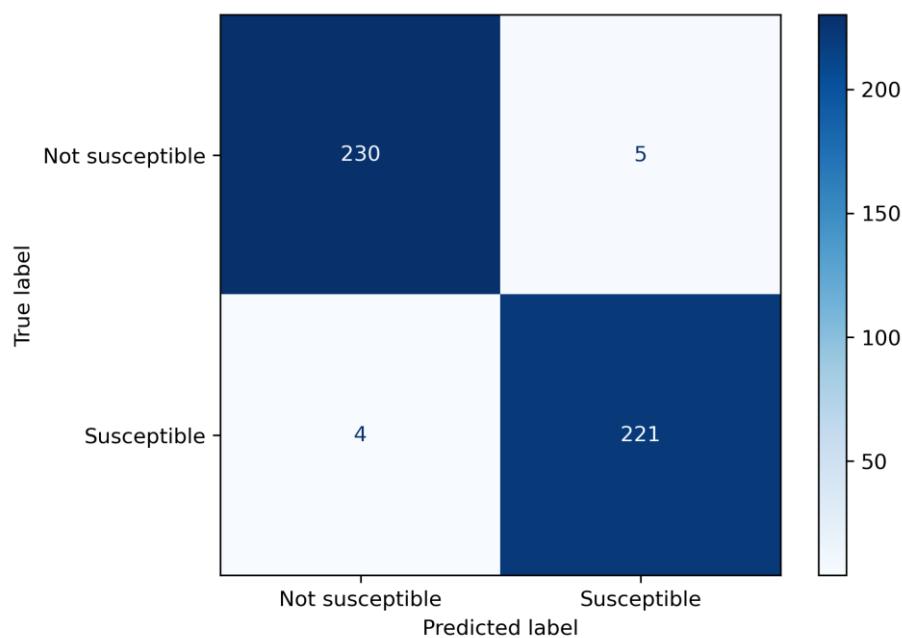


Figure 55. Confusion matrix of the optimized Gradient Boost on the validation set.

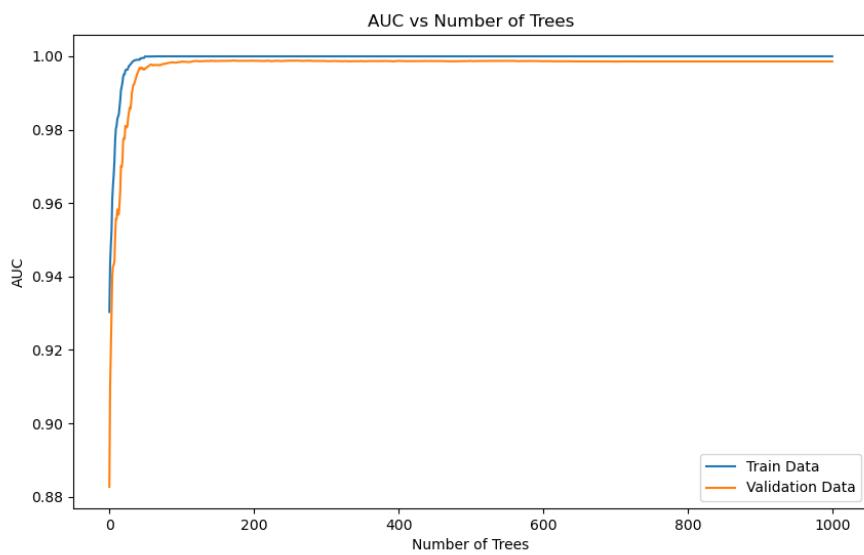


Figure 56. Graph of increasing accuracy with increasing number of estimators for the Gradient Boost model.

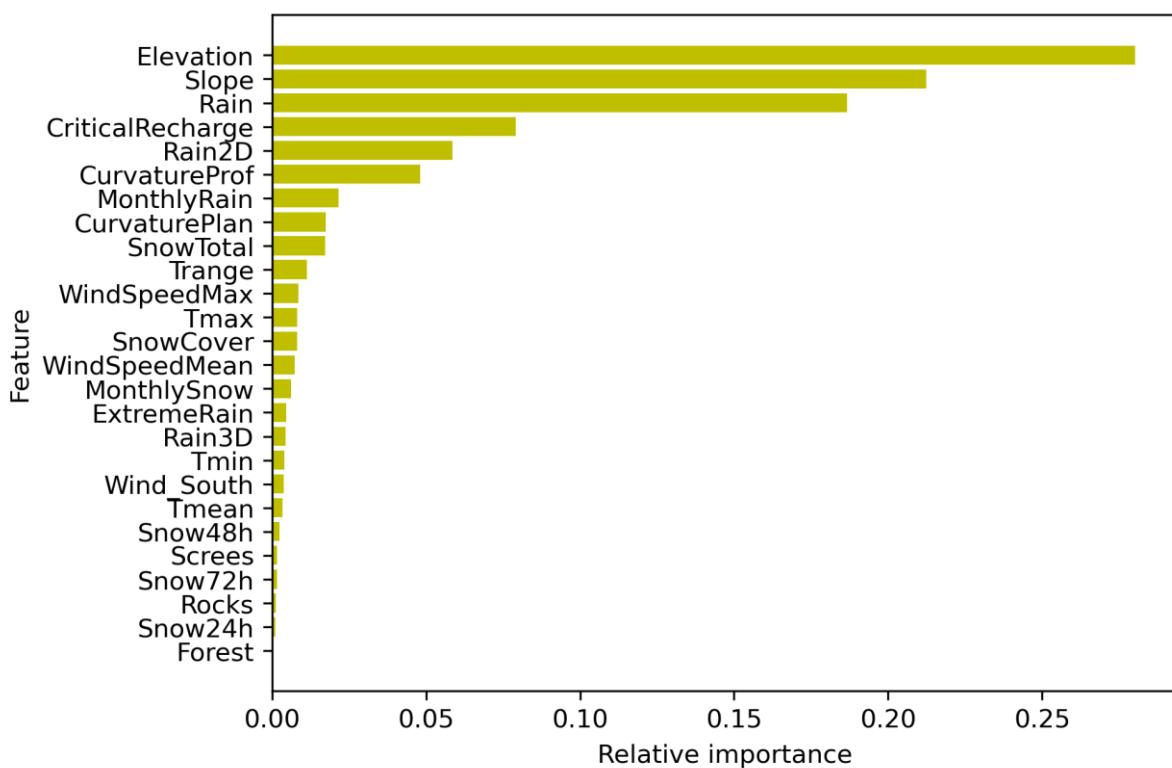


Figure 57. Feature importance graph of the Gradient Boost model.

## XGBoost

For the optimization of the XGBoost model the “GridSearchCV” method was applied again. A difference between this algorithm and the previous boosting models, is that, here, in order to avoid overfitting, an evaluation set is provided during the fit, which enables early stopping. Hence, the number of estimators is not a hyperparameter in this case. Figure 58 presents the training of the model using the optimized parameters. Once again, the accuracy of the XGBoost is very high, namely 98%, with very few misses and false positive events reported in the confusion matrix (Figure 59).

```
# Now that we have optimised the hyperparameters, we can build the final XGBoost model
clf_xgb = xgb.XGBClassifier(objective='binary:logistic',
                             seed=42,
                             early_stopping_rounds=10,
                             eval_metric='aucpr',
                             gamma=0,
                             learning_rate=0.3,
                             max_depth=3,
                             reg_lambda=1,
                             scale_pos_weight=1,
                             #subsample=0.9,
                             #colsample_bytree=0.5
)
clf_xgb.fit(X_train,
             y_train,
             verbose=True,
             eval_set=[(X_valid, y_valid)])
```

Figure 58. Training of the XGBoost model with the optimized hyperparameters.

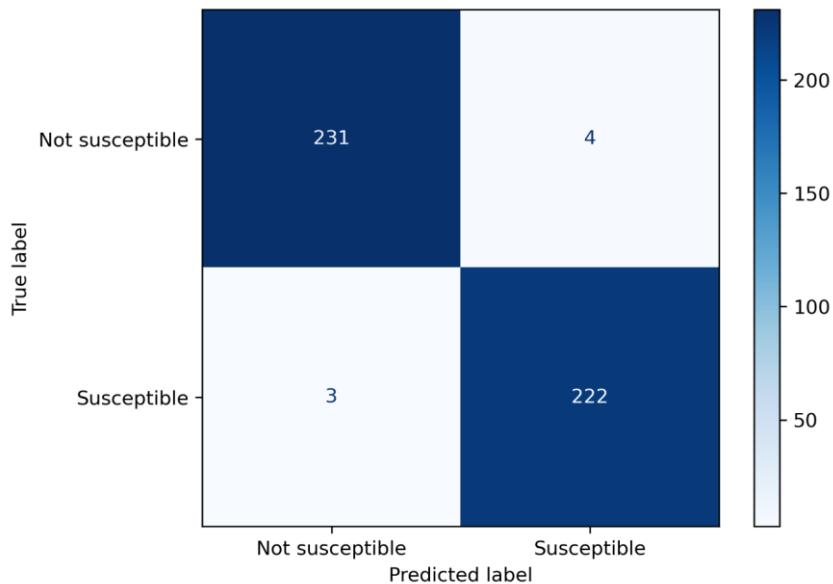


Figure 59. Confusion matrix of the optimized XGBoost on the validation set.

## Logistic Regression

Logistic Regression suffers from multicollinearity and, for this reason, the variables of accumulated rain and snow of 48 and 72 hours, the minimum and maximum temperature and the daily temperature range were removed from the database before the model development process. Additionally, standardization is fundamental for Logistic Regression, in order to make sure a feature's weight does not dominate over others. The model was trained with the optimized parameters as shown in Figure 60, and the accuracy is presented in the confusion matrix (Figure 61). More specifically, the total validation score was 84%, with the recall being 84% as well, and the precision 83%.

```
reg = LogisticRegression(solver='lbfgs', max_iter = 100, C=1, random_state=5)
reg.fit(X_train_mc, y_train)
```

Figure 60. Training of the Logistic Regression model using the optimized hyperparameters.

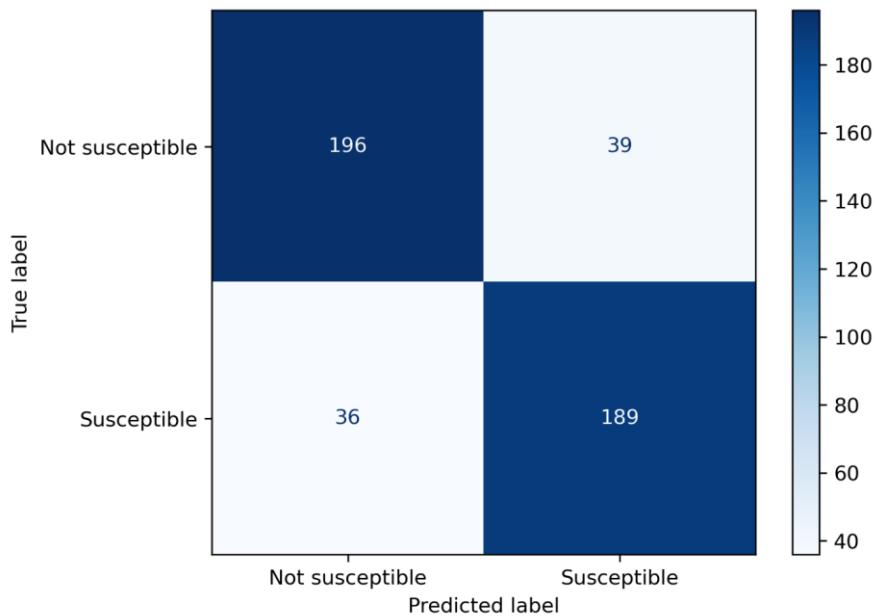


Figure 61. Confusion matrix of the optimized Logistic Regression on the validation set.

## Support Vector Machine

Being also influenced by multicollinearity effects, the SVM was optimized and trained (Figure 62) without the features that are highly correlated with others, using the “SVC” (Support Vector Machine for Classification) estimator of the sklearn package. Moreover, similar to Logistic

Regression, the Radial Basis Function (RBF), used as a kernel for the SVC in this case, assumes that the data are centered and scaled. Thus, standardization was performed prior to model development. Although SVMs are performing quite well out of the box, as confirmed by the confusion matrix before optimization (Figure 63), after the optimization the model predicts better regarding the false negatives and slightly worse as related to false positives (Figure 64).

```
clf_svm = SVC(random_state=42, C=100, gamma=0.01, kernel='rbf')
clf_svm.fit(x_train_mc, y_train)
```

Figure 62. Training of the SVC with the optimized hyperparameters.

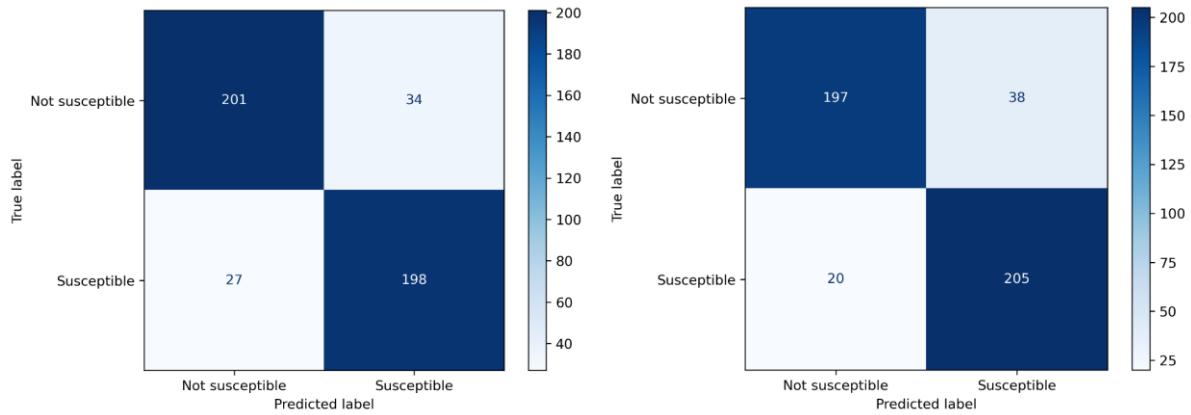


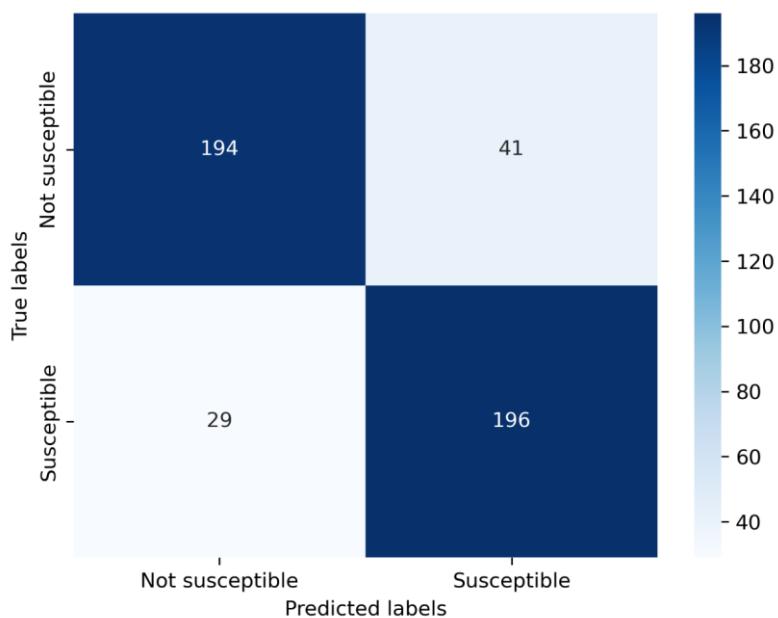
Figure 63. Confusion matrix of the SVC before optimization.

Figure 64. Confusion matrix of the SVC after optimization.

### Feedforward Neural Network

The Neural Network also requires the input data to be scaled but is not sensitive to features highly correlated with each other. The model was applied using the “Sequential” class of Keras, which groups a linear stack of layers into the model. The selected transfer function for the hidden layer was the “relu”, which applies the rectified linear unit activation function, while for the output layer the “sigmoid” function was selected. The number of nodes for the input layer was 64 and for the hidden layer 128, while the number of epochs was set to 256. The model was compiled using the “binary\_crossentropy” as a loss function, which is the standard for binary classification, and the “Adam” algorithm (Kingma & Ba, 2015), which is a stochastic gradient descent method that is based on adaptive estimation of first-order and second-order moments.

In general, too many epochs might lead to overfitting, while too few might result in an underfit model. Even if the epochs were fixed to a certain number, early stopping was applied in order to stop training once the model performance would stop improving on the validation dataset. In the case of the NN development in Keras there are no hyperparameters, although different optimization algorithms or different combinations of nodes and epochs can be tested. Since the sigmoid function provides an output between 0 and 1, a threshold of 0.5 was set to classify the classes into 1 or 0, signifying, as per usual in this study, susceptibility or no susceptibility, respectively. The acquired confusion matrix is shown in Figure 65. The achieved total accuracy was 85%, the recall 87% and the precision 83%.



*Figure 65. Confusion matrix of the feedforward neural network.*

Finally, all the developed models were deployed using the combined dataset of training and validation (consisting of  $1,380 + 460 = 1,840$  samples) and the respective optimized hyperparameters in each case. The evaluation of their performance using the testing dataset is presented in the next chapter.

## 5. RESULTS AND DISCUSSION

---

### 5.1. Testing and evaluation of the models

The accuracy of the created models was evaluated using “unseen” data (i.e., the test set) and is summarized in Table 6. As previously mentioned, there is variance in the error estimation of the “train-validation-test split” approach. In an effort to minimize the effect of this variance on the final, reported error, the accuracy presented in the summary table is, in fact, the average, derived from a 10-fold cross validation. Also, in the same table, a comparison of the test error can be made for both versions of the models that were built. The first version used the simpler database (including only topographical features) and the second one the entire database.

Table 6. Test error of both versions of models. The first version was developed using only topographical features and the second using the entire database.

	Predictors: Only topography	Predictors: Topography & Weather & Remote sensing & Shalstab
Model	Accuracy	
Decision Tree	80 %	82 %
Random Forest	82 %	86 %
AdaBoost	81 %	93 %
Gradient Boosting	82 %	94 %
XGBoost	83 %	95 %
Logistic Regression	81 %	80 %
SVM	82 %	80 %
Neural Network	82 %	86 %

## 5.2. Discussion

A first, prominent observation from Table 6 is that the additional predictors of weather, snow cover and the output of the shallow landslide model enhanced the performance of almost all the models. The boosting algorithms, in particular, presented a major improvement in their accuracy. Although it cannot be seen in the table, during model development of both versions of models it was observed that the inclusion of more features meant that the models were optimized for building more estimators, and the tree-based algorithms specifically, created deeper trees. This means that they utilized the extra inputs in every possible way to improve their accuracy. In the case of the Neural Network, the model ran more times and was able to recognize new patterns in the data when additional features were fed to it.

Another outcome of this comparison is that the bagging algorithm (i.e., the Random Forest) in both versions of models is, as expected, more robust than the single decision tree but performs worse than the boosting algorithms. This is due to the nature and ultimate goal of the ensemble methods, which is to combine the predictions of several base estimators in order to improve generalizability over a single estimator. However, the bagging methods solely average the predictions of independently built estimators, whereas, in the boosting methods, the estimators are built sequentially and the quality of the model is improved with every tree. This characteristic of the boosting algorithms reveals their superiority in terms of performance. On the other hand, their strategy of building several weak models in order to finally produce a powerful ensemble, indicates that one cannot follow along the model build, interpret the results and explain the high accuracy achieved. By contrast, interpretability is the strong element of decision trees. Although they present higher variance compared to ensemble methods, they are simple to understand, and they can be visualized, which translates to easily explorable results and more informed optimization.

As opposed to the tree-based algorithms, the Logistic Regression and the SVM performed worse when more predictors were used during model development. An explanation could be that Logistic Regression performs well when there is high linearity in the data, and, in addition, it is highly vulnerable to multicollinearity. When more data were fed to the Logistic Regressor, although the effect of multicollinearity was reduced by excluding the most correlated features, it is possible that the consideration of more predictors with high overlap and no clear linear patterns, caused the model to produce less accurate predictions. A similar assumption could be made about the reduced performance of the SVM. SVMs are aiming to make predictions by recognizing data that could be

grouped together, which apparently was easier when only few features were provided as input to the model.

### 5.3. Susceptibility mapping

After the boosting and bagging algorithms, that are not quite suitable for mapping, and the Neural Network that is a “black-box”, the DT is the next best model in terms of performance, with 82% accuracy. Although the Random Forest is, in fact, a “black-box” model, when implemented with the sklearn package there is the possibility of visualizing any of the built trees. However, this would not be any different than building and plotting a single DT. As mentioned in the previous section, the power of interpretability that the DT demonstrates is very suited for getting insights regarding the phenomenon itself and the features that contribute the most to it. In the current project, considering the above-mentioned advantages, the DT was the selected model to be used for producing a map of avalanche susceptibility. During optimization, it was deliberately chosen to proceed with the cost-complexity pruning, because it allowed explicit handling/calibration of the algorithm, so that it best describes the investigated phenomenon and enables the creation of a susceptibility map that serves the needs of the study but also of the final users.

As described in section 3.4, the optimized DT was fitted one last time to the whole dataset. The produced tree is shown in Figure 66 and the decision rules of the tree are depicted in Figure 67. The weights denote how many events were grouped as either class in each leaf, and, in effect, they provide information of the probability/accuracy of the predicted class in the leaf. The sequence of decisions until the end node of class 1 are highlighted with yellow color. In this case, following this sequence, the areas that are located higher than 2039.5 m, have an inclination bigger than  $37.5^\circ$ , and have a critical recharge (as acquired from Shalstab) lower than 2693.75, are going to be susceptible to avalanche release with a probability of 88%.

In the final susceptibility map, each potential release area is a grid cell of 70x70 m (exaggerated for visualization reasons). In other words, these areas represent locations that are highly possible (88% probability) to initiate an avalanche. The main road network of Andorra is also depicted in the map, in an effort to distinguish road axes that are located near potential release areas. As can be observed, areas in the North-West and North of the country are more vulnerable to avalanches, because there are primary and secondary road sections that are very close to initiation areas. These road axes are located at high altitude and serve the ski resorts of Arinsal and Arcalis, providing access to the country from both Spain and France. By contrast, in the East and South-East,

although there are many initiation points, there is less urban development, less activity related to tourism and, thus, less extended road network, which decreases the potential damages from snow avalanches.

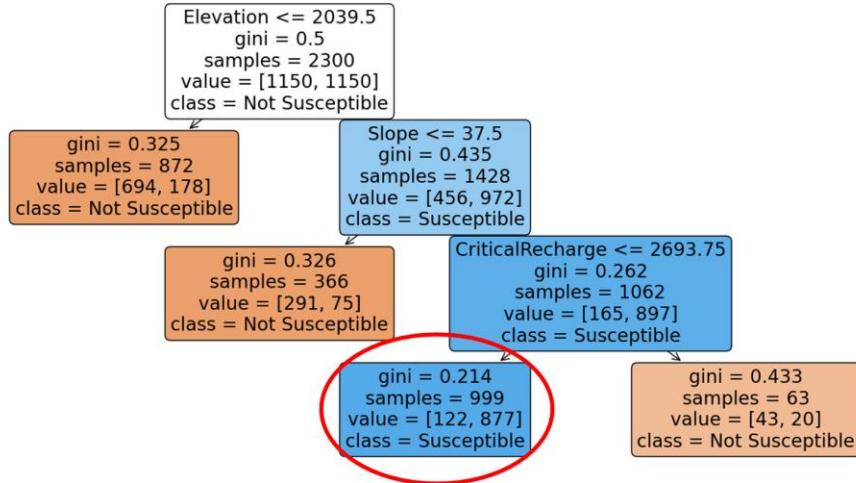


Figure 66. Plot of the optimized DT when fitted to the entire database.

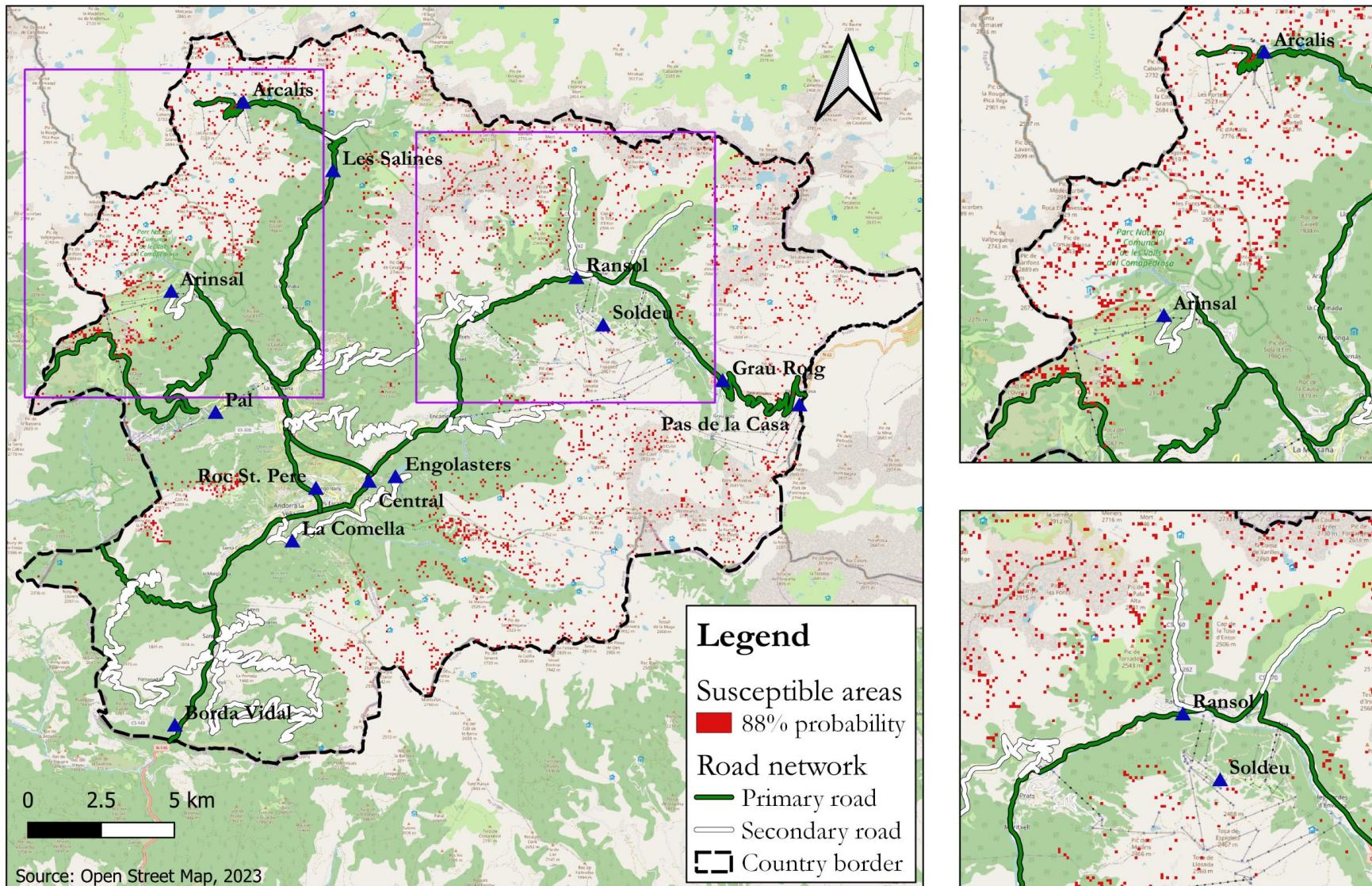
```

--- Elevation <= 2039.50
|--- weights: [694.00, 178.00] class: 0
--- Elevation > 2039.50
|--- slope <= 37.50
|--- weights: [291.00, 75.00] class: 0
|--- slope > 37.50
|--- CriticalRecharge <= 2693.75
|--- weights: [122.00, 877.00] class: 1
|--- CriticalRecharge > 2693.75
|--- weights: [43.00, 20.00] class: 0

```

Figure 67. Decision rules of the optimized DT fitted to the entire database.

# Avalanche susceptibility map



## 6. CONCLUSIONS AND RECOMMENDATIONS

---

### 6.1. Conclusions

The main aim of this thesis has been to investigate the suitability of data-driven models for avalanche susceptibility assessment, taking advantage of the inventory that the government of Andorra has been keeping since the 1980's. To achieve this aim, 1,122 historical avalanches were used and a set of objectives were set, which were gradually met throughout the completion of the thesis. In line with objective a, eight different machine learning models were brought to the test, namely the Decision Tree, Random Forest, AdaBoost, Gradient Boost, XGBoost, Logistic Regression, SVM, and Neural Network. The selection of models was such that a variety of algorithms could be tested, with different strengths and weaknesses each. Tree-based algorithms, bagging and boosting methods, linear models and “black-box” models were developed and compared. In relation to objective b, the obtained results highlight that all of the boosting algorithms, AdaBoost, Gradient Boost, and XGBoost, showcase high performances, that is 93%, 94%, and 95%, respectively. The next best models are the Random Forest and the Neural Network with 86% accuracy. In general, all of the tested models achieved accuracies of 80% or higher, which is considered satisfactory for the scope of this study.

The importance of the topographical characteristics for the classification is a profound outcome of the study. Elevation and slope systematically ranked the highest, while the contribution of curvature was also significant. This result reinforces the existing knowledge in the bibliography, according to which topographical features do not change as rapidly as weather or snow cover conditions, and, thus, account for reliable predictors with regard to avalanche susceptibility. In an attempt to explain the high performance of the boosting algorithms, the fact that a high-resolution DEM was available, combined with the robustness of these algorithms and the long avalanche inventory of Andorra, seems to have played a significant role.

Next to topography, daily rainfall but also accumulated rainfall of two days are ranked as two of the most important features. This result reveals the importance of precipitation for avalanche forecasting in the study area. It also highlights how mathematical and statistical techniques that enable the increase of dimensionality of the data during the preprocessing (in this case the calculation of the two-day rainfall accumulation) could enhance the performance of data-driven models. A relatively unexpected outcome in terms of feature importance is the low significance of

accumulated snow for the models, whereas, in practice, risk managers highly depend on 48 hours- or 72-hours snow accumulation measurements for assessing potential snow instability in an area. This result can be related to the few station measurements that were available, the quality of the data and also the interpolation method that was used, and sheds light to one of the biggest weaknesses of the study. Indeed, when developing machine learning forecasting models, one of the biggest challenges is to find reliable data.

On the other hand, one of the most interesting, acquired results is the contribution of the Shalstab's output to the performance of the majority of the models. Such a result confirms the relevance of the phenomena of shallow landslides and snow avalanches, as well as showcases the ability of machine learning models to recognize the similar patterns between them. Snow cover is also revealed as an important predictor, indicating the usefulness of remote sensing products even if ground-truth data are not available. When the two versions of models were compared, the ones that included the weather variables, the Shalstab output and the snow cover from Landsat clearly outperformed the ones built with less input data, with only a couple of exceptions that were previously discussed. The experimentation with the predictors and the exploration of the feature space enabled the objectives c and d to be successfully met, opened up possibilities for further research and underlined the innovation introduced by the current study.

Regarding the objective e, the Decision Tree, although not the best performing model (82% accuracy), proved to be very useful in terms of understanding the classification process and interpreting the results. Thus, it was the selected model based on which the susceptibility map was produced. The map is a simple visualization tool for narrowing down hot-spot areas of potential avalanche release. It is logical to assume that the local authorities responsible for avalanche control are well aware of the locations that are more vulnerable to the hazard. The approach developed in this thesis does not intend to substitute the conventional forecasting that depends on field measurements, snow profile/stability tests, constant monitoring, and weather analysis. On the contrary, the map can be used complementary to these methods and provide a good starting point for acquiring, at a glance, a quick idea of avalanche susceptibility. It can also assist the road maintenance experts, enable a better-informed strategy for land-use planning, and be used for suggesting the installation of protection measures in critical locations, based on evidence. As for any of the highly performing boosting algorithms, they could enhance the undergoing modelling effort in the country, providing predictions of initiation areas to the flow/dynamic models.

Overall, the findings of the study were able to answer both of the research questions, and, more specifically, they provided important insights into the contributing factors related to avalanche initiation in Andorra and proved that data-driven models are a suitable alternative for snow avalanche susceptibility assessment and mapping. Among the tested models, some produced promising results in capturing the patterns behind the release of the phenomenon, while others proved to be very efficient, highlighting the need for further research in the field but also in the study area.

## 6.2. Recommendations

In light of the findings of the current study and the conclusions outlined above, several key areas of future research are identified. The recommendations presented in this final section of the thesis aim to enhance the robustness and reliability of data-driven models with regard to snow avalanche susceptibility.

Firstly, future research should prioritize the inclusion of a more extensive network of weather stations throughout the study area. It has already been discussed that the wider area of the Pyrenees suffers from data scarcity and lack of metadata, which makes the quality control even harder. In Andorra, the last ten to twenty years many new meteorological stations have been established. However, in most cases their measurements have not been checked for outliers or other inaccuracies, and it is an undergoing effort and challenge for the local weather service to accomplish higher quality of the data in the following years. A denser network of weather stations will facilitate the collection of comprehensive meteorological data, allowing for a more detailed and accurate representation of the prevailing weather conditions. This, in turn, will enhance the precision of the models and enable better-informed decision-making by local authorities and avalanche forecasters. Additionally, the inclusion of real-time data from weather stations or regional weather forecasts can enhance the timeliness and responsiveness of avalanche warnings, contributing to the preparedness during critical weather events.

Secondly, it is essential to explore and implement advanced spatial interpolation methods for the weather inputs. Improved interpolation techniques can help fill data gaps in regions with limited weather station coverage, leading to a more complete picture of the meteorological conditions across the study area. By incorporating advanced interpolation methodologies, such as kriging or machine learning-based techniques, higher accuracy of the models could be achieved, reducing uncertainties associated with data scarcity.

Furthermore, a significant advancement in avalanche susceptibility assessment would involve incorporating historical avalanche event dates to acquire accurate weather conditions prevailing during the days of these events. By utilizing such information instead of timeseries averages, more accurate correlations between specific weather patterns and avalanche occurrences can be established, enabling more informed predictions of future avalanche events. In Andorra such data are not available in the existing avalanche inventory, although the last couple of years the occurred events have been recorded in a more detailed way, by taking aerophotographs, documenting the day and time, the weather conditions, and other information derived from field tests. If the methodology was to be implemented in another region where an avalanche inventory existed and the weather conditions of the events were known, the reliability of the avalanche susceptibility models and their potential for early warning systems could be enhanced significantly.

Moreover, an area that warrants particular attention, and has not been included at all in the current study, is the exploration of snowpack conditions and their influence on snow instability that could lead to an avalanche. In-depth investigation of snowpack properties, such as layering, density, and temperature gradients, can provide critical insights into the potential for avalanche initiation. Such research will aid in identifying specific snowpack characteristics associated with high avalanche susceptibility, ultimately leading to better-tailored risk mitigation strategies and avalanche forecasting methods.

Finally, the high importance of Shalstab as predictor to the models opens up new prospects of research and ideas for investigating the phenomenon of snow avalanches using existing, physically-based models developed for assessing shallow landslides. The two phenomena share a lot of similarities and the current study creates new ground on using and adjusting shallow landslide models for predicting avalanche initiation.

In conclusion, the future research directions outlined above hold substantial promise for enhancing avalanche susceptibility assessment in Andorra. By expanding the weather station network, exploring advanced spatial interpolation methods, investigating snowpack conditions, and complementing the existing inventory with new events and the weather conditions during them, the accuracy and reliability of avalanche susceptibility models, and, more specifically, data-driven models, can significantly improve. At the same time, physically-based models borrowed from shallow landslides can be a promising field of future research.

## REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., ... Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*. <https://doi.org/10.5281/zenodo.4724125>
- Acero, F. J., García, J. A., & Gallego, M. C. (2011). Peaks-over-threshold study of trends in extreme rainfall over the Iberian Peninsula. *Journal of Climate*, 24(4), 1089–1105. <https://doi.org/10.1175/2010JCLI3627.1>
- Akay, H. (2021). Spatial modeling of snow avalanche susceptibility using hybrid and ensemble machine learning techniques. *Catena*, 206(April), 105524. <https://doi.org/10.1016/j.catena.2021.105524>
- Apodaka, J., Pons, M., Trapero, L., Margalef, A., Albalat, A., Gallego, N., López-Moreno, J. I., & Furdada, G. (2018). Evaluation of 30 Years of Nivo-Meteorological and Avalanche Data in Andorra. *International Snow Science Workshop, Innsbruck, Austria, October 2018*. www.cenma.ad
- Arinsal: Sun direction.* (n.d.). Retrieved July 8, 2023, from <https://sun-direction.com/city/994/arinsal/>
- Baggi, S., & Schweizer, J. (2009). Characteristics of wet-snow avalanche activity: 20 years of observations from a high alpine valley (Dischma, Switzerland). *Natural Hazards*, 50(1), 97–108. <https://doi.org/10.1007/s11069-008-9322-7>
- Bakermans, L., & Jamieson, B. (2009). SWarm: A simple regression model to estimate near-surface snowpack warming for back-country avalanche forecasting. *Cold Regions Science and Technology*, 59(2–3), 133–142. <https://doi.org/10.1016/j.coldregions.2009.06.003>
- Barbolini, M., Pagliardi, M., Ferro, F., & Corradeghini, P. (2011). Avalanche hazard mapping over large undocumented areas. *Natural Hazards*, 56(2), 451–464. <https://doi.org/10.1007/s11069-009-9434-8>

- Bartelt, P., Salm, B., & Gruber, U. (1999). Calculating dense-snow avalanche runout using a Voellmy-fluid model with active/passive longitudinal straining. *Journal of Glaciology*, 45(150), 242–254. <https://doi.org/10.3189/s002214300000174x>
- Bellaire, S., van Herwijnen, A., Mitterer, C., & Schweizer, J. (2017). On forecasting wet-snow avalanche activity using simulated snow cover data. *Cold Regions Science and Technology*, 144(July), 28–38. <https://doi.org/10.1016/j.coldregions.2017.09.013>
- Bergua, S. B., Piedrabuena, M. Á. P., & Alfonso, J. L. M. (2018). Snow avalanche susceptibility in the eastern hillside of the aramo range (Asturian central massif, cantabrian mountains, nw spain). *Journal of Maps*, 14(2), 373–381. <https://doi.org/10.1080/17445647.2018.1480974>
- Blikra, L. H., & Nemec, W. (1998). Postglacial colluvium in western Norway: depositional processes, facies and palaeoclimatic record. *Sedimentology*, 45(5), 909–959. <https://doi.org/10.1046/j.1365-3091.1998.00200.x>
- Boyd, J., Haegeli, P., Abu-Laban, R. B., Shuster, M., & Butt, J. C. (2009). Patterns of death among avalanche fatalities: a 21-year review. *Canadian Medical Association Journal*, 180(5), 507–512. <https://doi.org/10.1503/cmaj.081327>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24, 123–140. <https://doi.org/https://doi.org/10.1007/BF00058655>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees (Wadsworth, Belmont, CA). ISBN-13, 978–412048418.
- Bühler, Y., Kumar, S., Veitinger, J., Christen, M., Stoffel, A., & Snehmani. (2013). Automated identification of potential snow avalanche release areas based on digital elevation models. *Natural Hazards and Earth System Science*, 13(5), 1321–1335. <https://doi.org/10.5194/nhess-13-1321-2013>
- Bühler, Y., Von Rickenbach, D., Stoffel, A., Margreth, S., Stoffel, L., & Christen, M. (2018). Automated snow avalanche release area delineation-validation of existing algorithms and proposition of a new object-based approach for large-scale hazard indication mapping. *Natural Hazards and Earth System Sciences*, 18(12), 3235–3251. <https://doi.org/10.5194/nhess-18-3235-2018>

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A Library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–40. <https://doi.org/10.1145/1961189.1961199>

Chen, J., Fyffe, B., Han, D., & Yang, S. (2022). Predicting mixed mode damage propagation in snowpack using the extended cohesive damage element method. *Theoretical and Applied Fracture Mechanics*, 122(September), 103567. <https://doi.org/10.1016/j.tafmec.2022.103567>

Chen, T., & Guestrin, C. (2016). {XGBoost}: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>

Chollet, F., & others. (2015). *Keras*.

Choubin, B., Borji, M., Hosseini, F. S., Mosavi, A., & Dineva, A. A. (2020). Mass wasting susceptibility assessment of snow avalanches using machine learning models. *Scientific Reports*, 10(1), 1–13. <https://doi.org/10.1038/s41598-020-75476-w>

Choubin, B., Borji, M., Mosavi, A., Sajedi-Hosseini, F., Singh, V. P., & Shamshirband, S. (2019). Snow avalanche hazard prediction using machine learning methods. *Journal of Hydrology*, 577(May), 123929. <https://doi.org/10.1016/j.jhydrol.2019.123929>

Christen, M., Bartelt, P., & Gruber, U. (2002). AVAL-1D: an avalanche dynamics program for the practice. *International Congress Interpraevent 2002 in the Pacific Rim, Matsumoto / Japan*, 2, 715–725.

Christen, M., Bartelt, P., Kowalski, J., & Stoffel, L. (2007). Calculation of dense snow avalanches in three-dimensional terrain with the numerical simulation program RAMMS. *Differential Equations, Proceeding*, 709–716.

Christen, M., Bühler, Y., Bartelt, P., Leine, R., Glover, J., Schweizer, A., Graf, C., Mcardell, B. W., Gerber, W., Deubelbeiss, Y., & Feistl, T. (2012). Integral Hazard Management Using a Unified Software Environment Numerical Simulation Tool “RAMMS.” *Congress Interpraevent*, 77–86.

Christen, M., Kowalski, J., & Bartelt, P. (2010). RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. *Cold Regions Science and Technology*, 63(1–2), 1–14. <https://doi.org/10.1016/j.coldregions.2010.04.005>

- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, 8(7), 1991–2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- Conway, H., Carran, W., Techel, F., & Carran, A. (2008). Guidelines for forecasting snow avalanches on the Milford Road, New Zealand. *Proceedings of the International Snow Science Workshop. Whistler, Canada.*
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297.
- Curvature function—ArcMap | Documentation.* (n.d.). Retrieved July 7, 2023, from <https://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/curvature-function.htm>
- Davis, R. E., 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 Regions Science and Technology*, 30, 79–89.
- De Haas, T., McArdell, B. W., Conway, S. J., McElwaine, J. N., Kleinhans, M. G., Salese, F., & Grindrod, P. M. (2019). Initiation and Flow Conditions of Contemporary Flows in Martian Gullies. *Journal of Geophysical Research: Planets*, 124(8), 2246–2271. <https://doi.org/10.1029/2018JE005899>
- Desktop Help 10.0 - How Curvature works.* (n.d.). Retrieved July 8, 2023, from <https://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/00q90000000t000000>
- Development and importance of tourism for Andorra.* (n.d.). Retrieved February 28, 2023, from <https://www.worlddata.info/europe/andorra/tourism.php>
- Dietrich, W. E., & Montgomery, D. R. (1998). *SHALSTAB: A Digital Terrain Model for Mapping Shallow Landslide Potential.* University of California. <https://books.google.es/books?id=ahz7uQEACAAJ>
- Dillon, J., & Hammonds, K. (2021). Brief Communication: Initializing RAMMS with High Resolution LiDAR Data for Avalanche Simulations. *The Cryosphere Discussions*, January, 1–11.

Dong, P., Yang, C., Rui, X., Zhang, L., & Cheng, Q. (2003). An Effective Buffer Generation Method in GIS. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 6(100101), 3706–3708. <https://doi.org/10.1109/igarss.2003.1295244>

*Double-mass curve - Glossary of Meteorology.* (n.d.). Retrieved July 7, 2023, from [https://glossary.ametsoc.org/wiki/Double-mass\\_curve](https://glossary.ametsoc.org/wiki/Double-mass_curve)

Durlević, U., Valjarević, A., Novković, I., Ćurčić, N. B., Smiljić, M., Morar, C., Stoica, A., Barišić, D., & Lukić, T. (2022). GIS-Based Spatial Modeling of Snow Avalanches Using Analytic Hierarchy Process. A Case Study of the Šar Mountains, Serbia. *Atmosphere*, 13(8). <https://doi.org/10.3390/atmos13081229>

Esteban, P., Jones, P. D., Martín-Vide, J., & Mases, M. (2002). Application of multivariate statistic techniques for the characterization of the surface circulation patterns related to heavy snowfalls in Andorra, Pyrenees. *Proceedings ISSW 2002. International Snow Science Workshop*, 401–406.

Esteban, P., Jones, P. D., Martín-Vide, J., & Mases, M. (2005). Atmospheric circulation patterns related to heavy snowfall days in Andorra, Pyrenees. *International Journal of Climatology*, 25(3), 319–329. <https://doi.org/10.1002/joc.1103>

*Fatalities – EAWS.* (n.d.). Retrieved February 22, 2023, from <https://www.avalanches.org/fatalities/>

Fohn, P. M. B. (1980). Snow transport over mountain crests. *Journal of Glaciology*, 26(94), 469–480. <https://doi.org/10.1017/S0022143000010984>

*Forecasts – BRASS Avalanche.* (n.d.). Retrieved February 23, 2023, from <https://brassavalanche.org/forecasts/>

Francisco, G., Apodaka, J., Travesset-Baro, O., Vilella, M., Margalef, A., & Pons, M. (2018). Exploring the potential of mobile phone data (call detail records) to track and analyze backcountry skiers dynamics in avalanche terrain. *International Snow Science Workshop, Innsbruck, Austria*, 3–6.

Freedman, D., Pisani, R., & Purves, R. (2007). Statistics (international student edition). *Pisani, R. Purves, 4th Edn. WW Norton & Company, New York.*

- Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1), 119–139. <https://doi.org/10.1006/jcss.1997.1504>
- Friedman, J. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://www.jstor.org/stable/2699986>
- Furdada, G., Margalef, A., Trapero, L., Pons, M., Areny, F., Baró, M., Reyes, A., & Guinau, M. (2020a). The Avalanche of Les Fonts d'Arinsal (Andorra): An Example of a Pure Powder, Dry Snow Avalanche. *Geosciences*, 10(4). <https://doi.org/doi:10.3390/geosciences10040126>
- Furdada, G., Margalef, A., Trapero, L., Pons, M., Areny, F., Baró, M., Reyes, A., & Guinau, M. (2020b). The avalanche of les fonts d'arinsal (Andorra): An example of a pure powder, dry snow avalanche. *Geosciences (Switzerland)*, 10(4), 1–26. <https://doi.org/10.3390/geosciences10040126>
- Gaume, J., Chambon, G., Eckert, N., & Naaim, M. (2013). Influence of weak-layer heterogeneity on snow slab avalanche release: Application to the evaluation of avalanche release depths. *Journal of Glaciology*, 59(215), 423–437. <https://doi.org/10.3189/2013JoG12J161>
- Gaume, J., Chambon, G., Eckert, N., Naaim, M., & Schweizer, J. (2015). Influence of weak layer heterogeneity and slab properties on slab tensile failure propensity and avalanche release area. *Cryosphere*, 9(2), 795–804. <https://doi.org/10.5194/tc-9-795-2015>
- Gaume, J., Herwijnen, A. Van, Gast, T., Teran, J., & Jiang, C. (2019). Investigating the release and flow of snow avalanches at the slope-scale using a unified model based on the material point method. *Cold Regions Science and Technology*, 168(July), 102847. <https://doi.org/10.1016/j.coldregions.2019.102847>
- Gaume, J., & Reuter, B. (2017). Assessing snow instability in skier-triggered snow slab avalanches by combining failure initiation and crack propagation. *Cold Regions Science and Technology*, 144(June), 6–15. <https://doi.org/10.1016/j.coldregions.2017.05.011>
- Gaume, J., Reuter, B., Herwijnen, A. Van, & Schweizer, J. (2016). A new crack propagation criterion for skier-triggered snow slab avalanches. *International Snow Science Workshop, Breckenridge, Colorado*, 484–491.

- Gaume, J., Schweizer, J., Herwijnen, A., Chambon, G., Reuter, B., Eckert, N., & Naaim, M. (2014). Evaluation of slope stability with respect to snowpack spatial variability. *Journal of Geophysical Research: Earth Surface*, 119(9), 1783–1799. <https://doi.org/10.1002/2014jf003193>
- Gaume, J., Van Herwijnen, A., Chambon, G., Birkeland, K. W., & Schweizer, J. (2015). Modeling of crack propagation in weak snowpack layers using the discrete element method. *Cryosphere*, 9(5), 1915–1932. <https://doi.org/10.5194/tc-9-1915-2015>
- Gaume, J., Van Herwijnen, A., Chambon, G., Wever, N., & Schweizer, J. (2017). Snow fracture in relation to slab avalanche release: Critical state for the onset of crack propagation. *Cryosphere*, 11(1), 217–228. <https://doi.org/10.5194/tc-11-217-2017>
- Gaume, J., van Herwijnen, A., Chambon, G., Wever, N., & Schweizer, J. (2017). Snow fracture in relation to slab avalanche release: Critical state for the onset of crack propagation. *Cryosphere*, 11(1), 217–228. <https://doi.org/10.5194/tc-11-217-2017>
- Gauthier, F., Germain, D., & Hétu, B. (2017). Logistic models as a forecasting tool for snow avalanches in a cold maritime climate: northern Gaspésie, Québec, Canada. *Natural Hazards*, 89(1), 201–232. <https://doi.org/10.1007/s11069-017-2959-3>
- Geitner, C., Mayr, A., Rutzinger, M., Löbmann, M. T., Tonin, R., Zerbe, S., Wellstein, C., Markart, G., & Kohl, B. (2021). Shallow erosion on grassland slopes in the European Alps – Geomorphological classification, spatio-temporal analysis, and understanding snow and vegetation impacts. *Geomorphology*, 373, 107446. <https://doi.org/10.1016/j.geomorph.2020.107446>
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3–42. <https://doi.org/10.1007/s10994-006-6226-1>
- Ghinoi, A., & Chung, C. J. (2005). STARTER: A statistical GIS-based model for the prediction of snow avalanche susceptibility using terrain features - Application to Alta Val Badia, Italian Dolomites. *Geomorphology*, 66, 305–325. <https://doi.org/10.1016/j.geomorph.2004.09.018>
- Gilaberte Búrdalo, M. (2018). *Effects of meteorological variability and climate change on ski tourism in the Spanish Central Pyrenees and Andorra: analysis of trends and future scenarios* [PhD Dissertation]. Universidad San Jorge Instituto.

Gleason, J. A., Birkeland, K. W., & Schweizer, J. (2008). Anticracks: A new theory Of fracture initiation and fracture propagation in snow. *International Snow Science Workshop*, 9–15. <https://doi.org/10.1016/j.coldregions.2007.09.004>

*Glossary* – *EAWS*. (n.d.-a). Retrieved February 22, 2023, from <https://www.avalanches.org/glossary/#avalanche-types>

*Glossary* – *EAWS*. (n.d.-b). Retrieved March 3, 2023, from <https://www.avalanches.org/glossary/#whumpfing-collapsing-sound>

Gobiet, A., Mitterer, C., Jöbstl, L., Steinkogler, W., Rieder, H., Olefs, M., Studeregger, A., Monti, F., & Bellaire, S. (2016). Operational forecasting of wet snow avalanche activity: a case study for the Eastern European Alps. *Proceedings ISSW*, 3–7.

González-Rouco, J. F., Jiménez, J. L., Quesada, V., & Valero, F. (2001). Quality control and homogeneity of precipitation data in the southwest of Europe. *Journal of Climate*, 14(5), 964–978. [https://doi.org/10.1175/1520-0442\(2001\)014<0964:QCAHOP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0964:QCAHOP>2.0.CO;2)

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. <https://doi.org/10.1016/j.rse.2017.06.031>

Granig, M., Sampl, P., Kofler, A., Fischer, J., & Jörg, P. (2016). Adaption and further development of the numerical solution in the avalanche simulation model SamosAT. *Interpraevent*, 284–289.

Guillet, L., Blatny, L., Trottet, B., & Gaume, J. (2023). *A Depth-Averaged Material Point Method for Shallow Landslides: Applications to Snow Slab Avalanche Release A Depth-Averaged Material Point Method for Shallow Landslides: Applications to Snow Slab Avalanche Release*. <https://doi.org/10.48550/arXiv.2301.06296>

Hall, D. K. and G. A. Riggs. (2021). MODIS/Terra Snow Cover Daily L3 Global 500m SIN Grid, Version 61 [Data Set]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. <https://doi.org/10.5067/MODIS/MOD10A1.061>. Date Accessed 07-17-2023.

Hastie, T. et. all. (2009). Statistics The Elements of Statistical Learning. In *Springer Series in Statistics*.  
<http://www.springerlink.com/index/D7X7KX6772HQ2135.pdf>

Hebertson, E. G., & Jenkins, M. J. (2003). Historic climate factors associated with major avalanche years on the Wasatch Plateau, Utah. *Cold Regions Science and Technology*, 37(3), 315–332.  
[https://doi.org/10.1016/S0165-232X\(03\)00073-9](https://doi.org/10.1016/S0165-232X(03)00073-9)

Heierli, J., Birkeland, K. W., Simenhois, R., & Gumbesch, P. (2011). Anticrack model for skier triggering of slab avalanches. *Cold Regions Science and Technology*, 65(3), 372–381.  
<https://doi.org/10.1016/j.coldregions.2010.10.008>

Heierli, J., Gumbesch, P., & Zaiser, M. (2008). Anticrack nucleation as triggering mechanism for snow slab avalanches. *Science*, 321(5886), 240–243. <https://doi.org/10.1126/science.1153948>

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

Hendrikx, J., Owens, I., Carran, W., & Carran, A. (2004). Overview of the spatial distribution of avalanche activity in relation to meteorological and topographic variables in an extreme maritime environment. *Proceedings of the International Snow Science Workshop*. Jackson Hole, Wyoming, USA.

Hendrikx, J., Owens, I., Carran, W., & Carran, A. (2005). Avalanche activity in an extreme maritime climate: The application of classification trees for forecasting. *Cold Regions Science and Technology*, 43(1–2), 104–116. <https://doi.org/10.1016/j.coldregions.2005.05.006>

Hincks, T. K., Aspinall, W. P., Sparks, R. S., Holcombe, E. A., & Kern, M. (2011). Landslide and avalanche hazards. In *Risk and Uncertainty Assessment for Natural Hazards* (Vol. 9781107006, Issue October 2021). <https://doi.org/10.1017/CBO9781139047562.010>

Hopfinger, E. J. (1983). Snow avalanche motion and related phenomena. *Annual Review of Fluid Mechanics*, 15, 47–76.

Hopfinger, E. J., & Tochon-Danguy, J.-C. (1977). A Model Study of Powder-Snow Avalanches. *Journal of Glaciology*, 19(81), 343–356. <https://doi.org/10.3189/s0022143000029373>

*Information Pyramid – EAWS.* (n.d.). Retrieved February 23, 2023, from <https://www.avalanches.org/standards/information-pyramid/>

Issler, D. (1998). Modelling of snow entrainment and deposition in powder-snow avalanches. *Annals of Glaciology*, 26, 253–258. <https://doi.org/https://doi.org/10.3189/1998AoG26-1-253-258>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*. Springer. <https://doi.org/10.1007/978-1-4614-7138->

Janeras Casanova, M., & Furdada Bellavista, G. (2002). Gestión del riesgo de aludes: metodología de análisis y planificación territorial y aplicación a l'Armiana de Canillo (Andorra) Snow avalanches risk management: method of analysis and land use planning and application to l'Armiana de Canillo (Andorra). *Estudios Recientes (2000-2002) En Geomorfología. Patrimonio, Montaña y Dinámica Territorial. VII Reunión Nacional de Geomorfología*, 99–108.

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.

Košová, V., Molokáč, M., Čech, V., & Jesenský, M. (2022). Avalanche Hazard Modelling within the Král'ova Hol'a Area in the Low Tatras Mountains in Slovakia. *Land*, 11(6). <https://doi.org/10.3390/land11060766>

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E*, 69(6), 66138.

Kumar, S., Srivastava, P. K., Snehmani, & Bhatiya, S. (2019). Geospatial probabilistic modelling for release area mapping of snow avalanches. *Cold Regions Science and Technology*, 165(September 2019), 102813. <https://doi.org/10.1016/j.coldregions.2019.102813>

Lang, M., Ouarda, T. B. M. J., & Bobée, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225(3–4), 103–117. [https://doi.org/10.1016/S0022-1694\(99\)00167-5](https://doi.org/10.1016/S0022-1694(99)00167-5)

Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U., & Zimmerli, M. (1998). A network of automatic weather and snow stations and supplementary model calculations providing

snowpack information for avalanche warning. *Proceedings of the International Snow Science Workshop “a Merging of Theory and Practice*, 225–233.

Lehning, M., Naaim, F., Naaim, M., Brabec, B., Doorschot, J., Durand, Y., Guyomarc'H, G., Michaux, J. L., & Zimmerli, M. (2002). Snow drift: Acoustic sensors for avalanche warning and research. *Natural Hazards and Earth System Sciences*, 2(3–4), 121–128. <https://doi.org/10.5194/nhess-2-121-2002>

Li, X., Sovilla, B., Jiang, C., & Gaume, J. (2021). Three-dimensional and real-scale modeling of flow regimes in dense snow avalanches. *Landslides*, 18(10), 3393–3406. <https://doi.org/10.1007/s10346-021-01692-8>

*Long-term statistics - SLF*. (n.d.). Retrieved February 22, 2023, from <https://www.slf.ch/en/avalanches/destructive-avalanches-and-avalanche-accidents/long-term-statistics.html>

Maggioni, M., Gruber, U., & Stoffel, A. (2002). Definition and characterisation of potential avalanche release areas. *ESRI International User Conference*, 11, in press. <http://www.esri.com/library/userconf/proc01/professional/papers/pap964/p964.htm>

*Mapa de Cobertes del Sòl d'Andorra* (2012). (n.d.). Retrieved July 8, 2023, from <https://www.iea.ad/mapa-de-cobertes-del-sol-d-andorra-2012>

*Mapa Forestal del Principat d'Andorra*. (n.d.). Retrieved July 8, 2023, from <https://www.iea.ad/mapa-forestal-del-principat-d-andorra>

Margalef, A., Pons, M., Apodaka, J., Vilella, M., Molné, T., Martín, G., & Gorospe, P. (2018). ALLAUS.AD: A new web-based platform for avalanche information and dissemination in Andorra. *International Snow Science Workshop, Innsbruck, Austria*, 1545–1547.

Masek, J. G., Wulder, M. A., Markham, B., McCorkel, J., Crawford, C. J., Storey, J., & Jenstrom, D. T. (2020). Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, 248(June), 111968. <https://doi.org/10.1016/j.rse.2020.111968>

Mast, C. M., Arduino, P., Miller, G. R., & Mackenzie-Helnwein, P. (2014). Avalanche and landslide simulation using the material point method: flow dynamics and force interaction with

- structures. *Computational Geosciences*, 18(5), 817–830. <https://doi.org/10.1007/s10596-014-9428-9>
- McClung, D. M. (1979). Shear fracture precipitated by strain softening as a mechanism of dry slab avalanche release. *Journal of Geophysical Research: Solid Earth*, 84(B7), 3519–3526.
- McClung, D. M. (1996). Effects of temperature on fracture in dry slab avalanche release. *Journal of Geophysical Research: Solid Earth*, 101(B10), 21907–21920. <https://doi.org/10.1029/95jb03114>
- McClung, D. M. (2000). Predictions in avalanche forecasting. *Annals of Glaciology*, 31(1993), 377–381. <https://doi.org/10.3189/172756400781820507>
- McClung, D. M. (2013). Effects of triggering mechanism on snow avalanche slope angles and slab depths from field data. *Natural Hazards*, 69(3), 1721–1731. <https://doi.org/10.1007/s11069-013-0771-2>
- McClung, D., & Schaefer, P. A. (1993). *The Avalanche Handbook*. Mountaineers Books.
- McKinney, W. (2010). Data Structures for Statistical Computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 56–61). <https://doi.org/10.25080/Majora-92bf1922-00a>
- Meister, R. (1989). Influence of strong winds on snow distribution and avalanche activity. *Annals of Glaciology*, 13, 195–201. <https://doi.org/10.1017/s0260305500007886>
- Michaux, J. L., Naaim-Bouvet, F., Naaim, M., Lehning, M., & Guyomarc'H, G. (2002). Effect of unsteady wind on drifting snow: first investigations. *Natural Hazards and Earth System Sciences*, 2(3–4), 129–136. <https://doi.org/10.5194/nhess-2-129-2002>
- Model Digital d'Elevacions. (n.d.). Retrieved July 8, 2023, from <https://www.iea.ad/model-digital-d-elevacions>
- Mohammed, R., & Scholz, M. (2023). Quality control and homogeneity analysis of precipitation time series in the climatic region of Iraq. *Atmosphere*, 14(2), 1–12. <https://doi.org/10.3390/atmos14020197>

Montgomery, D. R., & Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. *Water Resources Research*, 30(4), 1153–1171. <https://doi.org/10.1029/93WR02979>

Monti, F., Gaume, J., Van Herwijnen, A., & Schweizer, J. (2016). Snow instability evaluation: Calculating the skier-induced stress in a multi-layered snowpack. *Natural Hazards and Earth System Sciences*, 16(3), 775–788. <https://doi.org/10.5194/nhess-16-775-2016>

Munteanu, A., Nedelea, A., Comanescu, L., & Stoica, V. (2013). The Impact of Snow Avalanches on the Tourist Activities in the Upper Area of the Piatra Craiului Massif. *Journal of Environmental and Tourism Analysis*, 1(March), 21–29. <http://www.cabi.org/leisuretourism/FullTextPDF/2014/20143145949.pdf>

Naaim, M., Brugnot, G., & Charry, J. C. (1996). Avalanche modeling in France - Theory and applications. *International Snow Science Workshop, Banff, Canada*, 239–243.

Naaim, M., Naaim-Bouvet, F., Faug, T., & Bouchet, A. (2004). Dense snow avalanche modeling: Flow, erosion, deposition and obstacle effects. *Cold Regions Science and Technology*, 39(2–3), 193–204. <https://doi.org/10.1016/j.coldregions.2004.07.001>

Naaim-Bouvet, F., Naaim, M., & Faug, T. (2004). Dense and powder avalanches: Momentum reduction generated by a dam. *Annals of Glaciology*, 38, 373–378. <https://doi.org/10.3189/172756404781815185>

Okabe, A., Boots, B., Sugihara, K., & Chiu, S. N. (2009). *Spatial tessellations: concepts and applications of Voronoi diagrams*.

Oliver, M. A., & Webster, R. (1990). Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, 4(3), 313–332. <https://doi.org/10.1080/02693799008941549>

Oller, P., Janeras, M., de Buen, H., Arnó, G., Christen, M., García, C., & Martínez, P. (2010). Using AVAL-1D to simulate avalanches in the eastern Pyrenees. *Cold Regions Science and Technology*, 64(2), 190–198. <https://doi.org/10.1016/j.coldregions.2010.08.011>

Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have

arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157–175. <https://doi.org/10.1080/14786440009463897>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85), 2825–2830. <http://jmlr.org/papers/v12/pedregosa11a.html>

Peitzsch, E. H., Hendrikx, J., Fagre, D. B., & Reardon, B. (2012). Examining spring wet slab and glide avalanche occurrence along the Going-to-the-Sun Road corridor, Glacier National Park, Montana, USA. *Cold Regions Science and Technology*, 78, 73–81. <https://doi.org/10.1016/j.coldregions.2012.01.012>

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullet, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szenthimrey, T., Salinger, J., Førland, E. J., Hanssen-Bauer, I., Alexandersson, H., ... Parker, D. (1998). Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology*, 18(13), 1493–1517. [https://doi.org/10.1002/\(SICI\)1097-0088\(19981115\)18:13<1493::AID-JOC329>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1097-0088(19981115)18:13<1493::AID-JOC329>3.0.CO;2-T)

Pons-Pons, M., Johnson, P. A., Rosas-Casals, M., Sureda, B., & Jover, È. (2012). Modeling climate change effects on winter ski tourism in Andorra. *Climate Research*, 54(3), 197–207. <https://doi.org/10.3354/cr01117>

QGIS Development Team. (n.d.). *QGIS Geographic Information System*. <https://www.qgis.org>

Quinlan, J. R. (1993). Program for machine learning. *C4. 5.*

*Recognize and distinguish types of avalanche | LAB SNOW*. (n.d.). Retrieved February 22, 2023, from <https://www.ortovox.com/uk/safety-academy-lab-snow/01-avalanche-basics/avalanche-knowledge>

Redlands, C. E. S. R. I. (2011). *ArcGIS Desktop: Release 10*.

- Reiweger, I., Gaume, J., & Schweizer, J. (2015). A new mixed-mode failure criterion for weak snowpack layers. *Geophysical Research Letters*, 42(5), 1427–1432. <https://doi.org/http://dx.doi.org/10.1002/2014GL062780>
- Reuter, B., Richter, B., & Schweizer, J. (2016). Snow instability patterns at the scale of a small basin. *Journal of Geophysical Research: Earth Surface*, 121, 257–282. <https://doi.org/doi:10.1002/2015JF003700>
- Reuter, B., & Schweizer, J. (2012). The effect of surface warming on slab stiffness and the fracture behavior of snow. *Cold Regions Science and Technology*, 83–84, 30–36. <https://doi.org/10.1016/j.coldregions.2012.06.001>
- Riba Porras, S., García-Sellés, C., Bartelt, P., & Stoffel, L. (2018). Analysis of one avalanche zone in the Eastern Pyrenees (Val d' Aran) using historical analysis, snow-climate data and mixed flowing/powder avalanche modelling. *International Snow Science Workshop, Innsbruck, Austria*.
- Rosendahl, P. L., & Weißgraeber, P. (2020). Modeling snow slab avalanches caused by weak-layer failure - Part 2: Coupled mixed-mode criterion for skier-triggered anticracks. *Cryosphere*, 14(1), 131–145. <https://doi.org/10.5194/tc-14-131-2020>
- Salm, B., Burkard, A., & Gubler, H. U. (1990). *Berechnung von Fliesslawinen eine Anleitung fuer Praktiker mit Beispielen* (Issue 47).
- Saleem, S., Aslam, M., & Shaukat, M. R. (2021). A review and empirical comparison of univariate outlier detection methods. *Pakistan Journal of Statistics*, 37(4), 447–462.
- Sanders, D., Widera, L., & Ostermann, M. (2014). Two-layer scree snow-avalanche triggered by rockfall (Eastern Alps): Significance for sedimentology of scree slopes. *Sedimentology*, 61, 996–1030. <https://doi.org/10.1111/sed.12083>
- Sanz-Ramos, M., Bladé, E., Oller, P., Andrade, C., & Furdada, G. (2020). Role of friction terms in two-dimensional modelling of dense snow avalanches. *Natural Hazards and Earth System Sciences*, 20(March), e2019-423.
- Schweizer, J. (2002). Skier-triggered avalanches: observations and concepts. *Proceedings Snoskred Og Friluftsliv, Stryn, Norway, 24-27 October 2002, October*, 52–55.

Schweizer, J., & Camponovo, C. (2002). The temperature dependence of the effective elastic shear modulus of snow. *Cold Regions Science and Technology*, 35(1), 55–64. [https://doi.org/10.1016/S0165-232X\(02\)00030-7](https://doi.org/10.1016/S0165-232X(02)00030-7)

Schweizer, J., & Jamieson, J. B. (2008). *Dry-snow slab avalanche release revisited: shear vs. collapse?* 10(A), 10994.

Schweizer, J., Jamieson, J. B., & Schneebeli, M. (2003). Snow avalanche formation. *Reviews of Geophysics*, 41(4). <https://doi.org/10.1029/2002RG000123>

Schweizer, J., Reuter, B., Herwijnen, A. Van, Gauthier, D., & Jamieson, J. B. (2014). On how the tensile strength of the slab affects crack propagation propensity. *Proceedings, International Snow Science Workshop, Banff*, 164–168.

Schweizer, J., Reuter, B., van Herwijnen, A., & Gaume, J. (2016). Avalanche Release 101. *Proceedings of the International Snow Science Workshop, Breckenridge, CO*, 1–11. [http://www.slf.ch/ueber/mitarbeiter/homepages/schweizj/publications/Schweizer\\_etal\\_Avalanche\\_release\\_101\\_ISSW2016.pdf](http://www.slf.ch/ueber/mitarbeiter/homepages/schweizj/publications/Schweizer_etal_Avalanche_release_101_ISSW2016.pdf)

Searcy, J. K., & Hardison, C. H. (1960). Double-Mass Curves. In *WaterSupply Paper 1541B*. <http://dspace.udel.edu:8080/dspace/handle/19716/1592>

Selçuk, L. (2013). An avalanche hazard model for Bitlis Province, Turkey, using GIS based multicriteria decision analysis. *Turkish Journal of Earth Sciences*, 22(4), 523–535. <https://doi.org/10.3906/yer-1201-10>

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference, ACM 1968*, 517–524. <https://doi.org/10.1145/800186.810616>

Sigrist, C., & Schweizer, J. (2007). Critical energy release rates of weak snowpack layers determined in field experiments. *Geophysical Research Letters*, 34(3). <https://doi.org/10.1029/2006GL028576>

SLF WSL. (2022). RAMMS::AVALANCHE User Manual v1.8.0.

Součková, M., Juras, R., Dyrta, K., Moravec, V., Blöcher, J. R., & Hanel, M. (2022). What weather variables are important for wet and slab avalanches under a changing climate in a low-altitude

mountain range in Czechia? *Natural Hazards and Earth System Sciences*, 22(10), 3501–3525. <https://doi.org/10.5194/nhess-22-3501-2022>

Sovilla, B., McElwaine, J. N., & Louge, M. Y. (2015). The structure of powder snow avalanches. *Comptes Rendus Physique*, 16(1), 97–104. <https://doi.org/10.1016/j.crhy.2014.11.005>

Sovilla, B., Schaer, M., Kern, M., & Bartelt, P. (2008). Impact Pressures and flow regimes in dense snow avalanches observed at the Vallée de la Sionne test site. *Journal of Geophysical Research: Earth Surface*, 113(1), 1–14. <https://doi.org/10.1029/2006JF000688>

Stethem, C., Jamieson, B., Schaefer, P., Liverman, D., Germain, D., & Walker, S. (2003). Snow avalanche hazard in Canada - A review. *Natural Hazards*, 28(2–3), 487–515. <https://doi.org/10.1023/A:1022998512227>

Stoffel, A., Meister, R., & Schweizer, J. (1998). Spatial characteristics of avalanche activity in an Alpine valley - a GIS approach. *Annals of Glaciology*, 26, 329–336. <https://doi.org/10.3189/1998aog26-1-329-336>

Tarboron, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2), 309–319.

Techel, F., Jarry, F., Kronthaler, G., Mitterer, S., Nairz, P., Pavšek, M., Valt, M., & Darms, G. (2016). Avalanche fatalities in the European Alps: Long-term trends and statistics. *Geographica Helvetica*, 71(2), 147–159. <https://doi.org/10.5194/gh-71-147-2016>

*Tool SHALSTAB / SAGA-GIS Tool Library Documentation (v3.0.0)*. (n.d.). Retrieved July 7, 2023, from [https://saga-gis.sourceforge.io/saga\\_tool\\_doc/3.0.0/ta\\_slope\\_stability\\_2.html](https://saga-gis.sourceforge.io/saga_tool_doc/3.0.0/ta_slope_stability_2.html)

U.S. Geological Survey. (2019). Landsat 8 Data Users Handbook. In *USGS science for a changing world* (Vol. 8, Issue November). <https://landsat.usgs.gov/documents/Landsat8DataUsersHandbook.pdf>

van Herwijnen, A., & Heierli, J. (2009). Measurement of crack-face friction in collapsed weak snow layers. *Geophysical Research Letters*, 36(23), 1–5. <https://doi.org/10.1029/2009GL040389>

Van Herwijnen, A., & Jamieson, B. (2007). Snowpack properties associated with fracture initiation and propagation resulting in skier-triggered dry snow slab avalanches. *Cold Regions Science and Technology*, 50(1–3), 13–22. <https://doi.org/10.1016/j.coldregions.2007.02.004>

Van Rossum, G. (2023). The Python Language Reference: Expressions. In *Python Software Foundation*.

<https://docs.python.org/3/reference/%0Ahttps://docs.python.org/3/reference/expressions.html#id12>

Van Zadelhoff, F. B., Alaba, A., Cohen, D., Phillips, C., Schaeffli, B., Dorren, L., & Schwarz, M. (2022). Introducing SlideforMAP: a probabilistic finite slope approach for modelling shallow-landslide probability in forested situations. *Natural Hazards and Earth System Sciences*, 22(8), 2611–2635. <https://doi.org/10.5194/nhess-22-2611-2022>

Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, 185, 46–56. <https://doi.org/10.1016/j.rse.2016.04.008>

*Whitebox Geospatial*. (n.d.). Retrieved July 7, 2023, from <https://www.whiteboxgeo.com/>

Wilson, A., Schweizer, J., Johnston, C. D., & Jamieson, J. B. (1999). Effects of surface warming of a dry snowpack. *Cold Regions Science and Technology*, 30(1–3), 59–65. [https://doi.org/10.1016/S0165-232X\(99\)00014-2](https://doi.org/10.1016/S0165-232X(99)00014-2)

Wubalem, A. (2022). Landslide Inventory, Susceptibility, Hazard and Risk Mapping. In *Landslides*. IntechOpen. <https://doi.org/10.5772/intechopen.100504>

Zarrini, M. (2016). Calculation of Avalanche Flow Velocity and Its Run-out Distance by the Using Mathematical Modeling and Residual Stress. *International Journal of Applied and Computational Mathematics*, 2(2), 213–221. <https://doi.org/10.1007/s40819-015-0056-4>

Zhu, J., Zou, H., Rosset, S., & Hastie, T. (2009). Multi-class AdaBoost\*. *Statistics and Its Interface*, 2, 349–360.